



HAL
open science

A SIMD Programmable Vision Chip with High Speed Focal Plane Image Processing

Dominique Ginhac, Jérôme Dubois, Michel Paindavoine, Barthélémy Heyrman

► **To cite this version:**

Dominique Ginhac, Jérôme Dubois, Michel Paindavoine, Barthélémy Heyrman. A SIMD Programmable Vision Chip with High Speed Focal Plane Image Processing. EURASIP Journal on Embedded Systems, 2009, pp.13. 10.1155/2008/961315 . hal-00517911

HAL Id: hal-00517911

<https://hal.science/hal-00517911>

Submitted on 15 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A SIMD Programmable Vision Chip with High Speed Focal Plane Image Processing

Dominique Ginjac, Jérôme Dubois, Michel Paindavoine, and Barthélémy Heyrman

Abstract—A high speed analog VLSI image acquisition and low-level image processing system is presented. The architecture of the chip is based on a dynamically reconfigurable SIMD processor array. The chip features a massively parallel architecture enabling the computation of programmable mask-based image processing in each pixel. Extraction of spatial gradients and convolutions such as Sobel operators are implemented on the circuit. Each pixel include a photodiode, an amplifier, two storage capacitors, and an analog arithmetic unit based on a four-quadrant multiplier architecture. A 64×64 pixel proof-of-concept chip was fabricated in a $0.35 \mu\text{m}$ standard CMOS process, with a pixel size of $35 \mu\text{m} \times 35 \mu\text{m}$. A dedicated embedded platform including FPGA and ADCs has also been designed to evaluate the vision chip. The chip can capture raw images up to 10 000 frames per second and runs low-level image processing at a framerate of 2 000 to 5 000 frames per second.

Index Terms—CMOS Image Sensor, Parallel architecture, SIMD, High-speed image processing, Analog arithmetic unit.

I. INTRODUCTION

TODAY, digital cameras are rapidly becoming ubiquitous, due to reduced costs and increasing demands of multimedia applications. Improvements in the growing digital imaging world continue to be made with two main image sensor technologies: charge coupled devices (CCD) and CMOS sensors. Historically, CCDs have been the dominant image-sensor technology. However, the continuous advances in CMOS technology for processors and DRAMs have made CMOS sensor arrays a viable alternative to the popular CCD sensors. This led to the adoption of CMOS image sensors in several high-volume products, such as webcams, mobile phones, PDAs for example. Furthermore, new recent technologies provide the ability to integrate complete CMOS imaging systems at focal plane, with analog to digital conversion, memory and processing [1]–[5]. By exploiting these advantages, innovative CMOS sensors have been developed and have demonstrated fabrication cost reduction, low power consumption, and size reduction of the camera [6]–[8].

The main advantage of CMOS image sensors is the flexibility to integrate processing down to the pixel level. As CMOS image sensors technologies scale to $0.18 \mu\text{m}$ processes and under, processing units can be realized at chip level (*system-on-chip* approach), at column level by dedicating processing elements to one or more columns, or at pixel-level by integrating a specific unit in each pixel or local of neighboring pixels. Most of the researches deals with chip

and column-level [9]–[12]. Indeed, pixel-level processing is generally dismissed because pixel sizes are often too large to be of practical use. However, as CMOS scales, integrating a processing element at each pixel or group of neighboring pixels becomes more feasible since the area occupied by the pixel transistors decreases, leading to an acceptable small pixel size. A fundamental tradeoff must be made between three dependent and correlated variables: pixel size, processing element area, and fill-factor. This implies various points of view:

- 1) for a fixed fill-factor and a given processing element area, the pixel size is reduced with technology improvements. As a consequence, reducing pixel size increases spatial resolution for a fixed sensor die size.
- 2) for a fixed pixel size and a given processing element area, the photodiode area and the fill-factor increase as technology scales since the area occupied by the pixel transistors in each processing element decreases. It results in better sensibility, higher dynamic range and signal-to-noise ratio.
- 3) for a fixed pixel size and a given fill-factor, the processing element can integrate more functionalities since the transistors require less area as technology scales. Consequently, the image processing capabilities of the sensor increase.

In summary, each new technology process offers 1) to integrate more processing functions in a given silicon area, or 2) to integrate the same functionalities in a smaller silicon area. This can benefit the quality of imaging in terms of resolution, noise for example by integrating specific processing functions such as correlated double sampling [13], anti blooming [14], high dynamic range [15], and even all basic camera functions (color processing functions, color correction, white balance adjustment, gamma correction) onto the same camera-on-chip [16]. Furthermore, employing a processing element per pixel offers the ability to exploit the high speed imaging capabilities of the CMOS technology by achieving massively parallel computations [17]–[20].

In this paper, we discuss hardware implementation issues of a high speed CMOS imaging system embedding low-level image processing. For this purpose, we designed, fabricated, and tested a proof-of-concept 64×64 pixel CMOS analog sensor with per-pixel programmable processing element in a standard $0.35 \mu\text{m}$ double-poly quadruple-metal CMOS technology.

The rest of the paper is organized as follows. The Section II is dedicated to the description of the high speed algorithms embedded at pixel-level. The Section III is a general description

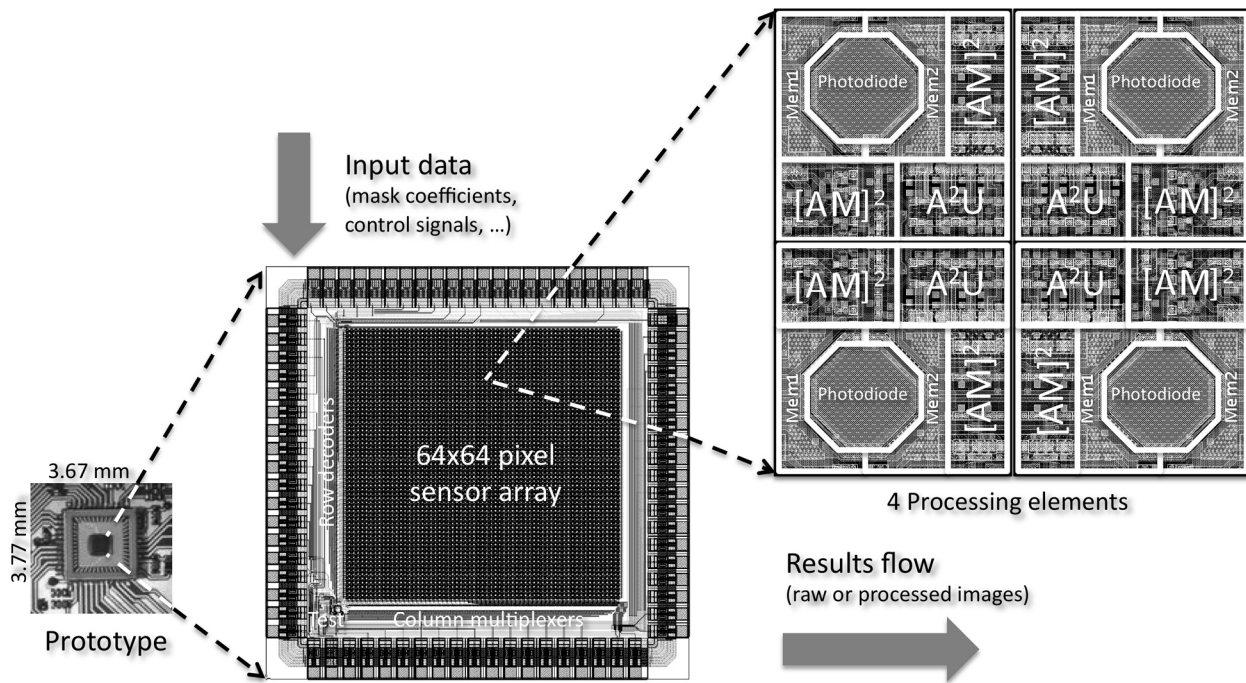


Fig. 1. Overview of the image sensor with a processor-per-pixel array

of the characteristics of the sensor. These characteristics are well detailed in the Section IV, which talks about the design of the circuit, with a full description of the main components such as the photodiode structure, the embedded analog memories, and the arithmetic unit are successively described. In the section V, we describe the test hardware platform and the chip characterization results, including an analysis of the fixed pattern noise. Finally, some experimental results of high speed image acquisition with pixel-level processing are provided in the last section of this paper.

