A Focal Plane Processor for Continuous-Time 1-D Optical Correlation Applications

G. Liñán-Cembrano, L. Carranza, B. Alexandre, A. Rodríguez-Vázquez,
P. de la Fuente, T. Morlanes

Abstract This chapter describes a 1-D Focal Plane Processor which has been designed to run continuous-time optical correlation applications. The chip contains 200 sensory processing elements which acquire light patterns through a $2\text{mm} \times 10.9\mu\text{m}$ photodiode. The photogenerated current is scaled at the pixel level by five independent 3-bit programmable-gain current scaling blocks. The correlation patterns are defined as five sets of 200 3-bit numbers (from 0 to 7) which are provided to the chip via a standard I^2C interface. Correlation outputs are provided in current form through 8-bit Programmable Gain Amplifiers (PGA) whose configurations are also defined via I^2C. The chip contains a mounting alignment help which consists of 3 rows of 100 conventional Active Pixel Sensors (APS) inserted at the top, middle, and bottom part of the main photodiode array. The chip has been fabricated in a standard 0.35μm CMOS technology and its maximum power consumption is below 30mW. Experimental results demonstrate that the chip is able to process interference patterns moving at an equivalent frequency of up to 500kHz.

1 Introduction

This chapter presents an *Application Specific Focal Plane Processor* (ASFPP) with dedicated architecture, sensory front-end, computing resources and external interface. The chip has been developed in the framework of an Industrial R+D project whose aim is to design a One-Dimension (1D) programmable opto-electronic de-

G. Liñan-Cembrano, L. Carranza, B. Alexandre, A. Rodríguez-Vázquez
Instituto de Microelectrónica de Sevilla CNM-CSIC / Universidad de Sevilla,
Americo Vespucio s/n 41092 Seville (Spain), e-mail: linan@imse-cnm.csic.es,
P. de la Fuente, T. Morlanes
Fagor Aotek, S. Coop,
Paseo Torrebaso, 4 - Aptdo. Corr. 50, 20540, Eskoriatza, Guipúzcoa (Spain),
e-mail: tmorlanes@fagorautomation.es

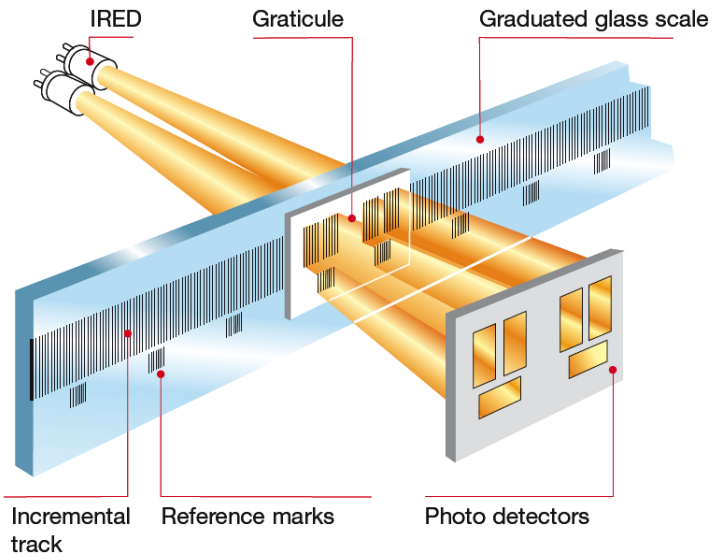


Fig. 1: Typical distribution of elements in an optical encoder (from [1]. Printed with permission)

vice able to acquire the light fringes produced in optical encoders applications [1], [2], [3], [4] and transform them into a set of electrical signals that can be employed to determine the relative, or absolute, movement of the object that the encoder is attached to.

Fig. 1 shows the typical distribution of elements in a transmission-type optical encoder [1]. Basically, four main blocks are identified. The first component is a near infrared[1] (NIR) LED which acts as light source. Light from this diode reaches a glass (second component) –also known as *scale*– which is fixed to the axis where the movement is to be detected. This glass contains a pattern of bars that allow/avoid the transmission of the light to the next element in the system. Third and fourth elements are mounted on the encoder's head, which is attached to the moving object. The third element is a regular scanning grid (again a pattern of bars for light transmission or oclusion) whose period is different from that on the *scale*. Finally, the fourth element is an opto-electronic device which must acquire the light fringes produced by either Moiré, Talbot, or Lau effects [5], [6], [7], and transform them into suitable length-measuring electrical signals.

Just for illustration purposes let us show a real example of how these fringes look like. In this example, the period of the fringes over the photodiodes is 509.5μm and, as one can easily see, what we obtain is not the typical ON/OFF pattern observed in simple relative encoders with only one scale – whose period is hundredths of

[1] λ =880nm in our case

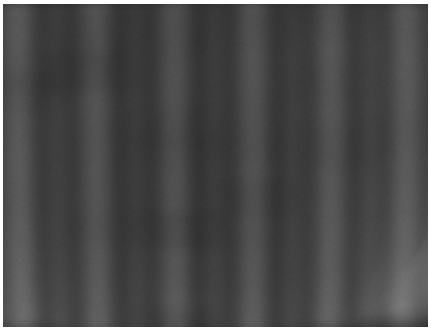


Fig. 2 Fringes pattern observed at the sensor's plane in a commercial encoder (acquired with a 640×480 7.2μm CCD camera). Notice that it is not the typical light ON/OFF pattern

micrometers – but a graded interference pattern instead. If we get slices of this image over the x axis (see Fig. 3 left) we observe the profile of the fringes along the axis where the movement is to be detected. This image has been captured with a relatively low-pitch (7.2μm) CCD camera in a real environment, this is why the lines (we are only plotting those corresponding to top and bottom rows of pixels) look so noisy. If we average pixel information along the y-axis and plot the resulting mean fringe pattern we obtain a cleaner view of how the fringes look like. Fig. 3 right shows this result [2]. One can easily identify a repeating pattern in the fringes whose

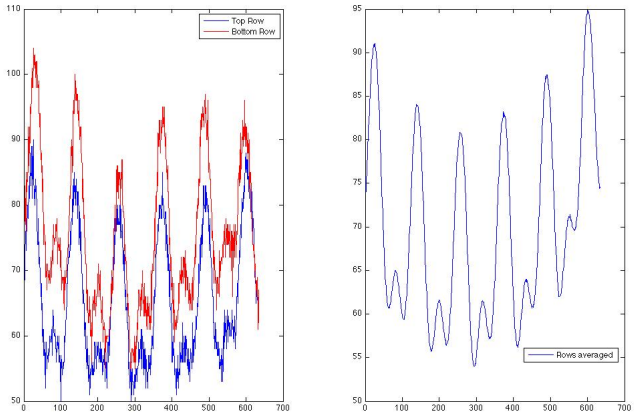


Fig. 3: (a) Fringe patterns observed in first and last row on the CCD. (b) Result from averaging over the y axis

period, for this particular application, is 509.5μm. We also observe the influence of

[2] Which is approximately equivalent –neglecting partial suppresion of reset noise due to averaging– to acquire the image with a 1D-CCD whose pixels have the same height as the array

using a single LED light source in the noticeable curvature appearing in the peaks of the fringes pattern. Obviously, when the object –with the encoder's head attached to it– suffers a displacement, these fringes also move on the focal plane[3] and, thus, the overall task of our optoelectronic device is to provide a few signals from which this movement can be measured precisely[4] . According to [4], the intensity of the light fringes produced over the detector's area along the x-axis can be compactly expressed as:

$$f(x,\theta) = A + B\sum_{n\geq 1}^{\infty} a_n \cos\left[\frac{2\pi n}{p}(x+\theta)\right] \tag{1}$$

where x is the axis of movement of the object, p is the period of the fringes, θ is the relative displacement between *scale* and *scanning* gratings[5], and the $\{a_n\}$ Fourier coefficients depend both on the optical and physical parameters of the system.
The goal in this application consists of producing two quadrature signals A and B:

$$\begin{aligned} A &= k_A \sin\left(\frac{2\pi}{p}\theta\right) \\ B &= k_B \cos\left(\frac{2\pi}{p}\theta\right) \end{aligned} \tag{2}$$

from which one can obtain the relative displacement θ using:

$$\theta = \frac{p}{2\pi}\arctan\left(\frac{k_B A}{k_A B}\right) \tag{3}$$

Obviously, the more accurate we are in the generation of A and B signals, the more precise will be the measurement of the displacement, since the better would be the interpolation over the Lissajous plot –which ideally should result in a perfect circle. However, generating precise quadrature signals from this pattern of fringes is far from being easy. As demonstrated in [3], [4], *sine* and *cosine* functions can be obtained from such light fringes pattern by using the following $g_A(x)$ and $g_B(x)$ *kernel* functions,

$$\begin{aligned} g_A &= \frac{1}{\pi x}\left[\sin\left(\frac{2\pi}{p_1}x\right) - \sin\left(\frac{2\pi}{p_2}x\right)\right] \\ g_B &= \frac{1}{\pi x}\left[\cos\left(\frac{2\pi}{p_1}x\right) - \cos\left(\frac{2\pi}{p_2}x\right)\right], \end{aligned} \tag{4}$$

where p_1 and p_2 are the periods of the two gratings. Then,

$$k_A \sin\left(\frac{2\pi}{p}\theta\right) = \int_{-\infty}^{+\infty} g_A(x) f(x,\theta) dx$$

