A 1000 FPS at 128/spl times/128 vision processor with 8-bit digitized I/O
— **Source link**

G.L. Cembrano, Ángel Rodríguez-Vázquez, R.C. Galan, F. Jimenez-Garrido ...+2 more authors

Related papers:

- A general-purpose processor-per-pixel analog SIMD vision chip

- A digital vision chip specialized for high-speed target tracking

- A multiresolution 100-GOPS 4-Gpixels/s programmable smart vision sensor for multisense imaging

- A programmable focal-plane MIMD image processor chip

- A 10 000 fps CMOS Sensor With Massively Parallel Image Processing

A 1000 FPS at 128 × 128 Vision Processor With 8-Bit Digitized I/O

Gustavo Liñán Cembrano, *Associate Member, IEEE*, Angel Rodríguez-Vázquez, *Fellow, IEEE*,
Ricardo Carmona Galán, *Associate Member, IEEE*, F. Jiménez-Garrido, Servando Espejo, and
Rafael Domínguez-Castro

Abstract—This paper presents a mixed-signal programmable chip for high-speed vision applications. It consists of an array of processing elements, arranged to operate in accordance with the principles of single instruction multiple data (SIMD) computing architectures. This chip, implemented in a 0.35-μm fully digital CMOS technology, contains $\sim$ 3.75 M transistors and exhibits peak performance figures of 330 GOPS (8-bit equivalent giga-operations per second), 3.6 GOPS/mm^2 and 82.5 GOPS/W. It includes structures for image acquisition and for image processing, meaning that it does not require a separate imager for operation. At the sensory side, integration and log-compression sensing circuits are embedded, thus allowing the chip to handle a large variety of illumination conditions. At the processing plane, analog and digital circuits are employed whose parameters can be programmed and their architecture reconfigured for the realization of software-coded processing algorithms. The chip provides, and accepts, 8-bit digitized data through a 32-bit bidirectional data bus which operates at 120 MB/s. Experimental results show that frame rates of 1000 frames per second (FPS) can be achieved under room illumination conditions — applications using exposures of about 50 μs have been recently reached by using special illumination setups. The chip can capture an image, run approximately 150 two-dimensional linear convolutions, and download the result in 8-bit digital format, in less than 1 ms. This feature, together with the possibility of executing sequences of user-definable instructions (stored on a full-custom 32-kb on-chip memory), and storing intermediate results (up to 8 grayscale images) makes the chip a true general-purpose sensory/processing device.

Index Terms—Focal plane array processors, mixed-signal SIMD vision chips, visual microprocessors.

I. INTRODUCTION

VISION is a computation-demanding activity which involves many tasks and data types. These tasks are clustered hierarchically from bottom to top as Fig. 1 illustrates: starting with *low-level image processing*, (basically 2-D spatial filtering), passing through the *estimation of features* (segmentation of textures, motion, etc.), and ending with the *analysis of images and objects* (image classification, identification, 3-D reconstruction, etc.). Low-level tasks consist of simple operations executed on a very large data set; actually the whole set of pixel values. These low-level tasks do not require great accuracy; six to seven equivalent bits are sufficient [1]–[3]. However,

operating with whole images means extensive accesses to the memory, and poses hard constraints on the bandwidth of the communications between the memory and the processor. Besides, using the front-end device of a vision system just for sensing implies wasting system resources to convert, process, and transmit a lot of redundant information.

The relaxed accuracy requirements of low-level vision tasks [1]–[3] render analog techniques suitable for high-speed vision. Because analog circuits with moderate accuracy are quite efficient in terms of area occupation and energy consumption, they are adequate for the design of fully parallel focal-plane processing vision devices. Actually, in the last few years different analog and mixed-signal architectures and chips have been proposed which combine sensing and low-level image processing to achieve high-speed operation within a single chip [4]–[8].

The chip presented here outperforms all these previous proposals in terms of complexity, computational capability, and performance. It is a fully programmable, reconfigurable device which embeds all the structures needed for it to operate as a software-programmable, general-purpose image processor. It has been conceived for application in two alternative scenarios. First of all, as the core element of a vision system. In such a case, it acquires images through its embedded sensors, processes these images according to a user-defined program (which may contain image combinations, data bifucartions, and conditional executions), and finally gives the result (either images or event addresses) in 8-bit format. In this approach, resolution is limited by the 128 × 128 chip size, but processing tasks run at the chip's maximum attainable speed, which is less than 3 μs for 3 × 3 convolutions, including internal self-calibration. Its second use is as an image coprocessor. In this case, higher resolution images are captured by a conventional imager, windowed in chip-sized pieces by a controller and transmitted to the chip in 8-bit format for processing. This approach unavoidably leads to lower frame rates but opens the door for low-frame-rate high-resolution applications.

II. DESCRIPTION OF THE ARCHITECTURE

A. Global Overview

The chip presented in this paper, depicted in Fig. 2, follows the single instruction multiple data (SIMD) computing paradigm [4], [5]. Here, an array of identical processing elements (PEs) executes a sequence of instructions, the same for all PEs, issued by a global controller which is shared by all the PEs of the array. Data are defined at the PE level, in such a way that each PE in a given position of the array corresponds to a pixel
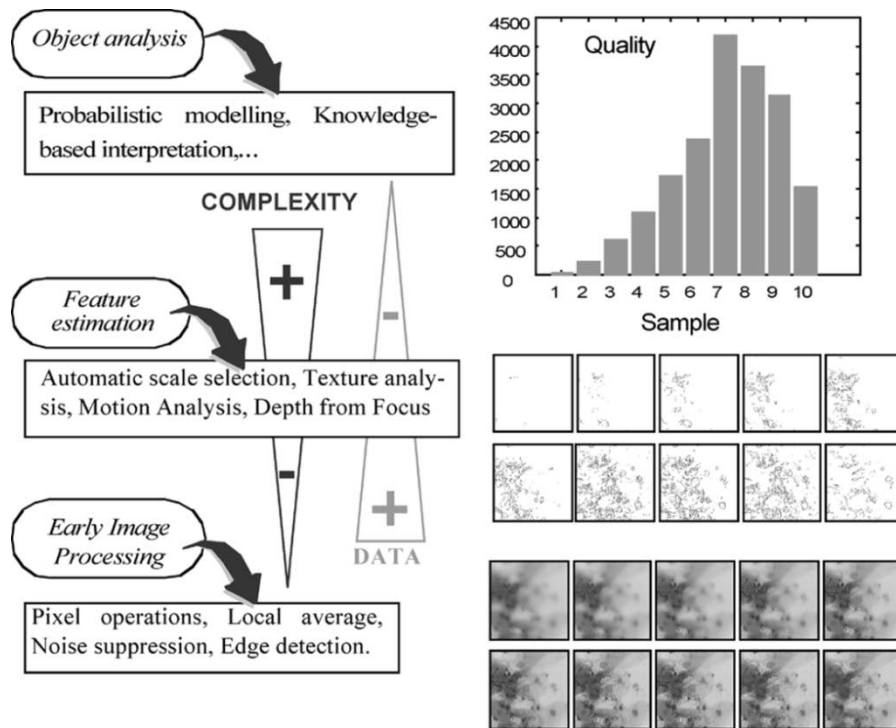
Fig. 1. Hierarchical classification of vision tasks as suggested in [3].
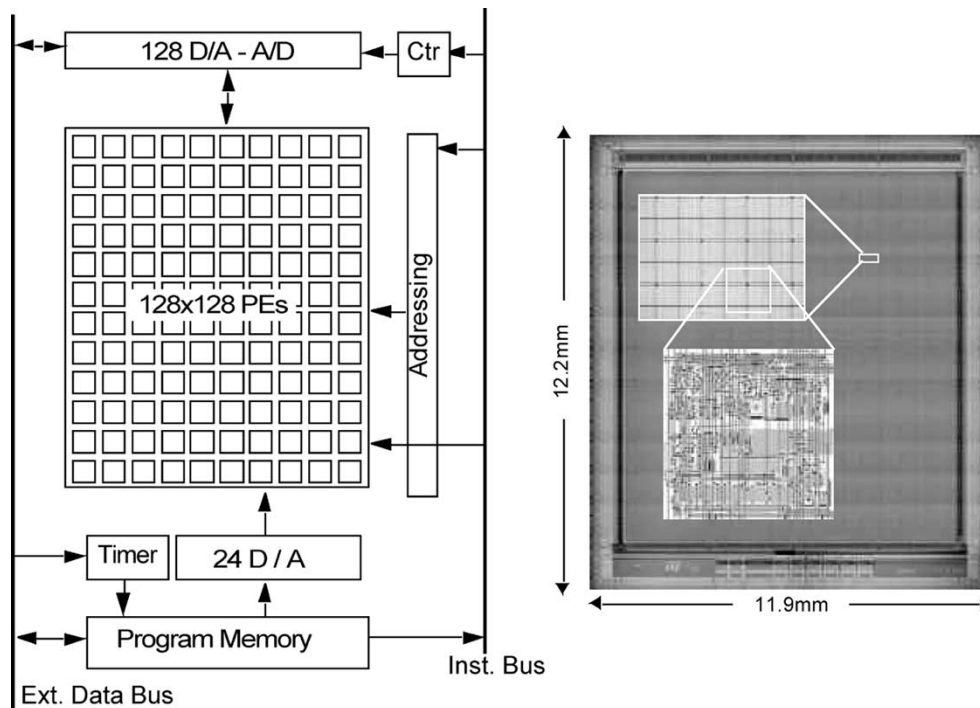
Fig. 2. Chip photograph and block diagram.