II. HIGH SPEED FOCAL PLANE IMAGE-PROCESSING CAPABILITIES

In an increasingly digital world, the most part of imaging systems has become almost entirely digital, using only an analog-to-digital (ADC) between the sensor and the processing operators. However, low-level image processing usually involves basic operations using local masks. These local operations are spatially dependent on other pixels around the processed pixel. Since the same type of operations is applied to a very large data set, these low-level tasks are computationally intensive and require a high bandwidth between the image memory and the digital processor. In this case, an analog or a mixed-approach can offer superior performance leading to a smaller, faster, and lower power solution than a digital processor [21]. Low-level image processing tasks are inherently pixel-parallel in nature. Integrating a processing element within each pixel based on a single instruction multiple data (SIMD) architecture is a natural candidate to cope with the processing constraints [18]. This approach is quite interesting for several aspects. First, SIMD image-processing capabilities at focal plane have not been fully exploited because the silicon area available for the processing elements is very limited.

Nevertheless, this enables massively parallel computations allowing high framerates up to thousands of images per second. The parallel evaluation of the pixels by the SIMD operators leads to processing times, independent of the resolution of the sensor. In a classical system, in which low-level image processing is externally implemented after digitization, processing times are proportional to the resolution leading to lower framerates as resolution increases. Several papers have demonstrated the potentially outstanding performance of CMOS image sensors [22]–[24]. Krymski et al. [22] describe a high speed (500 frames/s) large format 1024×1024 Active Pixel Sensor (APS) with 1024 ADCs. Stevanovic et al. [23] describe a 256×256 APS which achieves more than 1000 frames/s with variable integration times. Kleinfelder et al. [24] describe a 352×288 Digital Pixel Sensor (DPS) with per pixel bit parallel ADC achieving 10,000 frames/s or 1 Giga-pixels/s.

Secondly, the high speed imaging capability of CMOS image sensors can benefit the implementation of new complex applications at standard rates and improve the performance of existing video applications such as motion vector estimation [25]–[27], multiple capture with dynamic range [28]–[30], motion capture [31], and pattern recognition [32]. Indeed, standard digital systems are unable to operate at high framerates, because of the high output data rate requirements for the sensor, the memory, and the processing elements. Integrating the memory and processing with the sensor on the same chip removes the classical input output bottleneck between the sensor and the external processors in charge of processing the pixel values. Indeed, the bandwidth of the communication between the sensor and the external processors is known as a crucial aspect, especially with high resolution sensors. In such cases, the sensor output data flow can be very high, and needs a lot of hardware resources to convert, process and

transmit a lot of information. So, integrating image processing at the sensor level can alleviate the high data rate problem because the pixel values are pre-processed on-chip by the SIMD operators before sending them to the external world via the communication channel. This will result in data reduction, which allows sending the data at lower data-rates, and reduces the effect of the computational-load bottleneck.

Thirdly, one of the main drawbacks to design specific circuits integrating sensing and processing on the same chip is that these vision chips are often built as special-purpose devices, performing specific and dedicated tasks, and not reusable in another context [33]. So, it can be widely beneficial to integrate a versatile device, whose functionality can be easily modified. Moreover, except the basic operations such as convolutions with small masks, the majority of computer vision algorithms requires the sequential execution of different successive low level image processing on the same data. So, each processing element must be built around a programmable execution unit, communication channels, and local memories dedicated to intermediate results. Because of the very limited silicon area, the processing units are necessarily very simple, providing the best compromise between various factors such as versatility, complexity, parallelism, processing speeds and resolution.

To sum up, the flexibility to integrate processing down to the pixel level allows us to rearchitect the entire imaging system to achieve much higher performances [34]. The key idea is 1) to capture images at a very high framerate, 2) to process the data on each pixel with a SIMD programmable architecture exploiting the high on-chip bandwidth between the sensor, the memory and the elementary processors and 3) to provide results at the best framerate depending on the complexity of the image processing. In this paper, we present our approach to the design of a massively parallel, SIMD vision chip based implementing low level image processing based on local masks. Our analog processing operators are fully programmable devices by dynamic reconfiguration, and can be viewed as a software-programmable image processor dedicated to low-level image processing. The main objectives of our design are: 1) to evaluate the potential for high speed snap-shot imaging and, in particular, to reach a 10 000 frames/s rate, 2) to demonstrate a versatile and reconfigurable processing unit at pixel-level, and 3) to provide an original platform for experimenting with low-level image processing algorithms that exploit high-speed imaging.

III. DESCRIPTION OF THE ARCHITECTURE

The proof-of-concept chip presented in this paper is depicted in Fig. 1. The core includes a two-dimensional array of 64×64 identical processing element (PE). It follows the single instruction multiple data (SIMD) computing paradigm. Each of the PE is able to convolve the pixel value issued from the photodiode by applying a set of mask coefficients to the image pixel values located in a small neighborhood. The key idea is that a global control unit can dynamically reconfigure the convolution kernel masks and then implements the most part of low-level image processing algorithms. This confers

the functionality of programmable processing devices to the PEs embedded in the circuit. Each individual PE includes the following elements:

- a photodiode dedicated to the optical acquisition of the visual information and the light-to-voltage transduction,
- two *Analog Memory, Amplifier and Multiplexer* structures called $[AM]^2$, which serve as intelligent pixel memories and are able to dissociate the acquisition of the current frame in the first memory and the processing of the previous frames in the second memory,
- an *Analog Arithmetic Unit* named A^2U based on four analog multipliers, which performs the linear combination of the four adjacent pixels using a 2×2 convolution kernel.

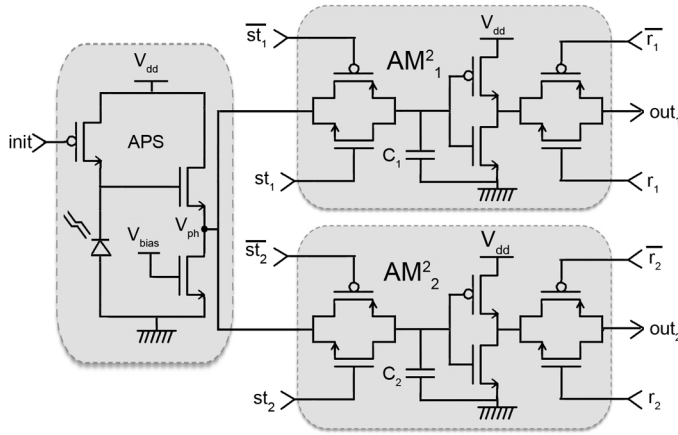
In brief, each PE includes 38 transistors integrating all the analog circuitry dedicated to the image processing algorithms. The global size of the PE is $35 \mu\text{m} \times 35 \mu\text{m}$ ($1225 \mu\text{m}^2$). The active area of the photodiode is $300 \mu\text{m}^2$, giving a fill-factor of 25 %. The chip has been realized in a standard $0.35 \mu\text{m}$ double-poly quadruple-metal CMOS technology and contains about 160 000 transistors on a $3.67 \text{ mm} \times 3.77 \text{ mm}$ die (13.83 mm^2). The chip also contains test structures on the bottom left of the chip. These structures are used for detailed characterization of the photodiodes and processing units.

IV. CIRCUIT DESIGN

A. Pixel Structure

Each pixel in the CMOS image sensor array consists of a photodiode and a processing unit dedicated to low-level image processing based on neighborhoods. In our chip, the type of photodiodes is one of the simplest photo element in CMOS image sensor technology. It consists of N-type photodiodes based on an n^+ -type diffusion in a p-type silicon substrate. The depletion region is formed in the neighborhood of the photodiode cathode. Optically generated photocarriers diffuse to neighboring junctions [35]. In order to achieve good performances, the photodiodes should be designed and optimized carefully, in order to minimize critical parameters such as the dark current and the spectral response [36]. The shape of photodiode layout, the structure of the photodiode, and the layout have significant influences on the performance of the whole imager [37], [38]. The active area of the photodiode absorbs the illumination energy and turns that energy into charge carriers. This active area must be large as possible in order to absorb a maximum of photons. In the mean time, the control circuitry required for the readout of the collected charges and the inter-element isolation area must be as small as possible in order to obtain the best fill factor. We have theoretically analyzed, designed and benchmarked different photodiodes shapes [39], and finally, an octagonal shape based on 45° structures was chosen (see Fig. 1). More details about the photodiodes design can be found in the aforementioned paper. Here, only the basic concept behind the photodiodes design has been briefly overviewed.