[3] Indeed movement is amplified at the focal plane due to the optical setup of the system

[4] *Precisely* means errors below 1μm per meter in this case

[5] Which indeed corresponds to the displacement of the object

$$k_B \cos\left(\frac{2\pi}{p}\theta\right) = \int_{-\infty}^{+\infty} g_B(x) f(x,\theta) dx \tag{5}$$

where $f(x,\theta)$ is the light intensity of the fringes pattern on the focal plane. Kernel functions $g_A(x)$ and $g_B(x)$ are displayed in Fig. 4.

2 Description of the Operation and Design Specifications

Obviously, when mapping these equations into an electronic circuit one must consider different simplifications. First of all, we do not have the light intensity function $f(x,\theta)$ as input, but the current generated by discrete photodiodes which are uniformly distributed along the focal plane. Hence, each of these diodes is providing a kind of averaging of the impinging light intensity over its area of influence –which, in the simpler case, will correspond to its area. Furthermore, we will not implement functions $g_A(x)$ and $g_B(x)$ in their continuous form but using a discrete approximation –with an equivalent number of bits. Finally, for implementation and versatility purposes it is preferable to have separate outputs for the positive and negative parts of the A, and B quadrature signals. Thus, instead of $g_A(x)$ and $g_B(x)$, we will have $g_{A_+}(x)$, $g_{A_-}(x)$, $g_{B_+}(x)$, and $g_{B_-}(x)$, let us simply denote them by $g_k(x)$ where the k index specifies whether it is an A or B function and whether it is the positive or the

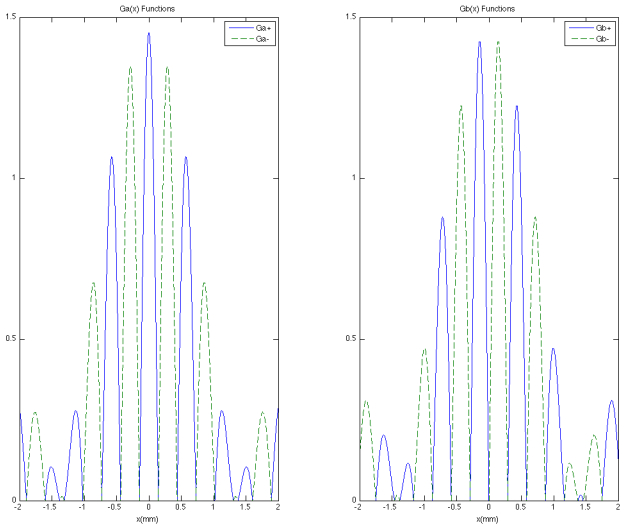


Fig. 4: $g_B(x)$ (left) and $g_A(x)$ signals (right) –negative part of the functions are sign-inverted and plotted in dashed line

negative part, $\{A_+, A_-, B_+, B_-\}$. Each of the outputs, O_k, of the chip is calculated as:

$$O_k(\theta) = \sum_{j=1}^{N_{pixels}} Iphoto_j(\theta) m_{k,j} \tag{6}$$

where each $m_{k,j}$ coefficient corresponds to the mean value of function $g_k(x)$ within the area of influence of diode j-th,

$$m_{k,j} = \int_{X_L}^{X_R} g_k(x) dx \tag{7}$$

In the chip, we have implemented a 3-bit representation of these coefficients –after an analysis where performance and area occupation were balanced– which results in the functions in Fig. 5. Besides, we have added the possibility to adjust the global gain of each output channel by means of a Current-Mode Programmable-Gain Amplifier (CM-PGA). Therefore, *Sine* and *Cosine* outputs are obtained as:

$$k_A \sin\left(\frac{2\pi}{p}\theta\right) = \alpha_1 O_{A_+} - \alpha_2 O_{A_-}$$

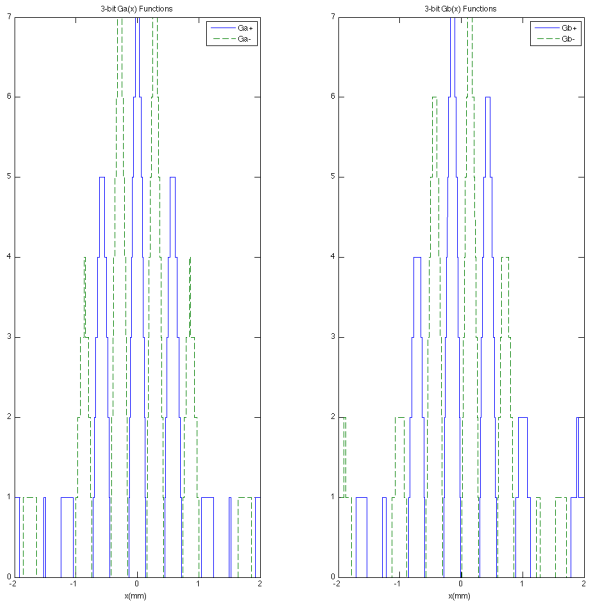


Fig. 5: 3-bit representation of $g_B(x)$ (left) and $g_A(x)$ signals (right) –negative part of the functions are sign-inverted and plotted in dashed line

$$k_B \cos\left(\frac{2\pi}{p}\theta\right) = \alpha_3 O_{B_+} - \alpha_4 O_{B_-} \tag{8}$$

Finally, the chip also provides a fifth output O_R calculated as,

$$O_R = \alpha_5 \sum_{j=1}^{N_{pixels}} Iphoto_j \cdot R_j; \qquad R_j\text{=[0, 1, 2, 3, 4, 5, 6, 7]} \tag{9}$$

which can be used to different purposes. Most commonly, it will be employed as a mechanism to obtain average illumination over the chip –by programming all R_j coefficients to the same value–, and adjust, in a feedback loop, the current through the NIR LED to guarantee that the amplitude of *Sine* and *Cosine* outputs remains within an appropriate margin. Another possible use could be as a way to read absolute reference positions in double-chip configurations on the same head, where one chip is given the task of incremental displacement measurements whereas the second chip performs readings of absolute marks.

2.1 Physical Information and System Requirements

The chip has been implemented using a 0.35μm 4M-2P technology available through Europractice. This technology offers *Nwell-Psubstrate* photodiodes with a sensitivity around 0.3 A/W@880nm. The following list summarizes the most relevant information and constraints for our design.

- Fringes period is **509.5μm**.
- Monochromatic light; **λ=880nm**.
- Incident light power at the focal plane between **[5.5, 55]μW/mm^2**. Hence, photogenerated current density will vary between **[1.65, 16.5]μA/mm^2**.
- Maximum frequency of fringes, due to head movement, at the focal plane is **500kHz** – head moving at **20m/s**.
- Fringe contrast is **12**%.
- Diodes pitch and fringes period must be relative prime numbers.
- Sensing part must allocate at least four periods of fringes – array lenght $\geq$2.04mm[6].
- Minimum diode's height is **1.5mm**.
- Power consumption below **100mW** – using 3.3V single power supply.
- Interfacing through **I^{2}C** 100kbps standard only.
- **Continuous time** operation – conventional reset-exposure-readout operation is not allowed.

[6] Using four periods of fringes and having relative prime numbers in diode's pitch and fringe period improves the quality of the interpolation process over the Lissajous plot when measuring the displacement.