in the same position in the input or output images. PEs are conceived as programmable processing devices. Most internal analog and digital blocks can be reconfigured as needed to implement most low-level [3] vision processing tasks. Each PE includes the following.

- Photosensors. For optical acquisition of the visual information. An additional mechanism for loading images elec-

trically onto the chip is provided to enable the operation as coprocessor.
- Pixel memories. To store different inputs or the results of previous processing steps and to allow for the execution of bifurcated-flow algorithms.
- Programmable Boolean operator. To perform image-wise Boolean combinations, such as XOR (image#1, image#2).

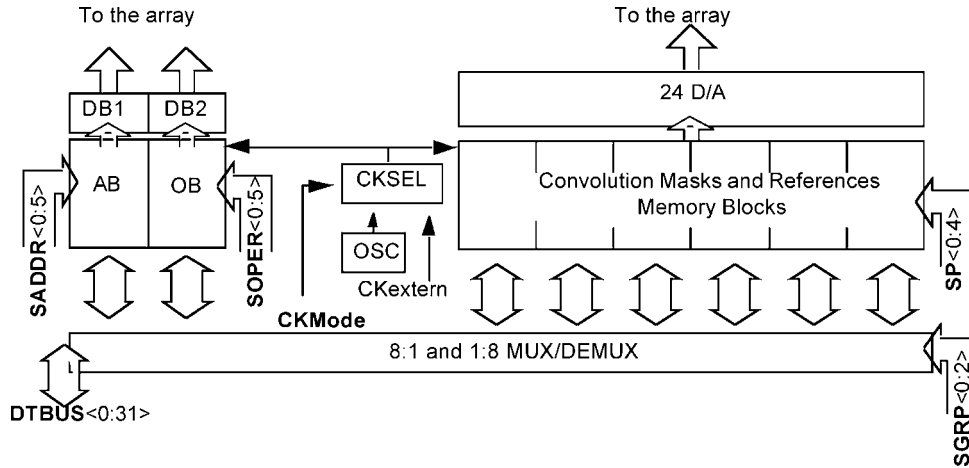Fig. 3.   Structure of the program memory in the chip which is employed for programming both instructions to be executed and convolution masks to be applied.

- Twelve programmable analog multipliers. To implement convolution kernels, which can also be employed for pixel scaling.
- Address event module. To make the chip output addresses of active pixels in the array, instead of whole images. This allows for a very fast extraction of the information contained within sparse black and white images, a quite common outcome of many processing tasks.
- Diffusion network. For the fast execution of the heat-diffusion equation. Indeed, the time constant is so fast that the associated network reaches its steady state in less than 200 ns, thus providing an ultrafast mechanism to calculate the average gray level of an image, which is extremely useful for adjusting the exposure in optical integration modes.

The chip has been realized in a digital CMOS 0.35-μm five-metal one-poly technology and contains about 3.75 M transistors, 85% of them working in analog mode. It reaches peak computing figures [1] of 330 GOPS, 3.6 GOPS/mm^2, and 82.5 GOPS/W, also providing and accepting 8-bit digitized images at 120 MB/s through a 32-bit data bus.

The chip core, its computing engine, is defined by an array of PEs, described in Section III. This core operates basically in analog mode. Nevertheless, in order to make the system versatile and usable in real-life applications, we have provided it with a standard digital interface which is described in the following section.

B. Program Block

The program block adds the capability of storing the vision processing algorithm on-chip. Although its components are conceptually simple, basically, this block consists of SRAMs and 8-bit digital-to-analog (D/A) converters. Its role is crucial for real-life applications. The purpose of this block, depicted in Fig. 3, is twofold. A first sector of the memory, blocks AB and OB in the figure, stores the machine code of the algorithm to be implemented. Every instruction here consists of a 2 × 32-bit digital word which defines the state of several reconfiguration switches used in the PEs and configures the I/O digital port.

[1]These are experimental results.

TABLE I
MAXIMUM INL AND DNL FOR CONVERTERS IN THE CHIP

Converter	DNL(LSBs)	INL(LSBs)
Reference	0.16	0.4
Weight	0.21	0.7
D/As in I/O Port	0.48	0.48
A/Ds in I/O Port	0.48	0.48

Statistical results obtained from 10 samples.

A second sector of the memory is employed to store, in 8-bit format, 32 sets of 24 coefficients which determine internal analog references (the chip does not require any analog input during its normal use), and codifies the electrical signals associated to the coefficients in the convolution masks to be applied (weights). Since processing blocks in the PE are basically analog, the outputs of this second sector of the memory are connected to 24 8-bit D/A converters, implemented as a resistor ladder and an analog multiplexer [9]. The outputs of these converters drive a spatially distributed bank of buffers [10] which transmit the corresponding analog value to each position in the array. Table I shows differential (DNL) and integral nonlinearity (INL) measurements of these converters, which demonstrate that the required 8-bit accuracy has been achieved.