The second part of the pixel is the analog processing unit. Existing works on analog pixel-level image processing can be classified into two main categories. The first one is

Fig. 2. The $[AM]^2$ structure

intrapixel, in which processing is performed on the individual pixels in order to improve image quality, such as the classical Active Pixel Sensor or APS [9], [40]. The second category is interpixel, where the processing is dedicated to groups of pixels in order to perform some early vision processing and not merely to capture images. Our work clearly takes place in this category because our main objective is the implementation of various in-situ image processing using local neighborhoods. Based on this design concept, this forces a re-thinking of the spatial distribution of the processing resources, so that each computational unit can easily use a programmable neighborhood of pixels. For this purpose, the pixels are mirrored about the horizontal and the vertical axis in order to share the different *Analog Arithmetic Units* (A^2U). As example, a block of 2×2 pixels is depicted in Fig. 1. The main feature of its innovative distribution is to optimize the compactness of the metal interconnections with pixels, to contribute to a better fill factor, and to provide generality of high speed processing based on neighborhood of pixels.

B. Analog Memory, Amplifier and Multiplexer : $[AM]^2$

In order to increase the algorithmic possibilities of the architecture, the key point is the separation of the acquisition of the light inside the photodiode and the readout of the stored value at pixel-level [41]. Thus, the storage element should keep the output voltage of the previous frames whereas the sensor integrates photocurrent for a new frame. So, we have designed and implemented dedicated pixels including a light sensitive structure and two specific circuits called *Analog Memory, Amplifier, and Multiplexer* ($[AM]^2$), as shown in Fig. 2.

The system has five successive operation modes: reset, integration, storage, amplification, and readout. All these phases are externally controlled by global signals common to the full array of pixels. They all occur in parallel over the sensor (*snapshot* mode) in order to avoid any distortion due to a row-by-row reset. In each pixel, the photosensor is a N-type photodiode associated with a PMOS transistor reset. This switch resets the integrating node to the fixed voltage V_{dd} . The pixel array is held in the reset mode until the *init* signal raises, turning the PMOS transistor off. Then, the photodiode discharges for

a fixed period, according to the incidental luminous flow. The first NMOS transistor acts as a transconductance, producing the voltage V_{ph} , directly proportional to the incident light intensity. The integrated voltage is polarized around $V_{dd}/2$ by the second NMOS transistor. The calibration of the structure is ensured by the positive reference bias voltage ($V_{bias} = 1.35V$).

Following the acquisition stage, two identical subcircuits $[AM]^2_i$ (with $i = 1, 2$) take place to realize the storage phase of V_{ph} . Each $[AM]^2$ includes three pairs of NMOS and PMOS transistors and a capacitor which acts as an analog memory. The subcircuit $[AM]^2_i$ is selected when the st_i signal is turned on. Then, the associated analog switch is open allowing the integration of the photogenerated current in the corresponding C_i capacitor. Consequently, the capacitors are able to store the pixel value during the frame capture from one of the two switches. The capacitors are implemented with double-polysilicium. The size of the capacitors is as large as possible in order to respect the fill-factor and the pixel size requirements. The capacitors values are about 40 fF. They are able to store the pixel value for 20 ms with an error lower than 4 %. Behind the storage subcircuit, a basic CMOS inverter is integrated. This inverter serves as a linear high-gain amplifier since the pixel signal is polarized around $V_{dd}/2$. Finally, the last phase consists in the readout of the stored values in the capacitors C_i . The integrated voltage across the capacitor C_i can be readout on the output out_i through one of the two switches, controlled by the r_i signals.

Fig. 3 describes the experimental results of successive acquisitions in an individual pixel. The acquisitions occur when one of the two signals st_1 or st_2 goes high. The two combinations of acquisitions are presented in this example. After the first reset, st_2 is followed by st_1 whereas the inverted sequence of acquisitions is realized after the second reset. The signal V_{ph} gives the voltage corresponding to the incidental illumination on the pixel and the two outputs (out_1 and out_2) give the voltage stored in each of the capacitors when the associated readout signal raises.

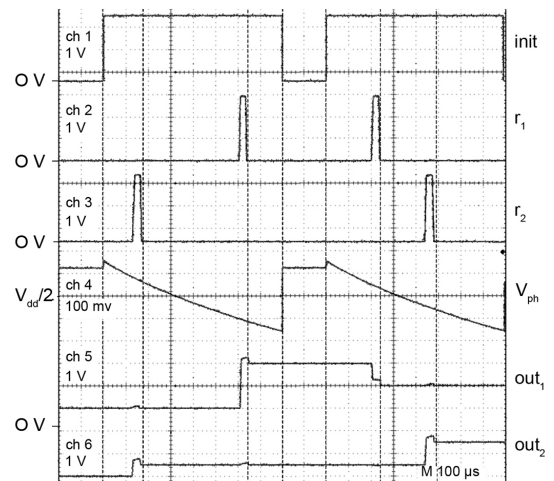


Fig. 3. High speed sequence capture with basic image processing

One of the main advantages of the two $[AM]^2$ structures is

that the capture sequence can be made in the first memory in parallel with a readout sequence and/or processing sequence of the previous image stored in the second memory, as shown in Fig. 4.

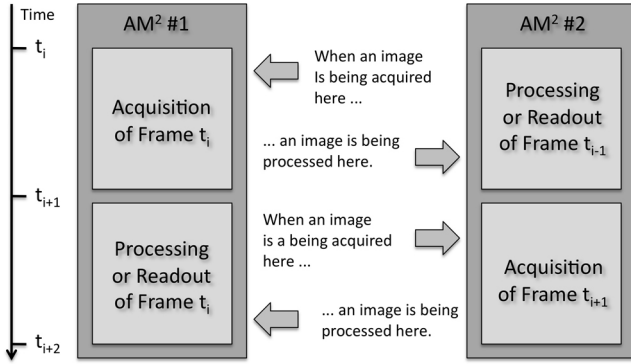


Fig. 4. Parallelism between capture sequence and readout sequence

Such a strategy has several advantages:

- 1) The framerate can be increased (up to 2x) without reducing the exposure time.
- 2) The image acquisition is time-decorrelated from image processing, implying that the architecture performance is always the highest, and the processing framerate is maximum,
- 3) A new image is always available without spending any integration time.

C. Analog Arithmetic Unit: A²U

When designing a pixel-level processing unit, you should consider adopting efficient strategies to minimize the silicon area occupied by the processor. To that end, we designed an analog arithmetic unit (A²U) which is able to perform convolution of the pixels with a 2x2 dynamic kernel. This unit is based on four-quadrant analog multipliers [42], [43] named M1, M2, M3, and M4, as illustrated in Fig. 5. Each multiplier requires 5 transistors. So, the transistor count of the complete unit is only 22 transistors. The last two transistors not depicted on the Fig. 5 serve as an output switch, driven by the column signal. It features relative small area, simplicity, and high speed. These characteristics make it an interesting choice in low-level image processing embedded at focal plane. Each multiplier M_i (with $i = 1, \dots, 4$) takes two analog signals V_{i1} and V_{i2} and produces an output V_{iS} which is their product. The outputs of multipliers are all interconnected with a diode-connected transistor employed as load. Consequently, the global operation result at the V_S point is a linear combination of the four products V_{iS} . Image processing operations such as spatial convolution can be easily performed by connecting the inputs V_{i1} to the kernel coefficients and the inputs V_{i2} to the corresponding pixel values.