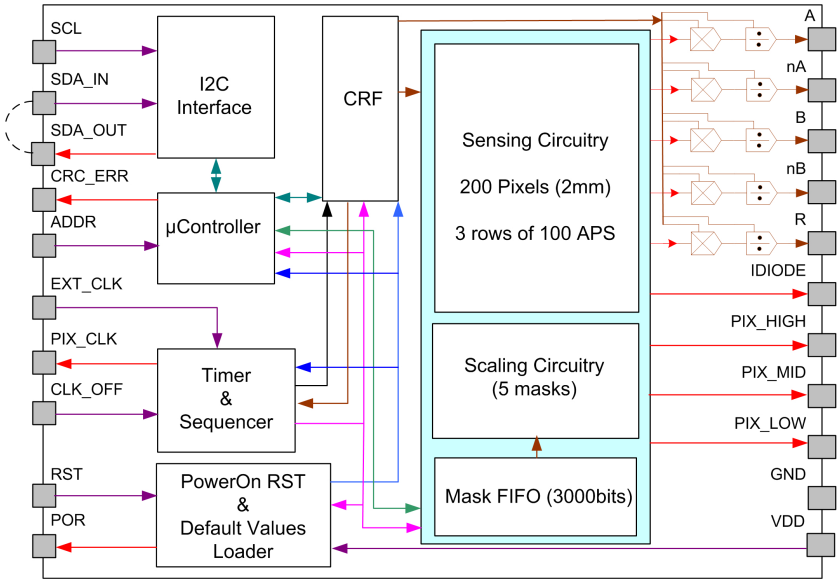


Fig. 6: Block Diagram of the Chip

3 Architecture of the Chip

Fig. 6 shows a simplified block diagram of the chip. As it has been mentioned, the main functionality of the chip is to provide five output currents,

$$O_k(t,\theta) = \alpha_k \cdot \sum_{j=1}^{N_{pixels}} Iphoto_j(t,\theta) \cdot m_{k,j}, \tag{10}$$

where $\{m_{k,j}\}$ are positive integer numbers in the range $[0,7]$ and the α_k coefficients are positive programmable gains which are defined as the quotient between any two 4-bit numbers,

$$\alpha_k = \frac{\sum_{n=0}^{3} 2^n}{\sum_{m=0}^{3} 2^m} \tag{11}$$

The chip includes the following main blocks:

- A standard I^2C interface which is the only mechanism to transmit commands and configuration setups to the chip. The chip address in the I^2C bus is determined by a built-in defined constant which is completed (LSB) with a bit provided through an external pin. Thus, two chips can be connected simultaneously to the same bus without conflicting. Readers with previous experience with the I^2C bus will notice that the chip uses SDA_IN and SDA_OUT lines (top right corner in the block diagram) instead of the conventional single SDA line. This is due to a lack of the proper bidirectional pad in the selected technology. SDA_IN uses a

conventional input pad whether SDA_OUT uses a conventional open-drain pull-down output pad. Both signals are connected together at the board to the SDA data line of the I^2C bus.
- A customized CISC microcontroller which implements 5 instructions –more details in Sect 4.
- A block of timers and prescalers which includes a 2-bit configurable (through binary divisions $\times 1$, $\times 1/2$, $\times 1/4$, $\times 1/8$) 55MHz oscillator[7].
- A Configuration Register File (CRF), which comprises 26 bytes, that defines the state of the different programmable modules of the chip.
- A Power-On-Reset and Bootloader unit which on the one hand, detect the events of power-up/down and executes a system reset during these events, and, on the other hand, loads the CRF with its by-default information.
- Five independent current-mode gain-programmable amplifiers (CM-PGA) which implement the α_k coefficient in (10).
- The main array, with 200 SPEs. Basically, each pixel contains a photodiode and five 3-bit programmable output branches of a current mirror. The array also contains, inserted at the top, bottom, and center, 3 rows of conventional APS sensors (100 pixels/row) that can be used during mounting of the head to help in the alignment process, or, in operation, to acquire the profile of the fringes being projected at a given time.
- A 3000-bit FIFO which stores the two hundred 3-bit coefficients that define each of the five – one per output – sets $\{m_{k,j}\}$.

4 Digital Part

The digital block of the chip has been designed around a specific purpose microcontroller which is in charge of chip control and configuration. Once programmed, the chip can operate autonomously as required in most applications. In addition, the controller contains access resources to the FIFO data memory, a set of timers and prescalers with sequencers plus a CRC calculation unit. The microcontroller architecture follows the CISC paradigm and incorporates five simple instructions which are summarized in Table 1.
The system only receives information through its I^2C interface and, as mentioned above, can be configured through an external pin to have two different positions in the bus, allowing two chips to be simultaneously connected to the same bus – which is very important in advanced heads containing two sensors. The design of the microcontroller has taken into account the analogue nature of the continuous time processing being performed by the system, this includes:

- To have low switching activity.

[7] Nominal frequency of the designed token ring oscillator; process corners, mismatching, power supply variations, and temperature affect this frequency which might move $\pm 50\%$

Table 1: Customized Microcontroller's Instruction Set

CMD	CODE	1^{st} Arg	2^{nd} Arg.	Data
WriCRF	0x01	[StartAddr]	N	{N bytes}
ReadCRF	0x02	[StartAddr]	N	{N bytes} read from chip
TIPS[a]	0x04			
WriFIFO	0x08			{K[b] bytes + CRC}
ReadFIFO	0x10			{CRC + 1000 bytes}

[a] Stands for Trigger Integrating Pixels Sequence.
[b] There are two modes of FIFO writing: (1) Complete transmission of FIFOLEN bytes –FIFOLEN is defined as a 16-bit number (in two CRF registers) which by default takes the value of 1000 (2) Marking the MSB of last byte to be transmitted with a logic 1. The specific mode during a transmission is defined in one of the registers in the CRF.

- To have maintain its performance when clocked with a low-precision frequency oscillator –internal token ring oscillator –, provided that its period remains between 12ns-140ns.
- To remain *almost* idle during normal operation of the chip. The only modules which stay active during the operation of the chip are the one which detects if the chip has been addressed after any I^2C start condition, and the sequencer of the Reset-Integrate-Readout process for the integration pixels – which indeed is not usually employed during normal use of the chip when making displacement measurements.

4.1 The Configuration Register File – CRF.

The Configuration Register File, or simply CRF, is a 26-byte memory composed of single write port and double read port 8-bit registers. The information stored in this register defines the status of all configurable options in the chip and it is written to the (safe) default value after a power-on reset or during a so-called *warm*[8] reset. The information stored in this register file can be divided into five groups of logic elements, namely:

1. The first group contains registers that inform about, the last command code executed by the microcontroller, FIFO data related information (the last datum written in the FIFO together with the last datum read from the FIFO), and the CRC value corresponding to the data stored within the FIFO. Although this information is not strictly necessary, it has been included for debugging and supervision purposes.
2. The second group contains the arguments that define the behavior of the FIFO accessing instructions.

[8] A reset commanded by the user by pulling down the **RST** pin of the chip

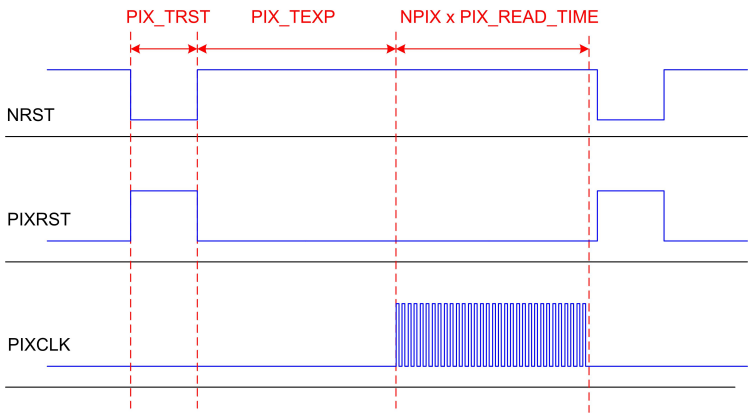


Fig. 7: Control Signals Generated by the APS Sequencer.

3. The third group stores the configuration that controls the APS Reset-Exposure-Read timing sequence.
4. The fourth group is a single register which controls (bit-masked) the operation of the analog core and the configuration of the built-in clock divider.
5. The fifth group, composed of five registers, defines the gain of each output channel $\{A_+, A_-, B_+, B_-, R\}$.