C. I/O Port

This structure allows the chip to interchange input and output images when connected as part of a conventional processing platform. Images are provided in digital format, using 8-bit codes for pixels, through a 32-bit data bus. Communications between the chip and the host are carried out by means of very simple handshaking protocols with all the addressing and timing signals being internally generated. The digital port employs one data converter per column which can be reconfigured to operate both as A/D and as D/A. It includes calibration blocks for fixed pattern noise (FPN) attenuation [11], [12] such as correlated double sampling (CDS), among others. Although solutions based on using one converter per column are not commonplace in conventional optical sensor arrays of
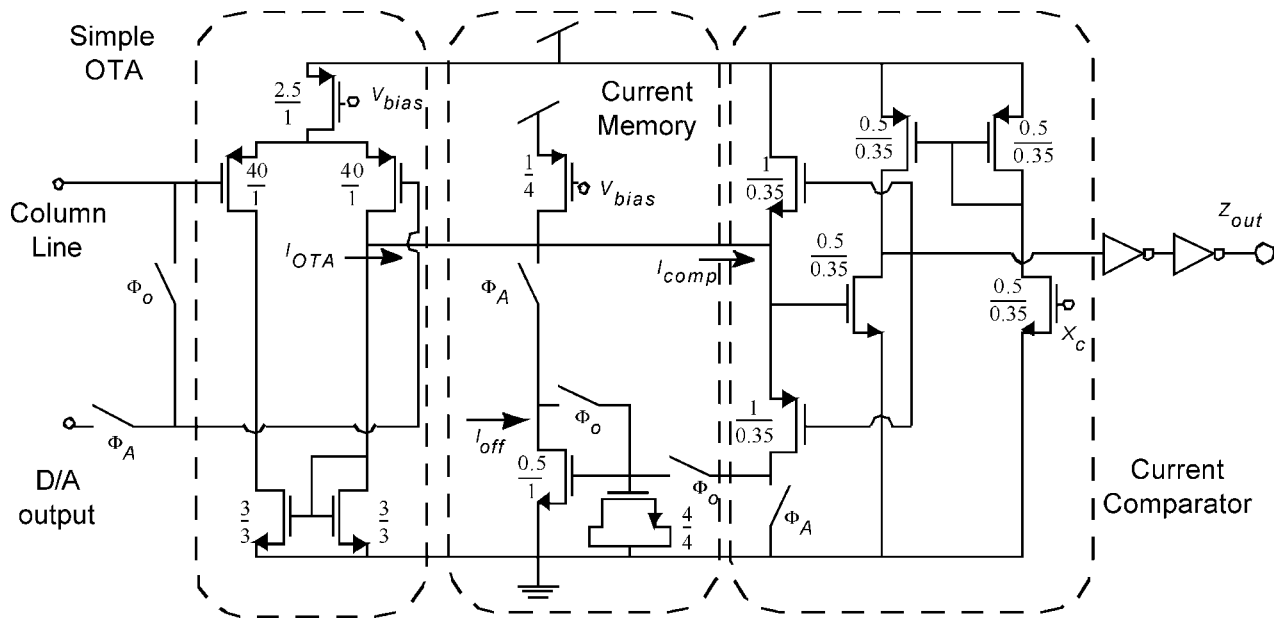
Fig. 4. Schematic of the comparator in the A/D successive approximation converter. A basic OTA, with a current memory for offset compensation, drives a nonlinear feedback current comparator [15].

moderate size because of the need to match the pixel pitch and that of the analog-to-digital (A/D) converters [13], they are well suited in our case since, on the one hand, they provide the maximum attainable frame rate due to its intrinsically parallel operation and, on the other hand, the converter must have the width of the PE and not that of a small pixel.

D/A converters, employed for image loading, are very similar to those in the program block, i.e., a resistor ladder, with resistors being implemented by not-silicided N-type diffusions, whose taps are connected to an analog multiplexer. The need to match the pitch of the converter to that of a single column (width of a PE), has made it necessary for the ladder to be laid out as 16 segments of 16 unitary resistors. In order to reduce the FPN produced by mismatching [14] in the elements of the ladder, the end taps of each of the 16 segments of 16 resistors in the ladder are connected to the equivalent taps in the converters on its right and left sides. This ensures that the output voltage of the ladders will coincide every 16 codes, and therefore it will reduce the discrepancies due to mismatch.

A/D converters, employed for image output, follow a successive approximation architecture [9]. Although these architectures might not be optimal for low resolution and moderate speed, they allow us to reuse the D/A converter employed for image loading. In addition, due to the parallelism of the operation of the converters, the control unit is shared by the 128 converters. Thus, in practice, the only modification needed to transform the D/A into an A/D is to add a comparator[2] and some reconfiguration switches. Regarding comparators, their input-referred offset voltages and their spatial distribution across the 128 converters might introduce additional FPN which, due to the special sensitivity of the human eye to stripe-like patterns, could strongly degrade the final quality of the output image. In order to attenuate these discrepancies, each comparator incorporates its own offset-correction block.

[2]Since the digital register was already needed to introduce the images.

Moreover, since calibration is performed just before converting every new pixel, such calibration may also be considered the first step of the successive approximation algorithm. Another interesting feature in this solution is that, since a new calibration is performed each time new data is going to be converted, signal dependent mismatching effects [3] can also be attenuated.

Fig. 4 shows the schematic of the comparator in the converter. It has two operating modes, as follows.

- The first one, corresponding to ϕ_A active, is the normal operating mode. In this mode, the comparator decides whether the voltage at the column line is larger or smaller than that provided by the D/A in every step of the successive approximation sequence, and its output fills the successive approximation register.
- The second mode, corresponding to ϕ_o active, defines the calibration phase. In this mode the offset current (output referred) of the input operational transconductance amplifier (OTA) in the comparator is sampled and stored in a conventional current memory.

The front-end of the comparator is a basic CMOS transconductance amplifier whose output is connected to a current memory in order to store its output offset current, and to a nonlinear feedback high-speed current comparator [15]. During the calibration phase, both inputs of the OTA are connected to the voltage level at the column line. Mismatching effects make the output current of the OTA nonnull. This current flows onto the calibration current memory, through the input branch of the current comparator, where it is stored. Note that since both inputs of the OTA are connected to the pixel value, common mode dependent errors are attenuated. Note also that, due to the negative feedback around the comparator, the voltage at its input node remains close to the inverter quiescent point [15].