In order to keep the analysis simple, it is assumed in this section that the contribution of parasitic capacitances is negligible. Considering the MOS transistors operating in sub-threshold region, the output node V_{iS} of a multiplier can be

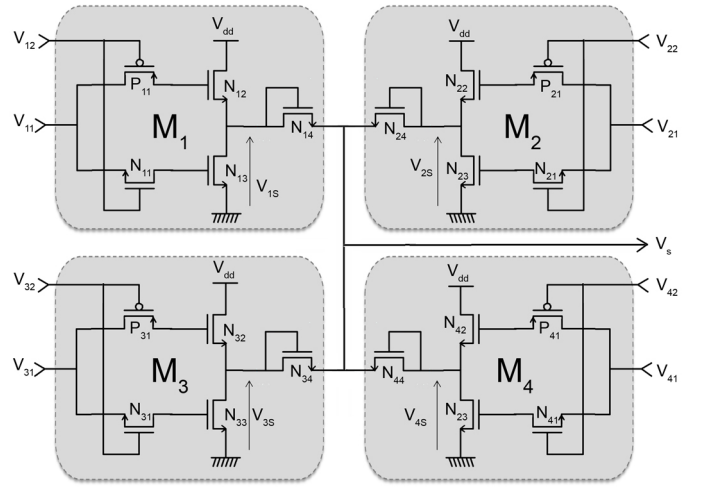


Fig. 5. The A²U structure

expressed as a function of the two inputs V_{i1} and V_{i2} as follows:

$$k_r (V_{ThN} - V_{i1}) (V_{i1} - V_{i2} - V_{ThN}) = (V_{i1} - V_{iS} - V_{ThN}) (V_{i2} - V_{iS} - V_{ThN} - V_{ThP}) \quad (1)$$

with k_r represents the transconductance factor, V_{ThN} and V_{ThP} are the threshold voltage for the NMOS and PMOS transistors. Around the operating point ($V_{dd}/2$), the variations of the output node mainly depend on the product $V_{i1}V_{i2}$. So, the equation 1 can be simplified and finally, the output node V_{iS} can be expressed as a simple first-order of the two input voltages V_{i1} and V_{i2} .

$$V_{iS} = M V_{i1} V_{i2} \quad \text{with,} \quad M = \frac{k_r - 1}{2V_{ThN} + V_{ThP}} \approx 8.07/V \quad (2)$$

The important value of the coefficient M gives to the structure a good robustness by limiting the impact of the second-order intermodulation products. The first consequence is a better linearity of our multiplier design integrating only 5 transistors.

The theoretical analysis has been validated by an experimental study realized on the test structures embedded on the chip. Fig. 6 shows the experimental measures obtained with two cosine signals as inputs of the multiplier structure.

$$V_{i1} = A \cos(2\pi f_1) \quad \text{with} \quad f_1 = 2.5 \text{ kHz} \quad (3)$$

$$V_{i2} = B \cos(2\pi f_2) \quad \text{with} \quad f_2 = 20 \text{ kHz} \quad (4)$$

In an ideal case, the voltage V_S at the output of the multiplier can be expressed as:

$$V_S = \frac{AB}{2} [\cos(2\pi(f_2 - f_1)) + \cos(2\pi(f_2 + f_1))] \quad (5)$$

The frequency spectrum, represented in Fig. 6(b) contains two main frequencies (17.5 kHz and 22.5 kHz) around the carrier frequency. The residues which appear in the spectrum are known as intermodulations products. Intermodulation is caused by non-linear behaviour of the of the structure (around 10 kHz and 30 kHz) and the insulation defects of input pads (at 40 kHz). Nevertheless, the amplitude of these intermodulation products is significantly lower than the two main frequencies.

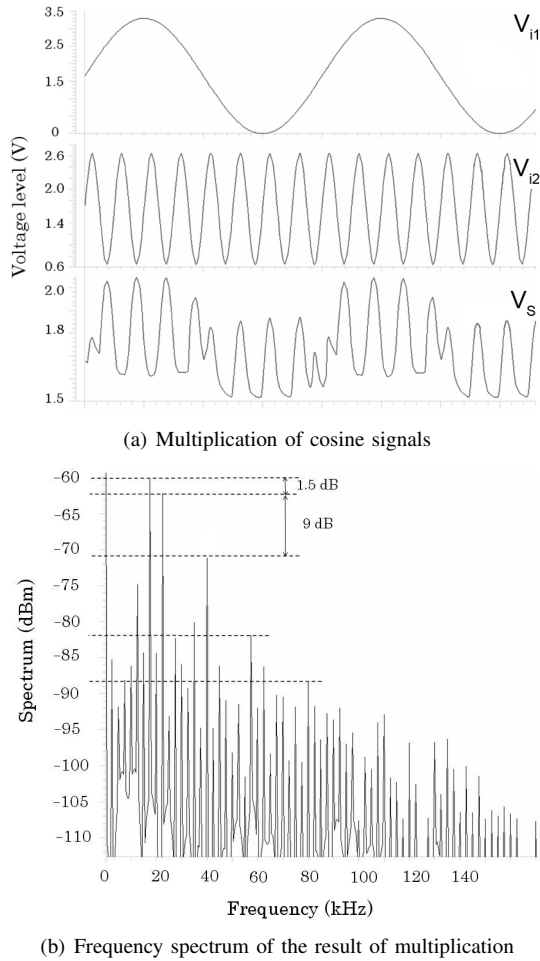


Fig. 6. Experimental results of the multiplication of two cosine signals by the four-quadrant multiplier

Indeed, at 40 kHz, the level of intermodulation is 9dB under the level of the main frequencies. Therefore, the contribution of the insulation defect is eight times smaller than the main signals. Furthermore, the experimental measures highlighted that the linearity of the multiplier is maximum for input signals with amplitude range between 0.6 V and 2.6 V. This corresponds to the typical range of values coming from the pixel since the $[AM]^2$ structures provide values around $V_{dd}/2=1.65V$.

V. CHIP CHARACTERIZATION

An experimental 64×64 pixel image sensor has been developed in a $0.35 \mu\text{m}$, 3.3 V, standard CMOS process with poly-poly capacitors. Its functional testing and its characterization were performed using a specific hardware platform. The hardware part of the imaging system contains a one million Gates Spartan-3 FPGA board with 32MB SDRAM embedded. This FPGA board is the XSA-3S1000 from XESS Corporation. An interface acquisition circuit includes three ADC from Analog Device (AD9048), high speed LM6171 amplifiers and others elements such as the motor lens. Fig. 7 shows the schematic and some pictures of the experimental platform.

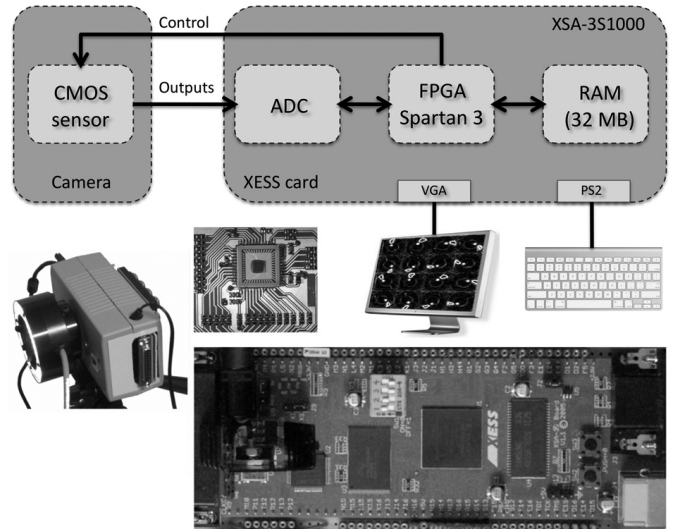


Fig. 7. Block diagram and pictures of the hardware platform including FPGA board and CMOS sensor

A. Electrical characterization

The sensor was quantitatively tested for conversion gain, sensitivity, fixed pattern noise, thermal reset noise, output levels disparities, voltage gain of the amplifier stage, linear flux, and dynamic range. Table I summarizes the main chip properties and the characterization results.

TABLE I
CHIP CHARACTERISTICS AND MEASUREMENTS

Technology	0.35 μm 2-poly 4-metal CMOS
Array size	64×64
Chip area	13.8 mm^2
Pixel size	$35 \mu\text{m} \times 35 \mu\text{m}$
Number of transistors	160 000
Number of transistors / pixel	38
Sensor Fill Factor	25 %
Dynamic power consumption	110 mW
Conversion gain	$54 \mu\text{V}/e^-$ RMS
Sensitivity	0.15 V/lux.s
Fixed Pattern Noise retina (FPN), dark	225 μV RMS
Thermal reset noise	68 μV RMS
Output levels disparities	4.3%
Voltage gain of the amplifier stage	12
Linear flux	98.5%
Dynamic range	68 dB

To determine these values, the sensor included specific test pixels in which some internal node voltages can be directly read. The test equipment hardware is based on a light generator with wavelength of 400 nm to 1100 nm. The sensor conversion gain was evaluated to $54 \mu\text{V}/e^-$ RMS with a sensitivity of 0.15 V/lux.s, thanks to the octagonal shape of the photodiode and the fill factor of 25%. At 10 000 frames/s, measured non-linearity is 0.12% over a 2 V range. These performances are similar to the sensor described in [24]. According to the experimental results, the voltage gain of the amplifier stage of

the two $[AM]^2$ is $Av = 12$ and the disparities on the output levels are about 4.3 %.