4.2 The APS Sequencer

The APS sequencer generates timing and control signals that command the Reset-Exposure-Readout cycle. Its operation can be cyclic, activating a specific bit in CRF4, or not cyclic using the special single command instruction *TIPS* (see Table. 1). The sequencer is fully programmable and completely idle when not in use. Fig. 7 shows the control signals generated by the sequencer. POINTRST initializes the circuitry that addresses the pixels during read time, while PIXRST is used to simultaneously initialize the photodiodes in all APS pixels. POINTRST and PIXRST are complementary signals and their duration is controlled by a programmable 16-bit timer (CRF5,CRF6), which consequently defines the reset time. PIXCLK drives the consecutive connection of each APS to its corresponding output node, defining the read time. The number of PIXCLK cycles is programmable (CRF9), and can be any number between 1 and 100. The duration of the PIXCLK cycle is programmable as well, and it is controlled by a 16-bit timer (CRF7, CRF8). The exposure time, defined as the time period between the end of reset time and the beginning of read time, is controlled by a programmable 24-bit timer (CRF10, CRF11, CRF12) . It is also possible to extend the these time intervals using a 16-bit prescaler (CRF13, CRF14), which can be activated if needed (by asserting a bit in CRF4). The se-

Table 2: The Configuration Register File

Position	Name	Description
0	LASTCMD	Last command received.
1	CRC	Last calculated CRC.
2	LASTREAD	Last byte read from the chip.
3	LASTWRI	Last byte written to the chip (no commands).
4	RCFP	Bit-masked configuration of the FIFO Write/Read and Integration Pixels. Some bits mask the operation of signals related to the operation of the integration pixels This includes whether to reset the pixels, whether to reset the pointer that addresses the pixels during readout, whether to activate the prescaler that defines the duration of reset time, exposure time, and readout time per pixel, and whether the Reset-Integrate-Readout process is to be executed continuously. Besides it also defines if FIFO write is controlled by the parameter FIFOLEN or by marking the last byte to be transmitted, whether FIFO readout (for test purposes) is destructive or not, and which clock has to be used during FIFO access operations (internal/SCL I^2C clock).
5	TRST_MSB	Definition of RST time for integration pixels (2-byte variable).
6	TRST_LSB	
7	TPIX_MSB	Definition of output time per pixel (2-byte variable).
8	TPIX_LSB	
9	LASTPIX	Definition of the position of the last integration pixel that must be read (100) by default.
10	TEXP_MSB	Definition of exposure time (3-byte variable).
11	TEXP_CSB	
12	TEXP_LSB	
13	PRESPIX_MSB	Definition of prescaler for integration pixel clock (2-byte variable).
14	PRESPIX_LSB	
15	PRESFIFO_MSB	Definition of prescaler for the FIFO (3-byte variable).
16	PRESFIFO_CSB	
17	PRESFIFO_LSB	
18	FIFOLEN_MSB	Number of bytes to be written to the FIFO (2-byte variable; 1000 by default).
19	FIFOLEN_LSB	
20	RCNA	Bit masked configuration of analogue blocks in the SPE. One bit defines whether the analogue section of the chip is ON or OFF. A set of 5 bits define whether or not to activate the different additional functionalities on the pixel. Finally, two bits define the division (x1, x.5, x0.25, x0.125) to apply to the on-chip built-in oscillator.
21	GAIN_A	Gain of the CM-PGA in channel A_+ (by default it is set to 0xFF, which means that the implemented gain is 15/15).
22	GAIN_nA	Gain of the CM-PGA in channel A_- (by default it is set to 0xFF, which means that the implemented gain is 15/15).
23	GAIN_B	Gain of the CM-PGA in channel B_+ (by default it is set to 0xFF, which means that the implemented gain is 15/15).
24	GAIN_nB	Gain of the CM-PGA in channel B_- (by default it is set to 0xFF, which means that the implemented gain is 15/15).
25	GAIN_R	Gain of the CM-PGA in channel R (by default it is set to 0xFF, which means that the implemented gain is 15/15).

quencer includes a programmable bit mask option that allows the user to deactivate (masking) any of the control signals generated by its circuitry (POINTRST, PIXRST, and PIXCLK).

4.3 Accessing the FIFO

The microcontroller has configuration options which make the access to the correlation pattern FIFO memory more flexible. All the accessing options are available by programming the CRF registers adequately. The FIFO registers main clock (used during read and write operations) is selectable; the user can extract it from the I^2C Serial Clock or use an internal programmable 16-bit clock timer which employs the built-in token ring oscillator.
Write operations consist of storing the data transmitted by and external source in the FIFO. The data must always be followed by the corresponding CRC. Users can mark the End of Transmission (EOT) in two ways, on the one hand, specifying the number of FIFO registers to be send, on the other hand, marking the EOT by asserting the MSB in the last byte to be transmitted (both options allow partial or total FIFO write operations). To verify the integrity of the received data, a CRC calculation unit performs CRC calculation during FIFO write operations. After the reception of the last datum, the controller compares the received and computed values and informs about matching status through a dedicated pin. As described above, users can also read the computed CRC value by downloading the information in CRF1.
Read operations involve the transmission to an external receptor of the total or partial content of the FIFO memory. The number of FIFO registers to be read is specified in variable FIFOLEN (CRF18, CRF19), although the external receptor can interrupt the transmission at any time by creating an I^2C stop condition. Read operations can be either destructive or non-destructive (by asserting a bit in CRF4). In the former case, every read and transmitted datum is eliminated. In the latter, during read operations, the controller interconnects the input and output ports of the FIFO. In this configuration, the FIFO is arranged in a ring structure, therefore data are simply shifted circularly during read operations, therefore a complete non-destructive read operation of the FIFO memory leaves its registers unchanged. The microcontroller oversees the FIFO configuration to avoid writing operations while the FIFO is disposed in ring structure. Therefore, even if the user accidentally leaves the non-destructive FIFO read access option established, write operations can always be executed.

5 The Mixed-Signal Processing Core

The computing core of the chip is an 1-D array of 200 programmable Sensory Processing Elements (SPEs). These SPEs transform the incident light (fringes) into a

photo-generated current and scale this current to produce five independent versions of it (one per output channel). Scaling coefficients are integer numbers in the range [0-7] and are locally stored within each SPE in a 15-bit shift register. Registers in physically adjacent SPE's are connected in series (output from left-side to input of right side) in such a way that a 3000-bit (15×200) shift register is formed –the previously described FIFO –, thus making the process of programming the coefficients quite straightforward. The SPE includes different configurable modules –whose state is defined in CRF20 – that allow for optimizing power consumption and accuracy according to the needs of the application at a given time. Thus, for instance, frequency response of the system can be modified –at the expense of power consumption– to allow processing fast moving fringes (20m/s). In addition to the main array, the mixed-signal core of the chip contains 3 rows of 100 APS pixels inserted at the top, middle, and bottom of the main diodes array. Finally, the mixed-signal core also contains five *Current-Mode* 8-bit *Programmable-Gain Amplifiers* which generate the output of the chip as expressed in (10). The following subsections describe, in details, the different modules in this mixed-signal processing core.

5.1 The Sensory-Processing Element

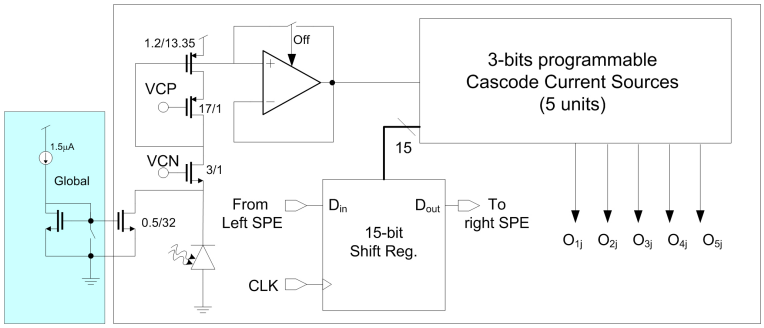


Fig. 8: Block Diagram of the SPE

Fig. 8 shows a block diagram –including a transistor-level representation of the Current-to-Voltage conversion unit– of the SPE. The blue-shaded area corresponds to a biasing unit which is shared by all SPEs in the chip and that is located at the periphery of the array. All biasing currents in the chip are obtained as scaled-up(/down) versions of a single 15μA source which is generated by an internal band-gap circuit, thus, the 1.5μA external source in Fig. 8 is obtained from the 15μA source and a $\times 10$ divider. Each SPE contains the following blocks

- A Nwell-Psubstrate photodiode that transforms incident light into a photogenerated current.