[3]For instance, errors induced by differences in the body factor of the input transistors in the comparator do actually depend on the input common mode, and hence, are content-sensitive.
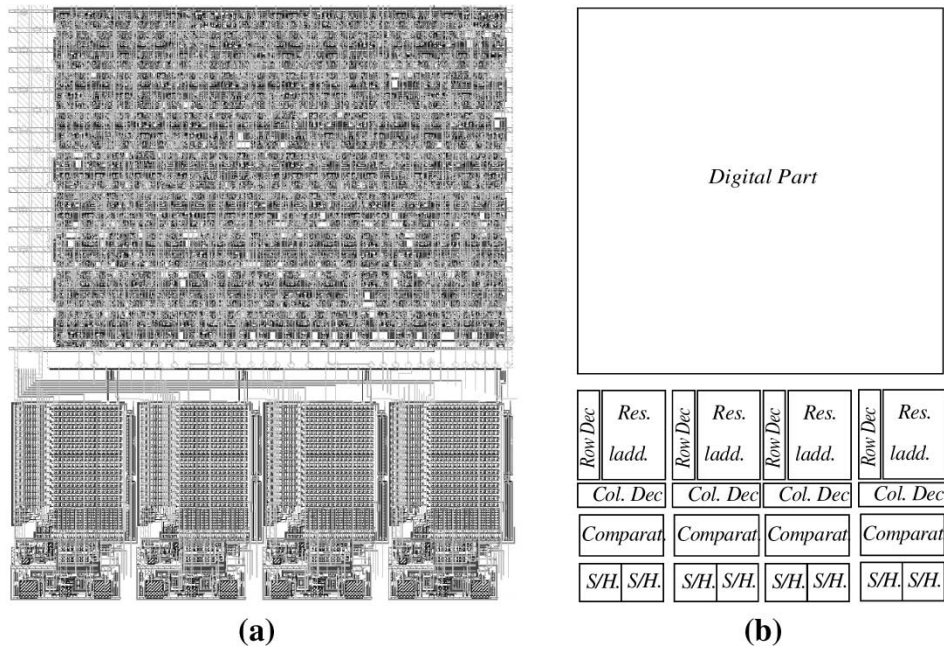
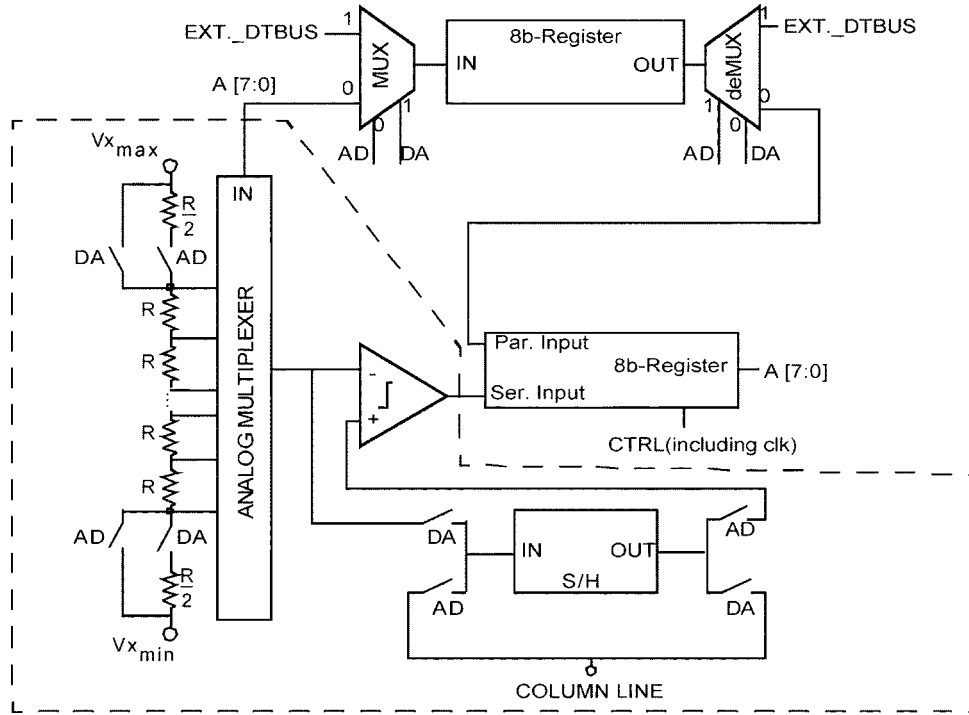Fig. 5.   The I/O 4-columns group. (a) Layout. (b) Floorplan.

Fig. 6.   Block diagram of an I/O block. A resistor ladder plus an analog multiplexer perform the D/A conversion. By controlling some switches and inserting a comparator the system becomes a successive approximation D/A converter.

Therefore, errors due to the finite output impedance of the simple OTA are highly reduced. During normal operation, such offset current is drained from the current provided by the OTA and, therefore, the current flowing onto the current comparator is, in first-order approach, offset-free.

Since the chip uses a 32-bit data bus and 8-bit coding for pixels, information in the digital data bus corresponds to pixels in four consecutive columns. For this reason, the digital part of the I/O block in each four columns are laid out together (see

Fig. 5) while only the analog part, in the box in Fig. 6, must match the pitch of the column. Experimental results, shown in Table I, confirm that the converters in this I/O block fulfill the requirements of 8-bit accuracy.

III. PROCESSING ELEMENT

Fig. 7 shows the block diagram of the PE [8]. Arrows indicate the flow of information. As may be seen, the PE contains
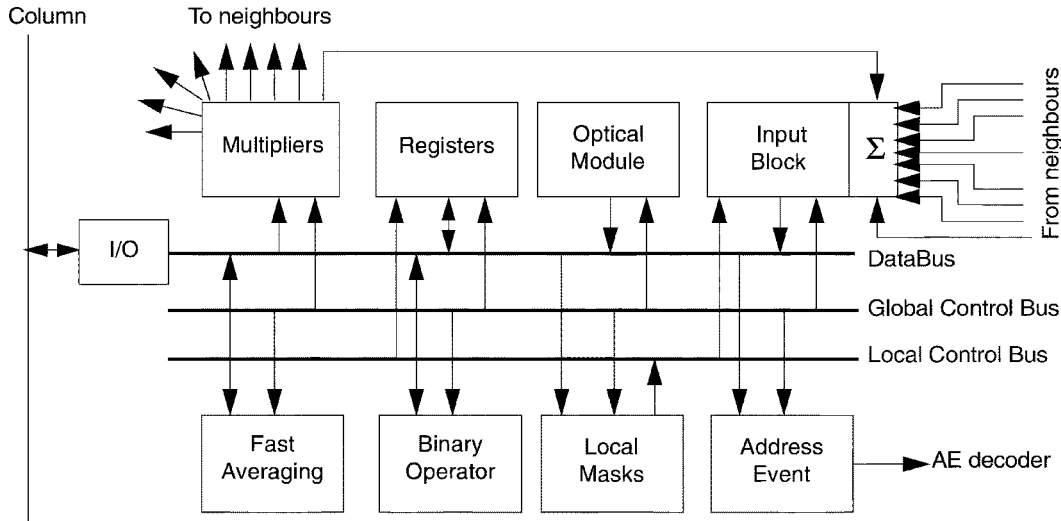
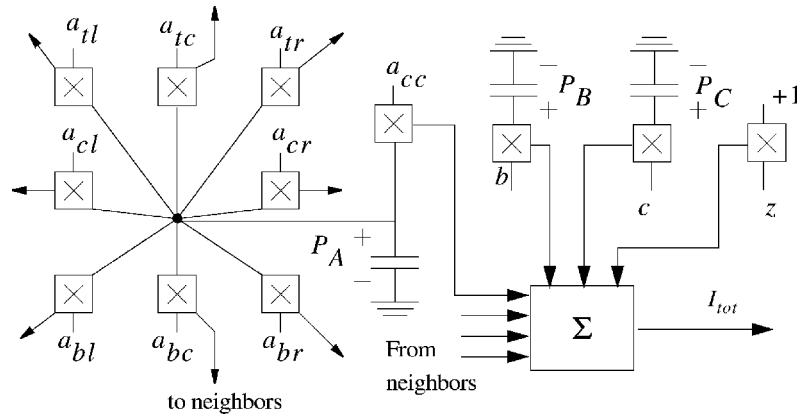Fig. 7.   Block diagram of the processing element.

Fig. 8.   The bank of programmable multipliers.

different building blocks which share a local data bus, a global control bus (Inst. Bus in Fig. 2) and a local control bus. Since the processing is done in the analog domain, such a local data bus is a single wire, which allows for great savings of routing area. Besides the main processing kernel for running 2-D 3 × 3 linear filters, each PE incorporates the following functional operators: 1) an analog register with capacity for 8 pixel values which are stored with an equivalent 8-bit resolution; 2) a two-input one-output programmable Boolean operator; 3) a multimode optical sensor [16]; 4) an address event downloading module; 5) a very fast averaging module which, when used, provides the mean value of an input image in less than 200 ns, which is very useful, for instance, for adapting the exposure in a supervised way when using integrating light sensing schemes; and 6) two local flags to allow for the conditional execution of two operations: running a convolution and updating an analog register.