B. Fixed pattern noise

Image sensors always suffer from technology related non-idealities that can limit the performances of the vision system. Among them, fixed pattern noise (FPN) is the variation in output pixel values, under uniform illumination, due to device and interconnect mismatches across the image sensor. Two main types of FPN occur in CMOS sensors. First, offset FPN which takes place into the pixel is due to fluctuations in the threshold voltage of the transistors. Second, the most important source of FPN is introduced by the column amplifiers used in standard APS systems. In our approach, the layout is symmetrically built in order to reduce the offset FPN among each block of four pixels and to ensure uniform spatial sampling, as already depicted in the layout of a 2×2 pixel block in Fig. 1.

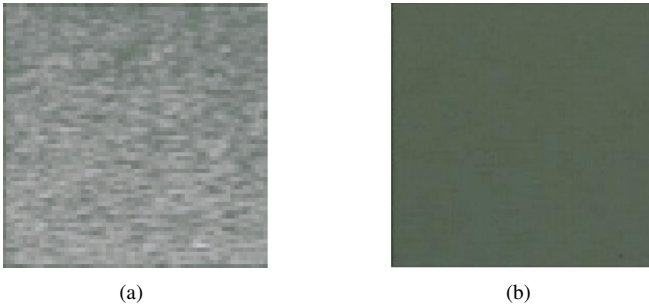


Fig. 8. Images of fixed pattern noise (a) without CDS and (b) with CDS for an integration time of 1 ms

Furthermore, our chip does not include any column amplifier since the amplification of the pixel values takes place into the pixel by means of an inverter. So, the gain FPN is very limited and only depends on the mismatch of the two transistors. FPN can be reduced by correlated double sampling (CDS). To implement CDS, each pixel output needs to be read twice, once after reset and a second time at the end of integration. The correct pixel signal is obtained by subtracting the two values. A CDS can be easily implemented in our chip. For this purpose, the first analog memory stores the pixel value just after the reset signal and the second memory stores the value at the end of integration. Then, at the end of the image acquisition, the two values can be transferred to the FPGA, responsible for producing the difference. In Fig. 8, the two images show fixed pattern noise with and without CDS using a 1-ms integration time. On the left image, the FPN is mainly due to the random variations in the offset voltages of the pixel-level analog structures. The experimental benchmarks of our chip reveal a FPN value of $225 \mu V$ RMS. The right picture shows the same image after analog CDS, performed as described above. The final FPN has been reduced by a factor of 34 to $6.6 \mu V$. In the rest of the results, CDS has not been implemented since FPN has low values. Only, an entire dark image is subtracted from the output images on the FPGA. Focus has been made on the development of low-level image processing using the two analog memories and the associated processing unit.

VI. HIGH-SPEED IMAGE PROCESSING APPLICATIONS

In this section, we provide experimental results of image processing implemented on our high-speed vision chip. First, we demonstrate the possibility of acquisition of raw images at different framerates, up to 10 000 frames/s. Secondly, we present an implementation of edge detection, based on the well-known Sobel operator.

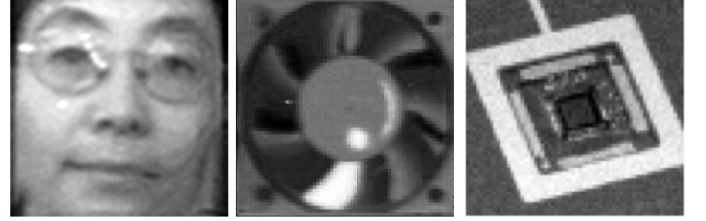


Fig. 9. Various raw images acquisition at 1 000, 5 000 and 10 000 frames/s

A. Sample Images

The prototype chip was used for acquisition of raw images. First, sample raw images of stationary scenes were captured at different framerates, as shown in Fig. 9. In the three views, no image processing is performed on the video stream, except for amplification of the photodiodes signal. From left to right, we can see a human face obtained at 1 000 frames/s, a static electric fan at 5 000 frames/s, and a electronic chip at 10 000 frames/s.

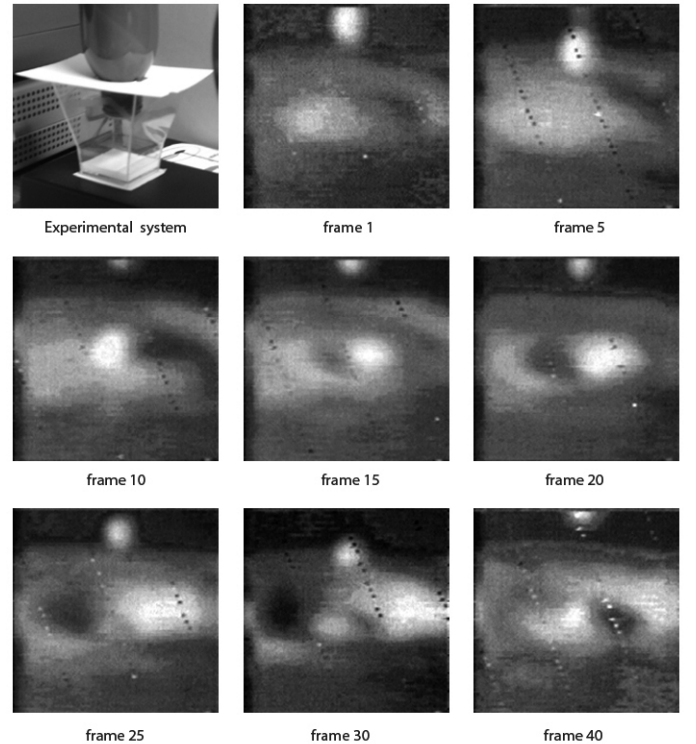


Fig. 10. A 2500 frames/s video sequence of a milk drop splashing

Fig. 10 represents different frames of a moving object, namely, a milk drop splashing sequence. In order to capture the

details of such a rapidly moving scene, the sensor operates at 2500 frames/s. and stores a sequence of 50 images. The frames 1, 5, 10, 15, 20, 25, 30 and 40 are shown on the Figure.

B. Sobel operator

The characteristics of our sensor, especially the analog processing unit, make it extremely useful for low-level image processing based on convolution masks. In this paragraph, we report an edge detector, the Sobel detector. The Sobel operator estimates the gradient of a 2D image. This algorithm is used for edge detection in the preprocessing stage of computer vision systems. The classical algorithm uses a pair of 3x3 convolution kernels (see Eq. 6), one to detect changes along the vertical axis (h_1) and another to detect horizontal contrast (h_2). For this purpose, the algorithm performs a convolution between the image and the sliding convolution mask over the image. It manipulates 9 pixels for each value to produce. The value corresponds to an approximation of the gradient centered on the processed image area.

$$h_1 = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad h_2 = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad (6)$$

The structure of our architecture is well-adapted to the evaluation of the Sobel algorithm. It leads to the result directly centered on the photo-sensor and directed along the natural axes of the image. The gradient is computed in each pixel of the image by performing successive linear combinations of the 4 adjacent pixels. For this purpose, each 3x3 kernel mask is decomposed into two 2x2 masks that successively operate on the whole image. For the kernel h_1 , the corresponding 2x2 masks are:

$$m_1 = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} \quad m_2 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \quad (7)$$

The Fig 11 represents the 3x3 mask centered on the pixel ph_5 . Each octagonal photodiode ph_i ($i = 1, \dots, 9$) is associated with a processing element PE_i , represented with a circle on the Figure. Each PE_i is positioned on the bottom right of its photodiode, as in the real layout of the circuit (see Fig 1). The first mask m_1 contributes to evaluate the following series of operations for the four PE_i s:

$$\begin{aligned} V_{11} &= -(V_{ph_1} + V_{ph_4}) \\ V_{12} &= -(V_{ph_2} + V_{ph_5}) \\ V_{14} &= -(V_{ph_4} + V_{ph_7}) \\ V_{15} &= -(V_{ph_5} + V_{ph_8}) \end{aligned} \quad (8)$$

and the second mask m_2 computes:

$$\begin{aligned} V_{21} &= +(V_{ph_2} + V_{ph_5}) \\ V_{22} &= +(V_{ph_3} + V_{ph_6}) \\ V_{24} &= +(V_{ph_5} + V_{ph_8}) \\ V_{25} &= +(V_{ph_6} + V_{ph_9}) \end{aligned} \quad (9)$$

with V_{ij} corresponding to the result provided by the processing element PE_j ($j = 1, 2, \dots, 9$) with the mask m_i ($i = 1, 2$), and V_{ph_k} ($k = 1, 2, \dots, 9$), the voltages representing

the incidental illumination on each photodiode ph_k . Then, the evaluation of the gradient at the center of the mask can be computed by summing the different values on the external FPGA. Note that $V_{12} = -V_{21}$ and $V_{15} = -V_{24}$. So, the final sum can be simplified and written as $V_{h_1} = V_{11} + V_{22} + V_{25} + V_{14}$. If we define a retina cycle as the time spent for the configuration of the coefficients kernel and the preprocessing of the image, the evaluation of the gradient on the vertical direction only spends a frame acquisition and two retina cycles. By generalization, the estimation of the complete gradient along the two axis spend 4 cycles because it involves 4 dynamic configurations.