- A re-configurable current to voltage conversion unit which transforms this photogenerated current into a voltage level.
- An analog buffer which transmits this voltage to a bank of five 3-bit programmable current sources.
- Five 3-bit programmable current sources which receive an input voltage from the analog buffer and transform it into five independent output currents.
- A 15-bit shift-register which stores the values of each of the 3-bit numbers that define the scaling factor that the SPE will apply on each of its five output currents.

Basically –leaving aside the optional features in some operations– the SPE operates as follows:

1. During the programming phase, shift registers in all SPEs are connected in series (receiving data from the left neighbor and providing data to the right neighbor) to form a 3000-bit shift register. Once the programming stream has been loaded into the array (through 3000 clock cycles) each SPE register contains five 3-bit numbers which define its scaling coefficients in (10).
2. The photodiode creates a photogenerated current which is –approximately– proportional to the power of the incident light.
3. The information stored in the shift register is automatically driven –its is wired[9]– to the programmable current sources.
4. This photogenerated current is transformed into a voltage by the input stage of a cascode PMOS current mirror, and copied, by the analog buffer, to the input node of 5 identical programmable cascode current sources. These 3-bit programmable current sources are designed as seven unitary elements with *common centroid* layout configuration –also including dummy elements to improve the matching–, in such a way that the disposition (from left-to-right) is Dummy-b_2-b_1-b_2-b_0-b_2-b_1-b_2-Dummy. Since we are using current mode outputs, we get the summation in (10) simply by connecting outputs in different SPEs to the same low-impedance node. This node, which is indeed the input stage of the CM-PGA in each correlation output channel of the chip, is described in Sect. 5.3.

The following subsections describe in more details the main subsystems in the SPE.

5.1.1 The Sensory Block

The sensory block is, obviously, one of the most important elements in the SPE. This block is formed by the photodiode which senses the light and the analogue circuitry

[9] Due to this direct connection, we could get big current peaks during programming since we are moving all bits in the shift register every clock cycle. In order to avoid this, the analogue part of the chip can be switched-off during the programming phase by asserting a particular bit in CRF20. Indeed, by default –i.e., after power-on or reset–, this bit is set to 0, to avoid any kind of trouble with this issue, and the user is always requested to activate the analogue part of the chip in order to get some current through its outputs. This option does not switch off the 3 rows of APS pixels, neither their output amplifiers, thus allowing to get information about correct positioning of the chip during the mounting of the head without needing to activate the five correlation outputs

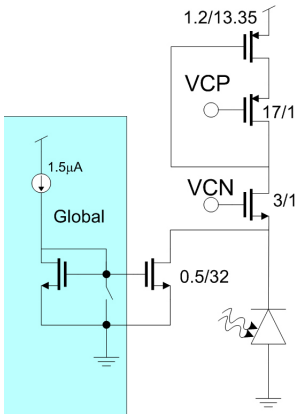


Fig. 9: Schematic of the Sensory Block

which transforms this current into a voltage which is suitable to be transmitted to the programmable current sources.

The sensory block, illustrated in Fig. 9, consists of:

- A 2000x10.9μm^2 Nwell-Psubstrate photodiode which provides the photogenerated current. According to the sensitivity value provided by the factory, and the expected incident power at the focal plane (see Sect. 2.1), expected photogenerated current will be in the range of [35-350]nA. The layout of this large photodiode includes contacts to the Nwell every 10μm, to reduce transit time of photogenerated carriers from the place where they are created to the place where they are collected. In addition to that, left and right (long) sides of the diode incorporate substrate contacts every 20μm as well, placed in such a way that substrate contact on one side, Nwell contact, and substrate contact on the other side are in a zig-zag disposition.
- A NMOS transistor, which keeps the reverse biasing of the photodiode to an almost constant voltage independently of the amount of photogenerated current. This transistor also serves to speed improvement purposes since the effect of the big parasitic capacitor of the photodiode over the frequency response of the system is largely attenuated by the cascading effect.
- The input state of a cascaded PMOS current mirror, which transforms the photogenerated current into a voltage. Unfortunately, the need for a continuous time operation avoids any possibility of using offset correction during photodiode's reset phases, then, one has to meet accuracy constraints by using large devices. However, making so big the transistor performing the current to voltage (*I*-*V*) conversion has a direct impact on its gate capacitance and therefore on the frequency response of the system. In order to overcome this limitation we introduced the next option in this block.

- A NMOS optional current source which can be added to the photogenerated current to improve the frequency response of the I-V block. This additional current will produce a shift in the location of the first pole of the system which instead of being proportional to $\sqrt{(I_{photo})}$ becomes proportional to $\sqrt{I_{photo}+I_{BIAS}}$. Obviously, since incident light power may vary within an order of magnitude, we may not require this block in the case of maximum illumination. Besides, adding this current degrades accuracy. First, it is evident that we must subtract the added offset current in a latter stage, and, of course, this subtraction is not error-free. Second, and not so evident, we must also consider mismatching dependence on the absolute current circulating through the mirror. We know that, neglecting output resistance effects, the mismatch in a simple mirror can be approximately expressed as[10]:

$$\frac{I_{out}}{I_{in}} = \frac{(\beta+0.5\Delta\beta)(V_{GE}+0.5\Delta V_{TH})^2}{(\beta-0.5\Delta\beta)(V_{GE}-0.5\Delta V_{TH})^2} \approx 1+\frac{\Delta\beta}{\beta}-\frac{2\Delta V_{TH}}{V_{GE}} \tag{12}$$

where V_{GE} is the *well-known* effective gate to source voltage[11]. As we see, there is a term which does not depend on the current through the mirror whereas there is another term which, via $V_{GE}=\sqrt{I/\beta}$, does depend on the current. Therefore, one can simply state that, for a given current mirror, while in saturation, matching improves as the current improves. However this is true for the total current through the mirror, and in our case, the signal is only a part of it. Therefore, we can find that the relative errors, defined as $\varepsilon=(I_{OUT}-I_{IN})/(I_{photo})$, are expressed as:

$$\begin{aligned} \varepsilon_{no\_IBIAS} &= \frac{\Delta\beta}{\beta}-2\Delta V_{TH}\sqrt{\frac{\beta}{I_{photo}}} \\ \varepsilon_{IBIAS} &= \frac{I_{BIAS}+I_{photo}}{I_{photo}}\frac{\Delta\beta}{\beta}-2\Delta V_{TH}\sqrt{\frac{\beta(I_{BIAS}+I_{photo})}{(I_{photo})^2}} \end{aligned} \tag{13}$$

hence, the error when adding I_{BIAS} to the photogenerated current is always larger and, therefore, one should only employ this extra current in those cases where the head is moving at top speed. Regarding the selection of a proper value for this offset current, we did a parametric analysis in which this current was varied between [0, 3]μA to find an optimum value. The result of this parametric analysis is shown in Fig. 10 where x-axis is the bias current and y-axis shows is the position of the first pole. According to this result, we have selected a *near*[12] to the peak value of 1.5μA which moves the first pole of the system to about 4.4MHz.

---

[10] we use the NMOS version for simplicity –i.e. not including V_{DD} in the equations

[11] $V_{GE}=V_{GS}-V_{TH}$

[12] Since degradation of performance beyond the optimum is quite abrupt –the PMOS input transistor of the mirror leaves the saturation region–, we preferred to move a little bit from the optimum value

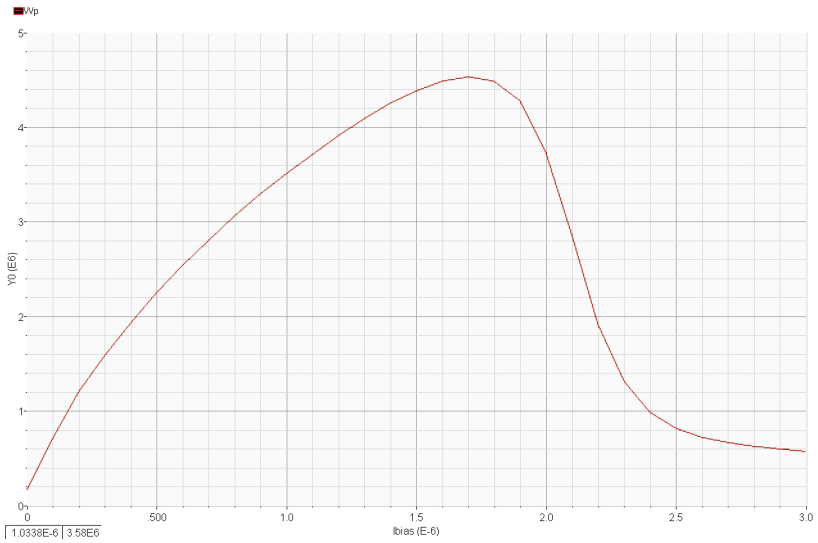


Fig. 10: Effect of I_{BIAS} level over the location of the first pole of the I-V block

As shown in Fig.8, we have also added a buffer inserted between the gate of the input transistor in the current mirror and the input of the 35 (7×5) programmable current sources. This buffer, which can be switched-off and bypassed when not required, has been added in order to reduce the capacitive load at the input node of the current mirror. Thus, instead of having 36 (35+1) equal transistors connected to this node, we only have 2 (the input transistor of the current mirror and the transistor at the positive input of the buffer). This buffer is a PMOS-input standard 5T Operational Transconductance Amplifier (OTA) which employs a 2.5μA bias current. Obviously, and similarly to the case of adding I_{BIAS}, the use of the buffer degrades accuracy performance due to the effect of its offset voltage[13].