A. Programmable Processing Kernel

Each PE can update its data depending on values stored in its internal analog registers, as needed in pixel-wise *cosmetic* operations [3]. Pixel values can also be updated by the weighted aggregation of information stored on the nearest PEs, as needed to perform basic spatial filtering operations [3]. The implementation of 2-D linear convolutions is accomplished by a bank of

programmable analog multipliers. Multipliers employ a single transistor technique elsewhere reported in [8], [17], [18]. Multipliers are driven by two voltages, determining the pixel and the weight respectively, and they provide an output current. The bank of multipliers, depicted at the conceptual level in Fig. 8, is driven by three pixel values, P_A, P_B, and P_C, in such a way that the input current which flows toward the input block, in Fig. 7, can be expressed as

$$I_{\text{tot}} = \mathbf{A} \bullet \mathbf{P_A} + b \cdot P_B + c \cdot P_C + z + I_{\text{off}} \qquad (1)$$

where $\mathbf{A}$ and $\mathbf{P_A}$ are defined as

$$\mathbf{A} = \begin{bmatrix} a_{tr} & a_{tc} & a_{tl} \\ a_{cr} & a_{cc} & a_{cl} \\ a_{br} & a_{bc} & a_{bl} \end{bmatrix} \quad \mathbf{P_A} = \begin{bmatrix} P_{A_{tl}} & P_{A_{tc}} & P_{A_{tr}} \\ P_{A_{cl}} & P_{A_{cc}} & P_{A_{cr}} \\ P_{A_{bl}} & P_{A_{bc}} & P_{A_{br}} \end{bmatrix}. \qquad (2)$$

The operator ($\bullet$) accounts for the convolution product of these matrices. I_{off} is an offset contribution (to be cancelled afterwards), which is intrinsic to the use of the single transistor multiplier [8], [17], [18]. Finally, indexes refer to PE position; thus, c stands for center, r for right, b for bottom, and t for top. The current in (1) is collected by an input block, in Fig. 9, through a virtual ground. Such a virtual ground plays an important role since the single transistor multiplier circuitry requires one of the diffusion terminals of the transistor to be connected to a
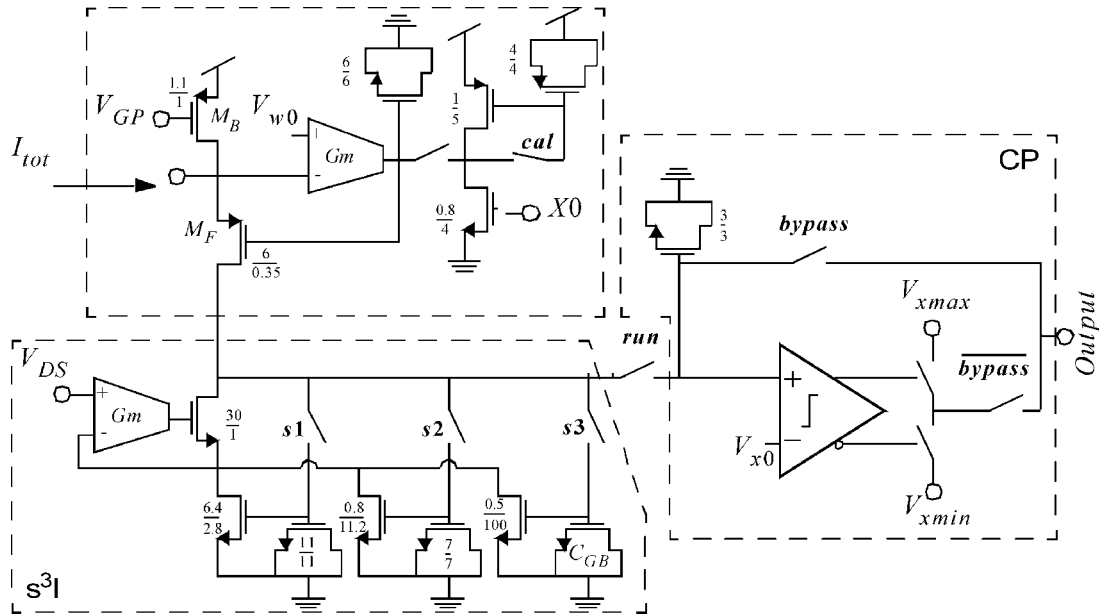
Fig. 9.   Schematic of the input block in Fig. 7.

fixed voltage reference. A simple current conveyor is employed for such task. The negative feedback loop set by the amplifier [4] and transistor M_F fixes the voltage at the input terminal of the block, whereas the total input current, plus the biasing current provided by M_B, flow through M_F toward an s^3I current memory and a basic current-processing block. The current conveyor also includes an offset calibration current memory to reduce the input-referred offset voltage of the OTA, which directly appears in the signal path and produces space-variant errors. The offset term generated by the multipliers, plus the biasing current from M_B, are stored, and then drained away from the input current, in an s^3I current memory, which is an extension of the s^2I approach in [19]. We will not discuss here the needs which motivated selecting such a memorization technique. Such current memory is designed to store a maximum input current of 15 μA and to produce a maximum memorization error of 2 nA.

Once offset and biasing currents are stored and eliminated from the total input current, the current which flows toward the current-processing block, CP in Fig. 9, is

$$I_{\text{in}} = \mathbf{A} \bullet \mathbf{P_A} + b \cdot P_B + c \cdot P_C + z. \qquad (3)$$

This current can be steered either to the global data bus, by asserting the global control line *bypass*, or to the input of a capacitive-input current comparator [15]. The output of this comparator determines whether we have to set the internal data bus to the maximum pixel value $V_{x\text{max}}$ or to the minimum pixel value $V_{x\text{min}}$. Hence, two different situations may appear depending on the state of the signal *bypass*.

- If it is not asserted, the voltage delivered to the data bus will correspond to the sign of I_{in}

$$\text{sign}\left(\mathbf{A} \bullet \mathbf{P_A} + b \cdot P_B + c \cdot P_C + z\right) \qquad (4)$$

and will produce a black-and-white image.

[4]A basic five-transistor OTA in our case.

- If it is asserted, analog current I_{in} can be routed to any of the capacitors [5] associated to the pixels and the output would be a grayscale image.

In the latter case above, the specific pixel capacitor(s) receiving I_{in} is (are) selected by the user through the activation of some bits in the global instruction. Obviously, the processing function being implemented will be described by a set of state equations whose actual expressions depend on the selected integrating capacitor, and also on the state of *bypass*. Therefore, different kinds of processing kernels are available. For instance, to run a convolution, the coefficients are entered into $\mathbf{A}$, the image into $\mathbf{P_A}$, and we make $c = z = 0$. Then, by enabling *bypass* and routing the input current toward C_B we get

$$C_B \frac{dP_B}{dt} = bP_B + \mathbf{A} \bullet \mathbf{P_A} \qquad (5)$$

whose steady state results in

$$P_B = -\frac{\mathbf{A}}{b} \bullet \mathbf{P_A}. \qquad (6)$$

Consider now that the input current is routed to C_A, that all but the central entries of matrix $\mathbf{A}$ in (2) are null, and that this central coefficient is $a_{cc} = -1$. This yields

$$P_A = b \cdot P_B + c \cdot P_C + z \qquad (7)$$

as it corresponds to the implementation of grayscale arithmetic operations on pixel values.

Note also that when updating of P_A is allowed, with some nonnull entry in $\mathbf{A}$, PEs become dynamically coupled and cellular neural network (CNN) temporal evolutions [6] may also be emulated.

Regarding the accuracy required for the implementation of (3), design is made in such a way that:

- registers provide pixels values with 8-bit equivalent precision;

[5]Also combinations of them.