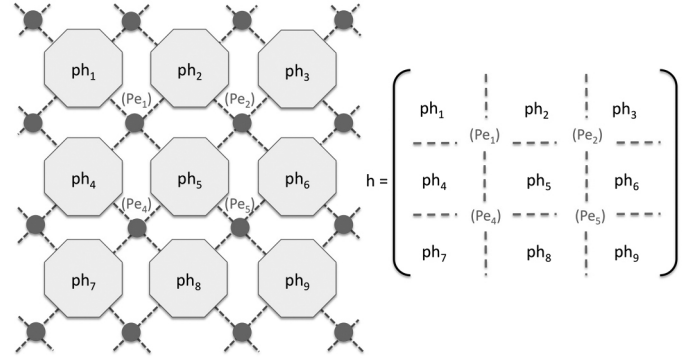


Fig. 11. 3x3 kernel used by the 4 processing elements

In short, the dynamic assignment of coefficient values from the external processor gives the system some interesting dynamic properties. The system can be easily reconfigured by changing the internal coefficients for the masks between two successive computations. First, this allows the possibility to dynamically change the image processing algorithms embedded in the sensor. Secondly, this enables the evaluation of some complex pixel-level algorithms, implying different successive convolutions. The images can be captured at higher framerates than the standard framerate, processed by exploiting the the analog memories and the reconfigurable processing elements and output at a lower framerate depending of the number of the dynamic reconfigurations. Moreover, the analog arithmetic units implementing these pixel-level convolutions drastically decrease the number of single operations such as additions and multiplications executed by an external processor (a FPGA in our case) as shown in Fig. 7. Indeed, in the case of our experimental 64×64 pixel sensor, the peak performance is equivalent to 4 parallel signed multiplications by pixel at 10000 frames/s, i.e. more than 160 million multiplications per second. With a VGA resolution (640×480), the performance level would increase to a factor of 75, leading to about 12 billion multiplications per second. Processing this data flow by external processors will imply important hardware resources in order to cope with the temporal constraints.

As an illustration of the Sobel algorithm, Fig. 12 is an example sequence of 16 images of a moving object, namely, an electric fan. Two white specific markers are placed on the fan, i.e a small circle near the rotor and a painted blade. The speed rotation of the fan is 3750 rpm. In order to capture such a rapidly moving object, a short integration time ($100 \mu s$) was

TABLE II
COMPARISON WITH OTHER SIMD SENSORS

Chip	This work	SCAMP-3 [44], [45]	MIMD IP Chip [46]	ACE16k [18], [47]	VCS-IV [48], [49]	PVLSAR2.2 [50]
Technology	0.35 μm	0.35 μm	1.2 μm	0.35 μm	0.35 μm	0.8 μm
Resolution	64\times64	128 \times 128	80 \times 78	128 \times 128	64 \times 64	128 \times 128
Pixel pitch	35 $\mu\text{m}$$\times$35 μm	49.35 μm \times 49.35 μm	45.6 μm \times 45 μm	75.5 μm \times 75.3 μm	67.4 μm \times 67.4 μm	60 μm \times 60 μm
Fill Factor	25 %	5.6 %	33 %	?	10 %	30 %
Transistors/PE	38 tr.	128 tr.	9 tr.	198 tr.	84 tr.	50 tr.
PE Type	Analog	Analog	Analog	Analog	Digital	Digital
Framerate	10000 fps	1000 fps	9600 fps	1000 fps	1000 fps	1000 fps
Image Processing	Mask-based Image Processing	Low-level Image Processing	Spatial Convolutions	Spatial Convolutions	Low-level Image Processing	Low to mid-level Image Processing

used for the frames acquisition. The Sobel algorithm allow to distinguish clearly the two white markers even with a high framerate.

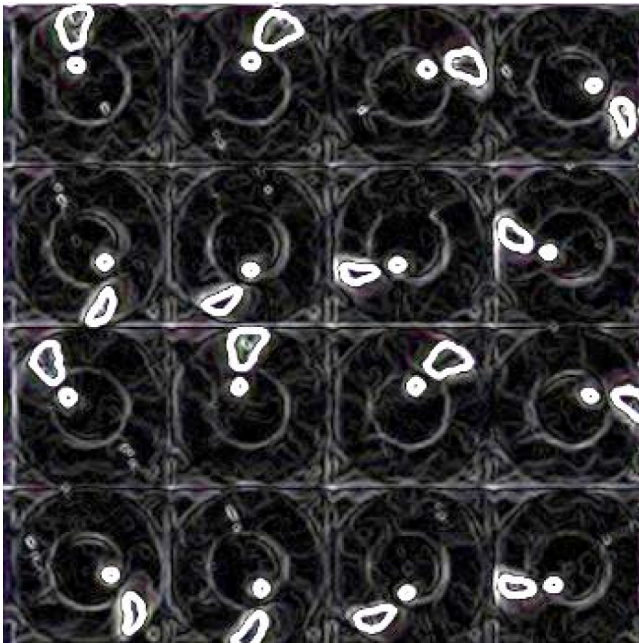


Fig. 12. Sequence of 16 images with Sobel operator

C. Strategies used for general spatial filter

In the preceding sections, we focused on 2x2 and 3x3 convolution masks. In the case of a 2x2 mask, the coefficients are fixed once before the beginning of the acquisition frame. In the case of a 3x3 mask, two possibilities can occur. First, the 3x3 mask presents some symmetrical properties (such as the Sobel kernel) and then the coefficients values can be fixed as in a 2x2 mask. Second, if the mask is not symmetric, it is necessary to dynamically reconfigure the coefficients during the acquisition frame. For masks which size is greater than 3x3 and more generally in the case of an $N \times N$ mask, a dynamic reconfiguration of coefficients is necessary during the acquisition frame in order to evaluate the successive values of the linear combinations of pixels.

VII. COMPARISON WITH OTHER SIMD VISION CHIPS

Table II shows an overview of some representative SIMD vision chips based on different alternatives for implementing vision processing at focal plane. The first column corresponds to the chip described in this paper. On the second column, we can find the main characteristics of SCAMP-3, a general purpose processor array implementing a variety of low-level image processing tasks at a high framerate. The third column describes a Multiple Instruction Multiple Data Image Processing Chip (MIMD IP Chip) performing spatial convolutions using kernels from 3 \times 3 to 11 \times 11. The fourth column presents a vision chip integrating an imager and an array of mixed-signal SIMD processing elements that can process grayscale images with Cellular Neural Network Universal Machines (CNN-UM) mode of operation. The fifth vision chip in the table is a programmable SIMD vision chip that can perform various early visual processing such as edge detection, smoothing or filtering by chaining processing elements and reconfiguring the hardware dynamically. Finally, the last one is a programmable artificial retina in which each pixel contains a tiny digital processing element capable of grey-level image processing from low to mid-level vision (motion detection, segmentation, pattern recognition).

Compared to this state-of-the-art of high-speed CMOS image sensors, one easily sees that the chip reported in this paper lead to some major improvements. With a pixel size of 35 μm by 35 μm , the pixel pitch is almost 1.5 more compact than the smallest pixel described in [44], [45]. This contributes either to a better resolution of the sensor for a given chip area or a lower cost for a given resolution. At the same time, as compared with the other processing elements, our solution relies on a compact analog arithmetic unit based on a four-quadrant multiplier architecture using only 38 minimal size transistors. This leads to a fill factor of about 25 %. Consequently, the active area of our photodiode is bigger compared to the major part of the other chips, providing a best sensibility to the sensor at high framerates of thousands of images per second.