5.1.2 The Current Scaling Block

The current scaling block provides the output current contribution of each SPE to (10). It consists of 35 identical cascode current source units (seven units per output) which also include the I_{BIAS} suppression circuitry. As described above, current sources in each output are laid out in a common centroid configuration (with dummies at both ends) in order to improve matching. Fig. 11 shows the schematic of one of this 3-bit programmable current sources.

[13] Indeed, this offset voltage plays the same role as a variation in the threshold voltage of input transistors in the programmable current sources.

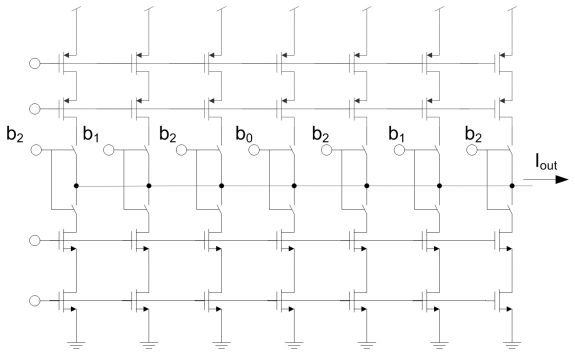


Fig. 11: Schematic of one of the programmable current sources in the SPE (transistor sizes as shown in Fig. 9)

5.1.3 The Memory Unit

The memory unit within the SPE is a simple 15-bit shift register which uses flip-flops from the available standard cells library. Its 15-bit parallel output (in parallel) drives the corresponding switches in the current scaling block. This memory unit also contains some clock buffering circuitry –end branch of the clock tree which is created for the whole array– in order to avoid any data corruption during shifting due to the use of very long clock wires with such huge (3000 registers) capacitive load.

5.2 Physical Details

As detailed in Sect. 5.1.1, the photodiode capturing fringes information has a pitch of 10.9μm (7.9 (active area) + 3 (separation between hot-wells)), and therefore, the pitch of the SPE should match this value. However, since standard cells have an height of 12μm in this technology, we opted for using a double-pitch layout for the SPE and locate SPEs both at the top and at the bottom of the photodiodes. Thus, every module within the SPE has been designed to match a pitch of 21.8μm. Obviously, with this configuration, odd SPEs have processing circuitry at one side of the array (bottom) whereas even SPEs have it at the other side (top). Consequently, the *bitstream* which is loaded into the chip to configure the different scaling coefficients must take this into account.
The processing part of each SPE is 1008,5μm height, with the occupation ratio detailed in Table 3.

Table 3: Area occupation per block within the SPE

Block	Height(μm)	(%)
Photodiode	2000	66.48
I-V	42	1.40
Current Scaling Blocks	5x102	16.95
Buffer	24	0.80
Test[1]	67.5	2.24
Registers	365	12.13

[1] The test circuitry consists of a switch which allows transmitting photodiode's current to a test-purpose output of the chip instead of to the input of the current mirror. Besides, the test block also contains a digital circuitry which acts as a pointer. This pointer selects the photodiode whose output is to be connected to the test pad. A reset pulse points to photodiode #0 (the leftmost device). Afterwards, consecutive clock pulses move this position to the right. In addition, another signal (a bit in CRF20) selects all diodes, only available in test mode, simultaneously, providing a fast mechanism to get total photogenerated current.

5.3 The Current-Mode Programmable Gain Amplifier

The chip provides its correlation outputs in current form through five Current-Mode Programmable Gain amplifiers (CMPGA). Each CMPGA must perform two important functions. First, it must accumulate current contributions from individual SPEs –implement the summation operation in (10). Second, it must scale this current up or down according to the gain programmed in the corresponding CRF register (CRF21-CRF25). Each function is implemented by a different subsystem, in the first case, accurate accumulation of SPEs current contributions is accomplished by a class-II current conveyor whereas its output -accumulated current– is scaled by a programmable current mirror, both subsystems are described in what that follows.

5.3.1 Accumulating the SPEs Contribution

The accumulation of the contribution of SPE's to the correlation output is accomplished by means of the virtual ground provided by a class-II current conveyor as shown in Fig. 12(a). The PMOS transistor and the amplifier are connected in a negative feedback look that maintains the voltage level at the input node independently of the the input current flowing through the transistor. Obviously, this simple description is far from what happens in practice, where one needs to consider the real input impedance at this virtual ground, and the output impedance of all current sources connected to it in order to extract useful design equations. Let us first consider the total output impedance of all (200) SPEs connected in a channel. Since we are using[14] cascaded current sources, the output conductance of the *k-th* SPE is:

[14] Assuming, for simplicity, that we are not using I_{BIAS}

$$G_o^k = \frac{g_{DS_p} \cdot g_{DScasc_p}}{g_{Mcasc_p}} \times m_k \tag{14}$$

where m_k is the scaling coefficient implemented by this SPE[15], and all other symbols are common in CMOS literature. Therefore, the total output impedance of all current sources in a correlation channel is simply:

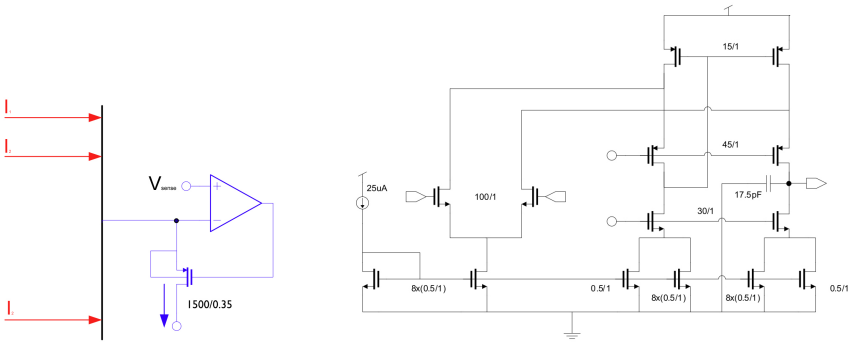
$$G_o = \sum_{k=1}^{k=200} G_o^k = \frac{g_{DS_p} \cdot g_{DScasc_p}}{g_{Mcasc_p}} \times \sum_{k=1}^{k=200} m_k \tag{15}$$

Since we are using a PMOS transistor in a negative feedback loop to collect current contributions from SPEs, we can define the error in current transmission as the difference between the current which is ideally provided by SPEs – let us denote it by I_{in} and the current flowing through the transistor in the feedback loop – I_{out}. After simple calculations one finds that:

$$I_{out} \approx I_{in} \times (1-\varepsilon) \qquad \text{with} \qquad \varepsilon = \frac{G_o}{g_{DS_{feedback}} + (A+1) g_{M\,feedback}} \tag{16}$$

where the *feedback* subindex refers to parameters of the PMOS transistor in the feedback loop, and we have assumed – which is indeed a design constraint – that:

$$g_{M\,feedback} \times g_{Mcasc_p} \times (A+1) \gg g_{DS_p} \times g_{DScasc_p} \times \sum_{k=1}^{k=200} m_k \tag{17}$$



(a) (b)

Fig. 12: Channel accumulation circuitry: (a) Current summation block. (b) Schematic of the amplifier

[15] Or, equivalently, the number of unitary current sources connected in parallel in this SPE to this accumulation node