Fig. 10. (a) Schematic of the optical input module. (b) Cross-section of the photosensor.

- multipliers are designed for a maximum harmonic distortion of 4% and 8-bit equivalent spatial resolution (3σ);
- input impedance of the current conveyor in Fig. 9 is designed to be much smaller than the combined impedance of the multipliers, and biasing current source, connected to it. Hence, summation is implemented almost free of errors.

B. Optical Input

The optical input module, shown in Fig. 10, [16], consists of a multimode sensor in which both the physical device employed as a photosensor and the transduction mechanism are programmable. A P-diff/N-well diode, an N-well/P-subs diode, or a P-diff/N-well/P-subs phototransistor are available. Though having such a complex sensor does reduce the fill-factor, the lack of previous experimental results with photosensors in this technology obliged us not to take a risk in selecting a single light-sensing device.

The selection of the photosensitive device is carried out by global instruction signals. Regarding the mechanism used to transform the photogenerated current into a voltage level, both linear integration and logarithmic compression [6] are available with proper definition of the digital instructions. Special attention has been paid to the design of the buffer in the photosensor, which consists of a basic OTA with pMOS input transistors (having source and bulk connected), and incorporating the required level shifting to accommodate the photosensor output to the pixel ranges. Moreover, since optical readouts are to be performed only occasionally, the buffer remains disconnected from the power supply when it is not used, thus saving on power consumption. In integrating modes, the sensor has a conversion factor of 18 $\mu V/e^-$ whereas the N-well which contains it occupies 9.8×9.8 μm^2, that is, 2% of the PE area.

IV. EXPERIMENTAL RESULTS

Table II shows the main electrical and physical characteristics of the chip. Its functional testing is performed by using a specific hardware–software environment designed by Analogic Computers Ltd.[7] The hardware part of the system contains various

[6]By driving a pMOS load in the subthreshold region.

[7]http://www.analogic-computers.com

TABLE II
CHARACTERISTICS OF THE CHIP

Technology	0.35 μm 5M-1P
# Array	128 × 128
# Transistors	3,748,170
# Transistors per PE	198
PE Dimensions (μm)	75.5 × 73.3
PE Density	180 cells/mm²
Conversion Factor of the Sensor	18μV/e^-
Fastest Time Constant for Convolutions	135ns
Power per Cell (at maximum speed)	180μW @ 3.3V
Precision in 8-bit equivalent convolutions.	> 99%
I/O Digital Rate	121 Mbytes/s
Conv. Masks in Memory	32
Instructions in Memory	4096
Chip Size (mm)	11.9 × 12.2

Fig. 11. Prototyping platform designed by Analogic Computers Ltd.

layers of boards of the same size, which connects by stacking. The first layer, shown in Fig. 11, hosts the chip, while the others are intended to provide the program, power, and data, and to accommodate inputs/outputs to/from the chip and from/to the PCI bus.
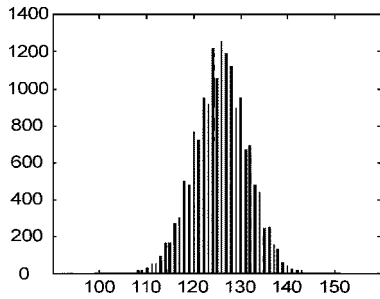
Fig. 12.   Fixed-pattern noise histogram. It has a 5 LSB standard deviation and is mainly due to the spatial distribution of the offsets of the on-PE readout buffer.
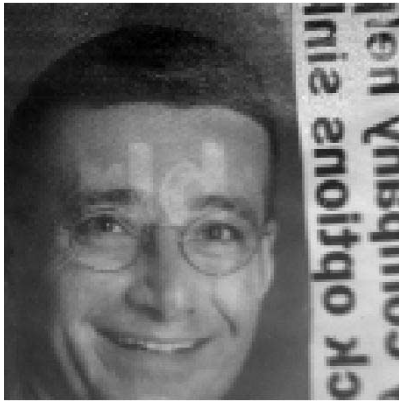

Fig. 13.   Result of an optical acquisition in typical conditions. Illumination is provided by a reading lamp at about 40 cm. Exposition time is 1 ms, selected sensor is the well-substrate diode.

A. Optical Input

As already mentioned, the optical input is a reconfigurable sensor with seven different sensing modes. Using such a complex, reconfigurable optical sensor [16] might have led to undesirably large FPN figures. However, it has not been the case in practice. First, because we designed both the schematic and the layout of the readout buffer by paying special attention to mismatching effects. We employed pMOS transistors in the differential pair with connected source and bulk just to avoid errors due to the random distributions of the body factor, which introduces signal-dependent FPN. Moreover, we took advantage of the analog registers to store and subtract the offset of the readout buffer already at the PE level—PE-level CDS. Finally, as mentioned earlier, the successive approximation A/D conversion algorithm also includes some steps to calibrate the offset of the comparators within the converter, thus helping to cancel most of the column-wise FPN.

Fig. 12 shows the histogram of FPN in one of the samples of the chip. Here, only column-wise FPN reduction was enabled while PE circuitry for FPN attenuation was disabled, in order to separate sensor performance from that of the A/D converter. Standard deviation amounts to 5 LSBs.

Fig. 13 shows the result of capturing an image by employing the well/substrate diode [8] and 1 ms of exposure. Light comes from a 60-W reading lamp at 30 cm. FPN suppression was fully enabled. Precise optical characterization is currently running.

[8]This has provided the larger sensitivity.

B. Image Processing Examples

Fig. 14 shows the experimental results of implementing two well-known 3×3 image processing kernels, namely a low-pass filter [Fig. 14(b)] and a horizontal Sobel filter [Fig. 14(c)] [3]. Using the algorithm in [20][9], and considering MATLAB results as ideal, the quality index of these images results in 99.8% and 99.6%, respectively.[10] Convolution kernels must run [11] for about 1 μs before the transient evolutions settle. Outputs can be stored within the processing element in any of its local pixel memories (as needed if they are going to be used in a complex algorithm) without effective degradation, keeping the 8-bit accuracy for 1 ms. Moreover, this accuracy is also kept after eight internal or external readouts. This is more than enough for standard algorithms. Complete uploading/downloading of an image only takes 135 μs.

C. Ultra-High Frame-Rate Visual Inspection

This chip has been incorporated by Analogic Computers Ltd. to its Bi-i Product, Binocular Eye. Recent developments demonstrate the capability of the chip to be the main component in a visual inspection system for classification and analysis of pills. The experimental setup, shown in Fig. 15(a), contains a rotating wheel moving at an equivalent linear speed of 1.3 m/s, a halogen lamp, and the Bi-i. The system is able to acquire and classify the objects at a rate of more than 10 000 FPS, thanks to the strong illumination. Fig. 15(b) shows the user panel developed by Analogic Computers.

V. Discussion

Table III shows an overview of representative chips using different alternatives for vision processing. The first row corresponds to a high-end DSP, the second row corresponds to an SIMD mixed-signal vision processor, the third row corresponds to a multiple instruction multiple data processor, and the others to SIMD-CNN processors. It is important to see how fully parallel analog and mixed-signal solutions obtain larger computation performances than a conventional digital solution.[12] Of course, those analog and mixed-signal chips are not capable of performing complex mathematical operations with 32-bit floating point accuracy. However, since early vision operations can be efficiently executed at moderate computational precision [4], analog operators may reach efficiency levels far above those of digital ones. By examining the information of the table, one easily sees how the chip reported in this paper has improved what was obtained in [17], that is, basically, less power consumption per PE and faster operation. Looking at the inputs in the table, most of them can be readily understood except those related to the peak computing power, referred to as Speed (OPS) in the table. For other authors, we simply picked data from the mentioned reference. In our chips, an equation is employed which combines the number of operations, additions and products, the time constant of the process, and the number

[9]It can be downloaded from http://www.cns.nyu.edu/~lcv/ssim/

[10]Considering that MATLAB results are also confined to the 8-bit range.