From the performance point of view, all the chips on the table implement low-level image processing by programming or dynamically configuring the processing elements. Our solution is able to capture image, run user-defined 3x3 convolution masks, and provide the results on the sensor output bus in less than 200 μs , giving a framerate of 5000 images per second.

These temporal performances are compatible and even slightly higher than the other sensors since each of them is only able to provide low-level image processing results at 1000 frames per second, except the MIMD IP Chip.

From the programmability point of view, our chip can suffer from less versatility compared to other programmable sensors. Indeed, the implementation of complex low-level image processing require successive reconfigurations of the internal masks coefficients of the processing elements. This task may be more difficult in comparison with algorithms easily written in a machine-level language for the programmable sensors, such as the SCAMP-3 sensor.

VIII. CONCLUSION AND PERSPECTIVES

An experimental pixel sensor implemented in a standard digital CMOS 0.35 μm process has been described in this paper. Each 35 μm \times 35 μm pixel contains 38 transistors implementing a circuit with photo-current integration, two [AM]² (Analog Memory, Amplifier, and Multiplexer), and an A²U (Analog Arithmetic Unit).

Experimental chip results reveal that raw image acquisition at 10000 frames per second can be easily achieved using the parallel A²U implemented at pixel-level. With basic image processing, the maximal framerate slows down to about 5000 fps. The potential for dynamic reconfiguration of the sensor was also demonstrated in the case of the Sobel operator.

The next step in our research will be the design of a similar circuit in a modern 130nm CMOS technology. The main objective will be to design a pixel of less than 10 μm \times 10 μm with a fill factor of 20%. A second possibility would be the design of a sensor with emergent technologies using amorphous silicon in which 3D pixels can be built. With a photosensitive layer placed just behind the optical part, the pixel fill factor can reach 100% since the processing elements can be packed in the empty space below the photodiode.

Thus, with the increasing scaling of the transistors in such technologies, we could consider the implementation of more sophisticated image processing operators dedicated to face localization and recognition. Previous works of our team [51] have demonstrated the needs of dedicated CMOS sensors embedding low-level image processing such as features extraction. Moreover, actual works [52] focus on a recent face detector named Convolutional Face Finder (CFF) [53]. CFF is based on a multi-layer convolutional neural architecture. The CFF consists of six successive neural layers. The first four layers extract characteristic features, and the last two perform the classification. Our objective would be to implement at pixel-level the first layers based on convolutions by different masks from 2 \times 2 to 5 \times 5.

In order to evaluate this future chip in some realistic conditions, we would like to design a CIF sensor (352 \times 288 pixels), which leads to a 3.2 mm \times 2.4 mm in a 130 nm technology. The exploitation of high FPS capability with this sensor could be achieved by two complementary ways. The first one is to integrate a dedicated Input/Output module, which is able to cope with a gigapixel per second bandwidth. Such modules have been already designed in other high-speed

CMOS sensors. As an example, we can cite the Digital Pixel Sensor [24] with his 64-bit (8-pixel) wide bus operating at 167 MHz, able to output 1.33 GB/s. The second way is to build a large sensor by assembling 64 \times 64 pixel modules with a dedicated output bus for each of them. For example, with a 384 \times 256 pixel sensor, this solution only requires 6x4=24 dedicated outputs. In the same time, we will focus on the development of a fast analog to digital converter (ADC). The integration of this ADC on future chips will allow us to provide new and sophisticated vision systems on chip (ViSOC) dedicated to digital embedded image processing at thousands of frames per second.

REFERENCES

- [1] E. Fossum, "Active pixel sensors: Are CCDs dinosaurs?" *International Society for Optical Engineering (SPIE)*, vol. 1900, pp. 2–14, 1993.
- [2] —, "CMOS Image Sensor : Electronic Camera On A CHIP," *IEEE Transactions on Electron Devices*, vol. 44, no. 10, pp. 1689–1698, October 1997.
- [3] A. El Gamal, D. Yang, and B. Fowler, "Pixel level processing – Why, What and How?" in *Proceedings of the SPIE Electronic Imaging '99 conference*, vol. 3650, January 1999, pp. 2–13.
- [4] P. Seitz, "Solid-State Image Sensing," *Handbook of computer Vision and Applications*, vol. 1, pp. 165–222, 2000.
- [5] D. Litwiller, "CCD vs. CMOS: Facts and Fiction," *Photonics Spectra*, pp. 154–158, January 2001.
- [6] C. H. Aw and B. Wooley, "A 128128-pixel standard-cmos image sensor with electronic shutter," *IEEE Journal of Solid State Circuits*, vol. 31, no. 12, pp. 1922–1930, 1996.
- [7] M. Loiaz, K. Singh, A. Blanksby, D. Inglis, K. Azadet, and B. Ackland, "A 200mv 3.3v CMOS Color Camera IC Producing 352 \times 288 24-b Video at 30 Frames/s," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2092–2103, 1998.
- [8] S. Smith, J. Hurwitz, M. Torrie, D. Baxter, A. Holmes, M. Panaghiston, R. Henderson, A. Murrayn, S. Anderson, and P. Denyer, "A single-chip 306x244-pixel CMOS NTSC video camera," in *In ISSCC Digest of technical papers*, San Francisco, CA, 1998, pp. 170–171.
- [9] O. Yadid-Pecht and A. Belenky, "In-Pixel Autoexposure CMOS APS," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 8, pp. 1425–1428, August 2003.
- [10] P. Acosta-Serafini, M. Ichiro, and C. Sodini, "A 1/3" VGA Linear Wide Dynamic Range CMOS Image Sensor Implementing a Predictive Multiple Sampling Algorithm With Overlapping Integration Intervals," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 9, pp. 1487–1496, September 2004.
- [11] L. Kozlowski, G. Rossi, L. Blanquart, R. Marchesini, Y. Huang, G. Chow, J. Richardson, and D. Standley, "Pixel Noise Suppression via SoC Management of Target Reset in a 1920 \times 1080 CMOS Image Sensor," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2766–2776, December 2005.
- [12] M. Sakakibara, S. Kawahito, D. Handoko, N. Nakamura, M. Higashi, K. Mabuchi, and H. Sumi, "A High-Sensitivity CMOS Image Sensor With Gain-Adaptative Column Amplifiers," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 5, pp. 1147–1156, May 2005.
- [13] R. H. Nixon, S. E. Kemeny, C. O. Staller, and E. R. Fossum, "128 x 128 CMOS photodiode-type active pixel sensor with on-chip timing, control, and signal chain electronics," in *Proc. SPIE Vol. 2415, p. 117-123, Charge-Coupled Devices and Solid State Optical Sensors V, Morley M. Blouke; Ed.*, ser. Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, M. M. Blouke, Ed., vol. 2415, Apr. 1995, pp. 117–123.
- [14] S. Wu, H. Chien, D. Yaung, C. Tseng, C. Wang, C. Chang, and Y. Hsiao, "A high performance active pixel sensor with 0.18um cmos color imager technology," *Electron Devices Meeting, 2001. IEDM Technical Digest. International*, pp. 555–558, 2001.
- [15] S. Decker, D. McGrath, K. Brehmer, and C. Sodini, "A 256256 cmos imaging array with wide dynamic range pixels and column-parallel digital output," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2081–2091, Dec 1998.
- [16] K. Yoon, C. Kim, B. Lee, and D. Lee, "Single-chip cmos image sensor for mobile applications," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 12, pp. 1839–1845, Dec 2002.