5.3.2 Scaling the Accumulated Current

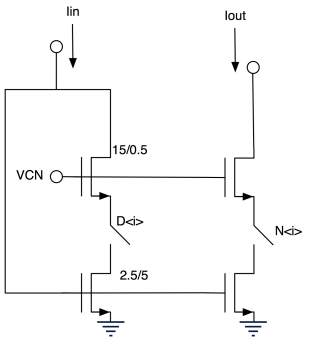


Fig. 13: Schematic of one-bit in the programmable gain current mirror. The programmable gain amplifier contains 16 of identical items in both input and output branches

The current flowing through the transistor in the feedback loop enters the input node of the Current-Mode Programmable Gain Stage. This unit – see Fig. 13 – is simply an all-NMOS current mirror with 16 identical input branches and 16 identical output branches. Clearly, the output current provided by this block is simply given by:

$$I_{Channel} = \frac{N}{D} \times I_{in} \tag{18}$$

where N is the number of active output units and D is the number of active input units – diode-configured transistors. N and D are configured by the user in CRF21-CRF25. There, each byte is divided in two octets (4-bit number) as $\{N_3N_2N_1N_0D_3D_2D_1D_0\}$, with two important considerations:

- This current gain stage has been included to guarantee that the chip will provide a sufficient amount of current in cases of very poor illumination conditions. By design, the maximum current through each bit-element in the input stage is 10μA. Currents beyond this limit will produce a saturation in the output channel[16] Thus, for instance, if a channel is producing a maximum current of 100μA the user must program **D** to be 10 or greater. It is clear that this limitation imposes a maximum output current per channel of 150μA which is a design specification fixed since the beginning of the project.
- One can wonder what happens if the user programs **D** to 0. In this case, there is not input stage receiving the current from the current conveyor. Hence, the

[16] It will make voltage at the input node to go above the limit imposed by the amplifier (V_{sense}), producing an instability in the circuitry due to the continuous transition from cut-off to conduction of the transistor in the feedback loop.

input node would increase its voltage until producing the same instability as in [16]. In order to avoid this, and to provide an additional feature, the controller checks whether D=0 for any of the output channel gains, and, if true, bypasses the current scaling stage providing the output current as it is collected by the current conveyor – including sign inversion. This option allows us to evaluate the operation of the current scaling block and to read correlation output currents larger than 150μA[17].

6 Chip layout

Fig. 14 shows the layout of the chip. It occupies 3.4x4.9 mm^2 and has been fabricated in a 0.35μm 4Metal-2Poly technology. The fringe sampling diodes are the vertical structures in the middle of the plot, being also easily visible the central row of APS pixels and the digital subsystem at the left side of the chip. In order to ease chip mounting on the encoder's head, the chip only contains pads in left and right sides. The CM-PGA are the vertical structures at the right side of the chip. Thus, all digital pads are on the left side whereas all analog pads are on the right side of the chip. On the one hand, it reduces noise in the analog lines, and, on the other hand, it allows for a *cleaner* design of the board which will host the chip.

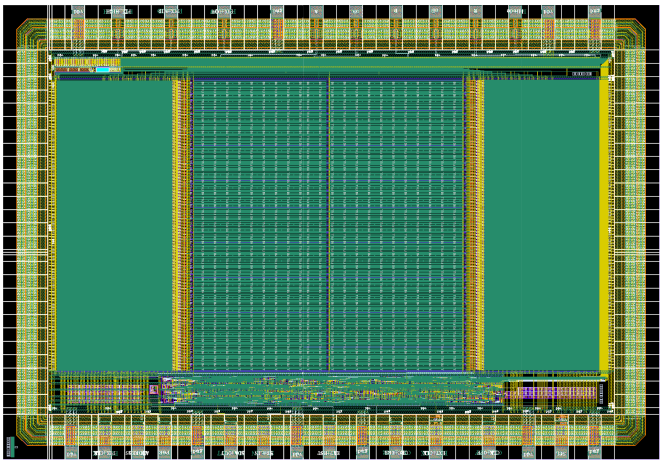


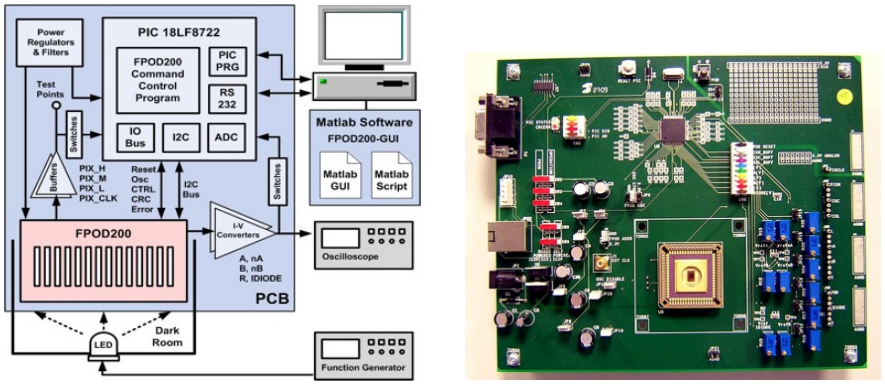
Fig. 14: Chip layout (rotated 90^o ccw)

[17] This situation is not very likely though. Considering that the average coefficient in each correlation channel is around 3, that the maximum expected photogenerated current is about 300nA (very optimistic supposition), and that we have 200 SPEs, the maximum expected output channel would be 180μA

7 Experimental Results

This section presents the experimental results obtained with the chip. All modules have been satisfactorily measured with experimental results meeting design specifications.

7.1 Test Setup



(a) (b)

Fig. 15: Chip Test (a) Block diagram of the test setup (b) The test board

Fig. 15 (a) shows a block diagram of our test setup. A control software, designed for MATLAB, communicates with the test board via RS232 protocol. Via this software we can command orders to the board controller – a 18LF8722 PIC, and, through this PIC, interact with the chip. Chip controls, CRF, and 3000-bit FIFO[18], are loaded into the chip by the PIC through its native I^2C interface. Analogue outputs of the chip can be read in two modes. In the first mode, a bank of switches connects all analogue output to different test points in the board. These test points are then read and digitized with an oscilloscope – which can be connected to MATLAB in the main computer as well. In the second mode, the bank of switches connects the analogue outputs of the chip to different analogue input channels on the PIC. Information is digitized by the ADC in the PIC and transmitted to the main computer via RS232. Using this second option is limited to DC – or low frequency – characterization measurements since the PIC only has one ADC which is time-multiplexed

[18] Though we can transmit byte by byte the information to be written in the FIFO from the PC via RS232 to the PIC and from the PIC to the chip, we have implemented a faster method by defining some preloaded FIFO configurations in the PIC memory so that one can simply select which configuration to write into the chip instead of transmitting it from the computer

when it is required to convert inputs from more than one channel.
Fig. 15 (b) shows the 4-layered test board designed to host the chip and the PIC. The chip is located inside the white square. Holes in the corners of this square are used to insert the screws that fix a black box which is placed on top of the chip during optical measurements. This black box contains on its top a 880nm LED whose current is modulated by a programmable function generator – we cannot produce fringes in this setup but modulated intensity patterns.