[11]Before running a kernel, different calibrations are optionally executed. When executing all of them the processing time increases up to 4 μs.

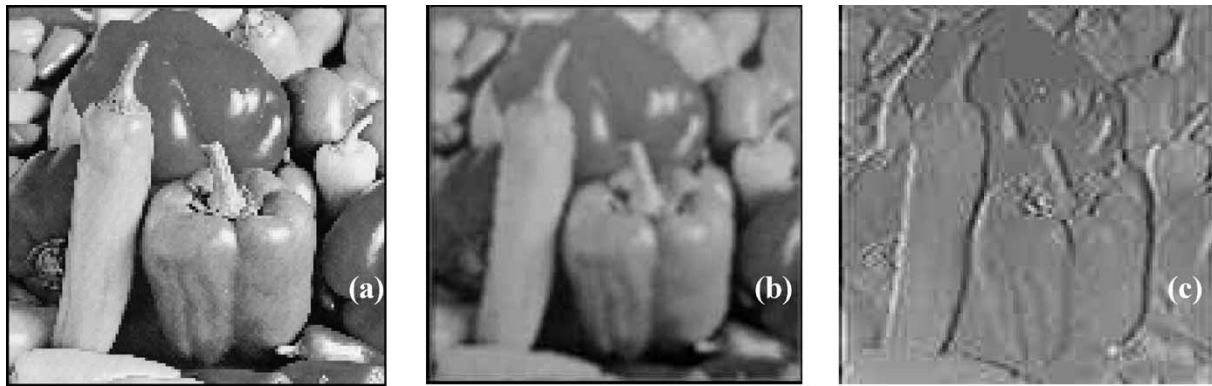[12]Of course, the DSP does many other things that our chip cannot do.

Fig. 14. Some image processing examples. (a) represents the input image. (b) shows the result of a 3 × 3 averaging filter. (c) shows the result of a Sobel filter. Quality of the results is checked with the algorithm in [20] and results in 99.8% and 99.6%, respectively.

Fig. 15. Visual inspection system by Analogic Computers Ltd.

TABLE III
REPRESENTATIVE VISION CHIPS

Ref.	Tech. (micron)	Style	Array	Pixel (format)	Memory	Speed(OPS)	Speed/Pow. (OP/Joule)	Photosensors
[21]	0.12	DSP	d.n.a.	Digital	32kB (L1) 8MB (L2)	4.8×10^9	3.2×10^9	No
[4]	0.6	SIMD	21 × 21	Analog	6 Gray	0.5×10^9	12×10^9	Yes
[7]	1.2	MIMD	80 × 78	Analog	N. A.	12.4×10^9	N.A.	Yes
[22]	0.5	CNN	48 × 48	Analog	4 Binary	500×10^9	1.6×10^{12}	No
[17]	0.5	CNN	64 × 64	Binary	4 Gray 4 Binary	47.5×10^9	39.5×10^9	Yes
This chip	0.35	SIMD/ CNN	128 × 128	Analog	8 Gray 4 Binary	330×10^9	82.5×10^9	Yes

of time constant units to get settling errors below a given limit. Thus, for convolutions, this yields

$$\text{OPS}_{\text{conv}} = \frac{(N_{\text{add}} + N_{\text{prod}}) \cdot N_c}{\tau_{\text{conv}} \cdot (n+1) \cdot Ln(2)} \qquad (8)$$

with N_c being the number of elements in the array (128×128), N_{add} the number of additions(8 in 3×3 conv.), N_{prod} the number of products (9 in 3×3 conv.), n the accuracy of the settling process in an equivalent number of bits, and τ_{conv} the time constant of the process in (5), about 135 ns for the largest b.

VI. Conclusion

Experimental evidence about the suitability of using mixed-signal vision sensor/processor chips to solve real time *low-level* vision problems has been shown in this paper. Such a possibility has been illustrated through a chip which contains an array of 128×128 processors. The chip is able to capture and process at about 1000 FPS under room illumination conditions. It can store different images on-chip and can run user-defined 3×3 convolution masks. Input images, or results can be arbitrarily combined on-chip by means of any linear operation—constant scaling, summation, and substraction—or a two-inputs Boolean function. The chip works at 3.3 V and provides peak computing figures of 330 GOPS, 3.6 GOPS/mm^2, and 82.5 GOPS/W, while achieving an equivalent accuracy close to 8 bits. Output images are provided to the hosting system in digital format at a maximum rate of 121 36 MB/s.

References

[1] R. H. Masland, "The fundamental plan of the retina," *Nature Neurosci.*, vol. 4, pp. 877–886, Sept. 2001.

[2] B. Roska and F. Werblin, "Vertical interactions across ten parallel, stacked representations in the mammalian retina," *Nature*, no. 410, pp. 583–587, Mar. 2001.

[3] B. Jahne, *Handbook of Computer Vision and Applications*, B. Jahne, H. Haubecker, and P. Geibler, Eds.   London, U.K.: Academic, 1999.

[4] P. Dudek, "A programmable focal-plane analogue processor array," Ph.D. dissertation, Univ. Manchester Inst. Sci. Technol., Manchester, U.K., 2000.

[5] D.A. Martin, H. S. Lee, and I. Masaki, "A mixed signal array processor with early vision applications," *IEEE J. Solid-State Circuits*, vol. 33, pp. 497–502, Mar. 1998.

[6] T. Roska and L. O. Chua, "The CNN universal machine: An analogic array computer," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 163–173, Mar. 1993.

[7] R. Etienne-Cummings, Z. K. Kalayjian, and D. Cai, "A programmable focal plane MIMD image processor chip," *IEEE J. Solid State Circuits*, vol. 36, pp. 64–73, Jan. 2001.

[8] G. Liñán, "Design of low-power mixed-signal programmable vision chips," Ph.D. dissertation, Univ. Seville, Spain, 2002.

[9] B. Razavi, *Principles of Data Conversion System Design*.   New York: IEEE Press, 1995.

[10] G. Liñán-Cembrano, A. Rodríguez-Vázquez, R. Carmona, S. Espejo, and R. Domínguez-Castro, "Analog weight buffering strategy for CNN chips," in *Proc. Int. Symp. Circuits and Systems (ISCAS 2003)*, vol. 3, Bangkok, Thailand, May 2003, pp. III-522–III-525.

[11] C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: Autozeroing, correlated double sampling, and chopper stabilization," *Proc. IEEE*, vol. 84, pp. 1584–1614, Nov. 1996.

[12] H. Wey and W. Guggenbuhl, "An improved correlated double sampling circuit for low noise charge-coupled-devices," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 1559–1565, Dec. 1990.

[13] G. Torelli, L. Gonzo, M. Gottardi, F. Maloberti, A. Sartori, and A. Simoni, "Analog-to-digital conversion architectures for intelligent optical sensor arrays," in *Proc. SPIE*, 1996, pp. 254–264.

[14] M. J. M. Pelgrom, A. C. J. Duinmaijer, and A. P. G. Welbers, "Matching properties of MOS transistors," *IEEE J. Solid-State Circuits*, vol. 24, pp. 1433–1440, Oct. 1989.