- [17] A. Krymsky and T. Niarong, "A 9-V/Lux 5000-Frames/s 512 x 512 CMOS Sensor," *IEEE Transactions on Electron Devices*, vol. 50, no. 1, pp. 136–143, January 2003.
- [18] G. Cembrano, A. Rodriguez-Vazquez, R. Galan, F. Jimenez-Garrido, S. Espejo, and R. Dominguez-Castro, "A 1000 FPS at 128 × 128 Vision Processor With 8-Bit Digitized I/O," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 7, pp. 1044–1055, July 2004.
- [19] L. Lindgren, J. Melander, R. Johansson, and B. Miller, "A Multiresolution 100-GOPS 4-Gpixels/s Programmable Smart Vision Sensor for Multi-sense Imaging," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 6, pp. 1350–1359, June 2005.
- [20] Y. Sugiyama, M. Takumi, H. Toyoda, N. Mukozaka, A. Ihori, T. Kurashina, Y. Nakamura, T. Tonbe, and S. Mizuno, "A High-Speed CMOS Image With Profile Data Acquiring Function," *IEEE Journal of Solid-State Circuits*, vol. 40, pp. 2816–2823, 2005.
- [21] D. Martin, H.-S. Lee, and I. Masaki, "A mixed-signal array processor with early vision applications," *Solid-State Circuits, IEEE Journal of*, vol. 33, no. 3, pp. 497–502, Mar 1998.
- [22] A. Krymski, D. Van Blerkom, A. Andersson, N. Bock, B. Mansoorian, and E. Fossum, "A high speed, 500 frames/s, 1024x1024 cmos active pixel sensor," *VLSI Circuits, 1999. Digest of Technical Papers. 1999 Symposium on*, pp. 137–138, 1999.
- [23] N. Stevanovic, M. Hillebrand, B. Hosticka, and A. Teuner, "A cmos image sensor for high-speed imaging," *Solid-State Circuits Conference, 2000. Digest of Technical Papers. ISSCC. 2000 IEEE International*, pp. 104–105, 449, 2000.
- [24] S. Kleinfelder, S. Lim, X. Liu, and A. El Gamal, "A 10 000 Frames/s CMOS Digital Pixel Sensor," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 12, pp. 2049–2059, December 2001.
- [25] D. Handoko, K. S. Y. Takokoro, M. Kumahara, and A. Matsuzawa, "A CMOS image sensor for local-plane motion vector estimation," in *Symposium of VLSI Circuits*, vol. 3650, June 2000, pp. 28–29.
- [26] S. Lim and A. El Gamal, "Integrating Image Capture and Processing – Beyond Single Chip Digital Camera," in *Proceedings of the SPIE Electronic Imaging '2001 conference*, vol. 4306, San Jose, CA, January 2001.
- [27] X. Liu and A. El Gamal, "Photocurrent estimation from multiple non-destructive samples in a CMOS image sensor," in *Proceedings of the SPIE Electronic Imaging '2001 conference*, vol. 4306, San Jose, CA, January 2001.
- [28] D. Yang, A. El Gamal, B. Fowler, and H. Tian, "A 640 x 512 CMOS Image Sensor with Ultra Wide Dynamix Range Floating-Point Pixel-Level ADC," *IEEE Journal of Solid-State Circuits*, vol. 34, pp. 1821–1834, December 1999.
- [29] O. Yadid-Pecht and E. Fossum, "CMOS APS with autoscaling and customized wide dynamic range," in *IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, vol. 3650, June 1999, pp. 48–51.
- [30] D. Stoppa, A. Somoni, L. Gonzo, M. Gottardi, and G.-F. Dalla Betta, "Novel CMOS Image Sensor With a 132-dB Dynamic Range," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 12, pp. 1846–1852, December 2002.
- [31] X. Liu and A. El Gamal, "Simultaneous image formation and motion blur restoration via multiple capture," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 3, 2001, pp. 1841–1844.
- [32] C.-Y. Wu and C.-T. Chiang, "A Low-Photocurrent CMOS Retinal Focal-Plane Sensor With a Pseudo-BJT Smoothing Network and an Adaptive Current Schmitt Trigger for Scanner Applications," *IEEE Sensors Journal*, vol. 4, no. 4, pp. 510–518, August 2004.
- [33] P. Dudek and P. Hicks, "An analogue simd focal-plane processor array," in *Proceedings of the International Symposium on Circuits and Systems (ISCAS)*, vol. 4, 6–9 May 2001, pp. 490–493 vol. 4.
- [34] A. El Gamal and H. Eltouky, "Cmos image sensors," *Circuits and Devices Magazine, IEEE*, vol. 21, no. 3, pp. 6–20, May-June 2005.
- [35] J. Lee and R. Hornsey, "CMOS Photodiodes with Substrate Openings for Higher Conversion Gain in Active Pixel Sensor," in *IEEE Workshop on CCDs and Advanced Image Sensors*, Crystal Bay, Nevada, June 2001.
- [36] C. Wu, Y. Shih, J. Lan, C. Hsieh, C. Huang, and J. Lu, "Design, optimization, and performance analysis of new photodiode structures for CMOS active-pixel-sensor (APS) imager applications," *IEEE Sensors Journal*, vol. 4, no. 1, pp. 135–144, February 2004.
- [37] I. Shcherback, A. Belenky, and O. Yadid-Pecht, "Empirical dark current modeling for complementary metal oxide semiconductor active pixel sensor," *Optical Engineering*, vol. 41, no. 6, pp. 1216–1219, June 2002.
- [38] I. Shcherback and O. Yadid-Pecht, "Photoresponse analysis and pixel shape optimization for CMOS active pixel sensors," *IEEE Transactions on Electron Devices*, vol. 50, no. 1, pp. 12–18, January 2003.
- [39] J. Dubois, D. Ginhac, M. Paindavoine, and B. Heyrman, "A 10 000 fps cmos sensor with massively parallel image processing," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 3, p. In Press, March 2008.
- [40] O. Yadid-Pecht, B. Pain, C. Staller, C. Clark, and E. Fossum, "CMOS Active Pixel Sensor Star Tracker with Regional Electronic Shutter," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 2, pp. 285–288, February 1997.
- [41] G. Chapinal, S. Bota, M. Moreno, J. Palacin, and A. Herms, "A 128 × 128 CMOS Image Sensor With Analog Memory for Synchronous Image Capture," *IEEE Sensors Journal*, vol. 2, no. 2, pp. 120–127, April 2002.
- [42] C. Ryan, "Applications of a four-quadrant multiplier," *IEEE Journal of Solid-State Circuits*, vol. 5, no. 1, pp. 45–48, Feb 1970.
- [43] S. Liu and Y. Hwang, "CMOS Squarer and Four-Quadrant Multiplier," *IEEE Transactions on Circuits and Systems-I:Fundamental Theory and Applications*, vol. 42, no. 2, pp. 119–122, Feb 1995.
- [44] P. Dudek, "Implementation of simd vision chip with 128×128 array of analogue processing elements," in *Proceedings of the International Symposium on Circuits and Systems (ISCAS)*, vol. 5, 2005, pp. 5806–5809.
- [45] P. Dudek and S. Carey, "General-purpose 128×128 SIMD processor array with integrated image sensor," *Electronic Letters*, vol. 42, no. 12, pp. 678–679, Jan 2006.
- [46] R. Etienne-Cummings, Z. Kalayjian, and D. Cai, "A programmable focal-plane mimd image processor chip," *Solid-State Circuits, IEEE Journal of*, vol. 36, no. 1, pp. 64–73, Jan 2001.
- [47] A. Rodriguez-Vasquez, G. Linan-Cembrano, L. Carranza, E. Roca-Moreno, R. Carmona, F. Jimenez-Garrido, R. Dominguez-Castro, and S. Meana, "ACE16k: the third generation of mixed-signal SIMD-CNN ACE chips toward VSoCs," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 5, pp. 851–863, 2004.
- [48] T. Komuro, I. Ishii, M. Ishikawa, and A. Yoshida, "A digital vision chip specialized for high-speed target tracking," *IEEE Transactions on Electron Devices*, vol. 50, no. 1, pp. 191–199, 2003.
- [49] T. Komuro, S. Kagami, and M. Ishikawa, "A dynamically reconfigurable SIMD processor for a vision chip," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 1, pp. 265–268, 2004.
- [50] F. Paillet, D. Mercier, and T. Bernard, "Second generation programmable artificial retina," 1999, pp. 304–309.
- [51] F. Yang and M. Paindavoine, "Implementation of an RBF neural network on embedded systems: real-time face tracking and identity verification," *IEEE Transactions on Neural Networks*, vol. 14, no. 5, pp. 1162–1175, Sept 2003.
- [52] N. Farrugia, F. Mamalet, S. Roux, F. Yang, and M. Paindavoine, "A Parallel Face Detection System Implemented on FPGA," in *Proceedings of the International Symposium on Circuits and Systems (ISCAS)*, New Orleans, USA, May 27–30 2007, pp. 3704–3707.
- [53] C. Garcia and M. Delakis, "Convolutional face finder: a neural architecture for fast and robust face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1408–1423, Nov 2004.