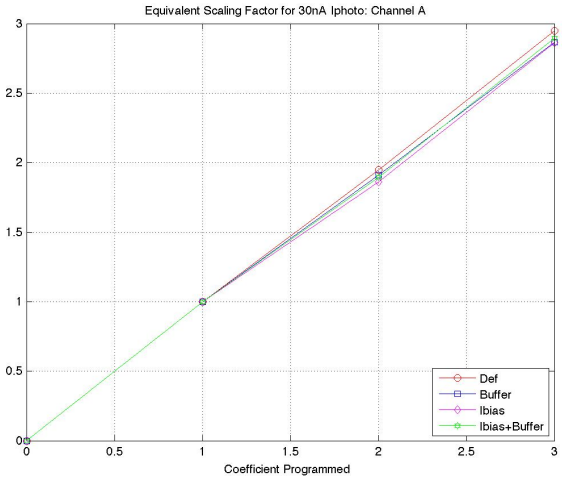
7.2 Scaling Coefficients Test: DC Response

Once we have checked that the digital subsystem operates correctly, we have evaluated the DC performance of the correlation channels. The test runs as follows:

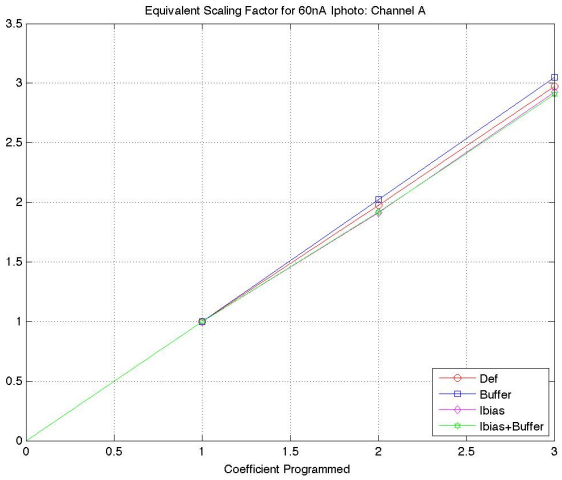
1. We configure the chip in test mode and enable all diodes simultaneously. With this configuration, we read the output current through the test pad, and vary the current through NIR LED until we get an equivalent current of 30nA[19] from each photodiode.
2. We load a FIFO stream (3000 bits) that configures coefficients in all SPEs to 1.
3. We measure the output current through all channels in four situations:
   a. In normal mode, i.e. not enabling the buffer neither the I_{BIAS} extra current.
   b. Enabling the buffer.
   c. Enabling the I_{BIAS} extra current.
   d. Enabling both the buffer and the I_{BIAS} extra current.
4. We re-write the FIFO by increasing the equivalent coefficient by 1 (while < 7) and repeat measurements
5. We go back to the first step, we increase the equivalent photogenerated current in 30nA (while $< 150nA$), and we repeat all measurements again.

Fig. 16 shows the equivalent gains obtained – we are displaying only one channel for visibility purposes – for photogenerated currents of 30nA and 60nA –worst cases regarding accuracy– , and in the four previously described operation modes.
Let us first comment that these results have been digitized by the DAC on the PIC. This DAC is a voltage mode DAC and, therefore, correlation output current from the chip has been converted to a voltage by means of a bank of programmable gain I-V converters – just a programmable resistor in a negative feedback loop around an operational amplifier. Due to noise in the board, and limitations imposed by the PIC DAC, our current mode LSB is about 150nA. Besides, in order to obtain the implemented coefficient, we are normalizing correlation output currents to the current obtained when we program all coefficients to 1 (this is why the coefficient 1 seems to be errorless). Results show that coefficients are satisfactorily implemented for a 3-bit representation of the information. Maximum obtained error is below 8%when

[19] Indeed we read 200×30nA=6μA



(a)



(b)

Fig. 16: Evaluation of equivalent gain for different photogenerated currents using the four modes of operation of the SPE. (1) Default (2) Enabling the buffer (3) Adding I_{BIAS} (4) Using the buffer and adding I_{BIAS}

we compute them by normalizing to the result of the implementation of coefficient one.
We can also compute the errors as the difference between the ideally produced output currents ($I_{photo} \times N_{pixels} \times Coeff$) vs. the measured ones. It is obvious that these errors would be bigger than those obtained when we normalized to the output current produced for coefficient one. However, we include them here to show overall deviation from ideality in the response of the chip. Fig. 17 shows these error computations (in %) for two extreme cases, the default configuration, in which we obtain the smallest error, and the result of adding I_{BIAS}, which produces the largest error[20]. Notice that we obtain errors that move between 5% and 8%, confirming that the required 3-bit implementation of the scaling coefficients is satisfactorily met.

7.3 Scaling Coefficients Test: Frequency Response

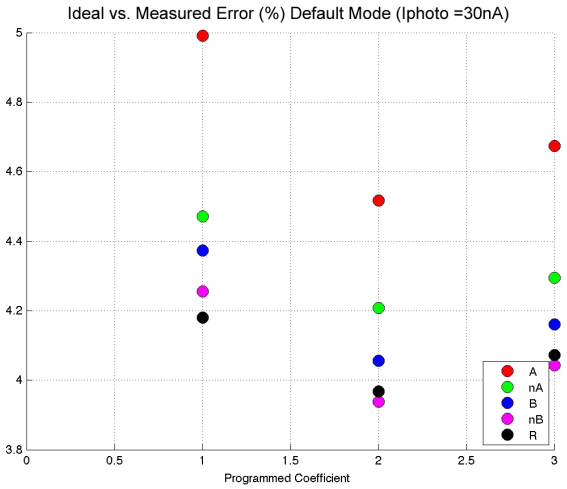
Frequency response of the correlation channels has been characterized using a 4-channels@500MHz Tektronix 3054DPO oscilloscope. The test works as follows:

1. We configure the chip in test mode and enable all diodes simultaneously – by enabling two bits in CRF20. With this configuration, we read the output current through the test pad, and vary the current through the photodiode until we get an equivalent sine current of 30nA with an optical contrast of 12% from each photodiode.
2. We load a FIFO stream (3000-bit) that configures all coefficients in all SPEs to 7[21].
3. We measure the output current through all channels in four situations at a very low frequency (200Hz):
   a. In normal mode, i.e. not enabling the buffer neither the I_{BIAS} extra current.
   b. Enabling the buffer.
   c. Enabling the I_{BIAS} extra current.
   d. Enabling both the buffer and the I_{BIAS} extra current.
4. We increase the frequency (while $<$ 5MHz) and repeat the measurements.
5. We find the -3dB frequencies for the different modes.

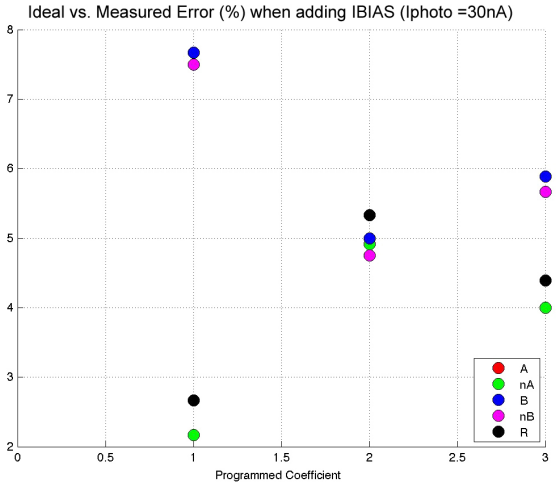
Table 4 shows averaged results of the cut frequency for different samples of the chip. Results show that, if properly configured, the chip can operate with frequency fringes moving at 500kHz (20m/s).

[20] Surprisingly, when we use the buffer and I_{BIAS}, the resulting error is smaller due to their different signs. Indeed, adding I_{BIAS} introduces a small systematic offset – due to non-total suppression of the I_{BIAS} in the SPE output current – which somehow compensates the small systematic – obviously not the random, but we observe averaged results since we measure the current from 200SPEs – component of the offset voltage of the buffer. Besides, this compensation is observed independently of the implemented coefficient since both terms scale as a function of the number of current sources connected to the SPE output node.

[21] This is the worst case since we are programming the maximum capacitive load



(a)



(b)

Fig. 17: Error computations for all output channels. (a) Default Mode – lowest error – (b) Adding I_{BIAS} –largest error

Table 4: Frequency response of the correlation channels in the different configuration modes

I_{BIAS}	Buffer	-3dB Frequency (kHz)
OFF	OFF	40
ON	OFF	140
OFF	ON	130
ON	ON	2700

Acknowledgements The authors would like to thank Dr. E. Roca from IMSE-CNM for her useful comments during pixel design. This work has been partially funded by CICE/JA, MICINN, and CDTI (Spain) through projects 2006-TIC-2352, TEC2009-11812, and Cenit EeE.

References

1. Fagor encoders catalog. URL http://www.fagorautomation.com/pub/doc/File/Catalogos/ingl/cat_captacion_general.pdf
2. D. Crespo, P. Alonso, T. Morlanes, E. Bernabeu, Optical Engineering **39**, 817 (2000)
3. D. Crespo, Nuevas herramientas aplicadas a la codificación óptica. Ph.D. thesis, Univ. Complutense de Madrid (2001)
4. T. Morlanes. Optical length measuring device with optoelectronic arrangement of photodetectors. European Patent Specification EP1164359B1
5. J. Tu, L. Zhan, Optics Communications **82**(3-4), 229 (1991)
6. H.F. Talbot, Philos. Mag. **9** (1836)
7. L. Liu, Appl. Opt. **28**, 4668 (1989)