[15] A. Rodríguez-Vázquez, M. Delgado-Restituto, J. L. Huertas, and F. Vidal, "Synthesis and design of nonlinear circuits," in *The Circuits and Filters Handbook*, W. K. Chen, Ed.   New York: IEEE Press, 1995.

[16] G. Liñán, A. Rodríguez-Vázquez, E. Roca, S. Espejo, and R. Domínguez-Castro, "A versatile sensor interface for programmable vision systems-on-chip," in *Proc. SPIE Electronic Imaging 2003*, Santa Clara, CA, Jan. 2003, pp. 38–47.

[17] G. Liñán, S. Espejo, R. Domínguez-Castro, and A. Rodríguez-Vázquez, "ACE4k: An analog I/O 64×64 visual microprocessor chip with 7-bit analog accuracy," *Int. J. Circuit Theory Applicat.*, vol. 30, no. 2–3, pp. 89–116, Mar. 2002.

[18] R. Carmona Galán, F. Jiménez-Garrido, R. Domínguez-Castro, S. Espejo, T. Roska, C. Rekeczky, and A. Rodríguez-Vázquez, "A bio-inspired two-layer mixed-signal flexible programmable chip for early vision," *IEEE Trans. Neural Networks*, vol. 14, pp. 1313–1336, Sept. 2003.

[19] J. B. Hughes and K. W. Moulding, "S^{2}I: A two-step approach to switched-currents," in *Proc. IEEE Int. Symp. Circuits and System*, May 1993, pp. 1235–1238.

[20] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, pp. 600–612, Apr. 2004.

[21] TMS320c6414 DSP. Texas Instruments, Inc., Dallas, TX. [Online]. Available: http://www.ti.com

[22] A. Paasio, A. Dawidziuk, K. Halonen, and V. Porra, "Minimum size 0.5 μm CMOS programmable 48×48 CNN test chip," in *Proc. Eur. Conf. Circuit Theory and Design*, Budapest, Hungary, Sept. 1997, pp. 154–156.

Gustavo Liñán Cembrano (A'03) received the Licenciado and Doctor (Ph.D.) degrees in physics, in the speciality of electronics, from the University of Seville, Spain, in 1996 and 2002, respectively.

In 1995, he was a granted student of the Spanish Ministry of Education at the Institute of Microelectronics of Seville CNM-CSIC. From 1997 to 1999, he received a doctoral grant at the Institute of Microelectronics of Seville CNM-CSIC funded by the Andalusian Government. Since 2000, he has been an Assistant Professor of the Department of Electronics and Electromagnetism, School of Engineering, University of Seville, and at the Faculty of Physics. His main areas of interest are the design and VLSI implementation of massively parallel analog/mixed-signal image processors.

Dr. Liñán Cembrano received the Best Paper Award 1999 from the *International Journal of Circuit Theory and Applications*. He was a corecipient of the Most Original Project Award, of the Salvà i Campillo Awards 2002, conceded by the Catalonian Association of Telecommunication Engineers.

Angel Rodríguez-Vázquez (M'80–F'96) is a Professor of electronics with the Department of Electronics and Electromagnetism, University of Seville, Spain. He is also a Member of the Research Staff of the Institute of Microelectronics of Seville (IMSE-CNM, CSIC), where he is heading a research group on Analog and Mixed-Signal VLSI. His research interests are in the design of analog interfaces for mixed-signal VLSI circuits, CMOS imagers and vision chips, neuro-fuzzy controllers, symbolic analysis of analog integrated circuits, and optimization of analog integrated circuits.

Dr. Rodríguez-Vázquez served as an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART I (TCAS-I) from 1993 to 1995, as Guest Editor of the IEEE TCAS-I special issue on Low-Voltage and Low-Power Analog and Mixed-Signal Circuits and Systems in 1995, as Guest Editor of the IEEE TCAS-II special issue on Advances in Nonlinear Electronic Circuits (1999), and as Chair of the IEEE Circuits and Systems Society Analog Signal Processing Committee in 1996. He was a corecipient of the 1995 Guillemin-Cauer award of the IEEE Circuits and Systems Society and the Best Paper Award of the 1995 European Conference on Circuit Theory and Design. In 1992 he received the Young Scientist Award of the Seville Academy of Science. In 1996 he was elected as a Fellow of the IEEE for "contributions to the design and applications of analog/digital nonlinear ICs."

Ricardo Carmona Galán (A'02) received the Licenciado and Doctor (Ph.D.) degrees in physics, in the speciality of electronics, from the University of Seville, Spain, in 1993 and 2002, respectively. From 1994 to 1996, he was a granted student at the National Centre for Microelectronics at Seville, funded by IBERDROLA S.A. From 1996 to 1998, he was a Research Assistant at the Electronics Research Laboratory, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley. He is a member of the Department of Analog Design in the Microelectronics Institute of Seville (CNM-CSIC). Since October 1999, he has been an Assistant Professor of the Department of Electronics and Electromagnetism, School of Engineering, University of Seville. His main areas of interest are linear and nonlinear analog and mixed-signal integrated circuits, in particular, the design and VLSI implementation of cellular neural networks and analog memory devices for real-time image processing and vision chips.

He received the Best Paper Award 1999 from the *International Journal of Circuit Theory and Applications*. He is a corecipient of the Most Original Project Award of the Salvà i Campillo Awards 2002, conceded by the Catalonian Association of Telecommunication Engineers.

F. Jiménez-Garrido received the B.S. degree in physics in 1998 and the B.S. degree in electronic engineering in 2002 from the University of Seville, Spain. He is currently working toward the Ph.D. degree in the Department of Electronics and Electromagnetism, University of Seville.

Since 1999, he has been with the Department of Analog Circuit Design, Spanish Microelectronics Center (Institute of Microelectronics of Seville, IMSE). His research interests are in linear and nonlinear analog and mixed-signal integrated circuits for image processing and communication devices.

Servando Espejo received the Licenciado en Física degree, the M.S. equivalent in microelectronics, and the Doctor en Ciencias Físicas degree from the University of Seville, Spain, in 1987, 1989, and 1994, respectively.

From 1989 to 1991 he was an intern with AT&T Bell Laboratories, Murray Hill, NJ, and an employee of AT&T Microelectronics of Spain. He is currently a Professor in electronic engineering in the Department of Electronics and Electromagnetism, University of Seville, and also with the Department of Analog Circuit Design of the Spanish Microelectronics Center. His main areas of interest are linear and nonlinear analog and mixed-signal integrated circuits, including neural networks electronic realizations and theory, vision chips, massively-parallel analog processingsystems, chaotic circuits, and communication devices.

Dr. Espejo was a corecipient of the 1995 Guillemin-Cauer Award of the IEEE Circuits and Systems Society, the Best Paper Award of the 1995 European Conference on Circuit Theory and Design, and the Best Paper Award of the 1999 *International Journal of Circuit Theory and Applications*.

Rafael Domínguez-Castro received the Licenciado en Física Electrónica degree in 1987 and the Ph.D. degree in 1993 from the University of Seville, Spain.

Since 1987, he has been with the Department of Electronics and Electromagnetism, University of Seville, where he is currently an Associate Professor. He is also with the Institute of Microelectronics of Seville (IMSE-CNM, CSIC), where he is a member of a research group on Analog and Mixed-Signal VLSI. His research interests are in the design of embedded analog interfaces for mixed-signal VLSI circuits, design of CMOS imagers and CMOS focal plane array processors, and development on CAD for automation of analog design.

Dr. Domínguez-Castro was a corecipient of the 1995 Guillemin-Cauer Award of the IEEE Circuits and Systems Society and the Best Paper Award of the 1995 European Conference on Circuit Theory and Design.