# Smart-Pixel Cellular Neural Networks in Analog Current-Mode CMOS Technology

S. Espejo, A. Rodríguez-Vázquez, *Member, IEEE*, R. Domínguez-Castro, J. L. Huertas, and E. Sánchez-Sinencio

*Abstract*—This paper presents a systematic approach to design CMOS chips with concurrent picture acquisition and processing capabilities. These chips consist of regular arrangements of elementary *units*, called *smart pixels*. Light detection is made with vertical CMOS-BJT's connected in a Darlington structure. Pixel smartness is achieved by exploiting the Cellular Neural Network paradigm [1], [2], incorporating at each pixel location an analog computing cell which interacts with those of nearby pixels. We propose a current-mode implementation technique and give measurements from two 16 × 16 prototypes in a single-poly double-metal CMOS *n*-well 1.6-$\mu$m technology. In addition to the sensory and processing circuitry, both chips incorporate *light-adaptation* circuitry for automatic contrast adjustment. They obtain smart-pixel densities up to 89 units/mm$^2$, with a power consumption down to 105 $\mu$W/unit and image processing times below 2 $\mu$s.

## I. INTRODUCTION

COMMON architectures for image-processing systems use a front-end sensory plane with digital-encoding of the pixel values, and serial transmission of these digital data for subsequent processing using either ASIC's or general-purpose computers. Contrary to this approach, *smart-pixel* chips [3] incorporate an analog computing cell at each sensory point, achieving high speed and low area occupation in the combined sensory/processing functions by fully exploiting *parallelism*. The combined spatial distribution of sensory and processing circuitry eliminates the time required for data transmission from the sensory to the processing plane during the image acquisition process. In addition, in some image-processing applications, the relevant information contained in the output image can be described by a reduced number of variables, allowing a fast downloading of the results for subsequent evaluation.

CMOS technologies offer unique features for the design of smart-pixel chips. On one hand, MOS transistor operation under normal biasing in strong inversion is not drastically affected by incident light; on the other, photosensitive CMOS devices can be built by exploiting the many junction devices available in CMOS technologies [4]. However, previous approaches to CMOS design of smart-pixel chips lack generality, as they rely on implementation methods suitable for specific

S. Espejo, A. Rodríguez-Vázquez, R. Domínguez-Castro, and J. L. Huertas are with the Centro Nacional de Microelectrónica-Universidad de Sevilla, Edificio CICA, C/Tarfia sn, 41012-Sevilla, Spain.

E. Sánchez-Sinencio is with the Department of Electrical Engineering, Texas A&M University, College Station, TX 77843 USA.

applications. In some cases, the processing-task performed at each pixel does not imply collective computation [3], while most of the approaches for "pixel-smartness" are based on active implementations of resistive-grid networks [5], [6].

The paradigm of Cellular Neural Networks (CNN) [1], [2] is a very suitable framework for *systematic* design of parallel sensory-processing chips. On one hand, CNN's consist of regular arrangements of *cells*- -topologically identical to smart-pixel chips. On the other, their cells are only locally connected, and thus, require simple routing. Also, the vast body of literature on CNN theory and applications demonstrates outstanding features of this paradigm for array-processing [7]. In particular, resistive grids have recently been demonstrated as a particular CNN class [8].

No experimental smart-pixel CNN chips have been reported to date. This paper outlines a design approach using Darlington phototransistors and current-mode processing circuitry. It is based on a modified version of the original CNN model which enables optimum speed/power and area occupation in VLSI design [9], [10]. The sensors include an automatic adjustment circuitry which ensures proper behavior under different illumination conditions. Our proposals are demonstrated via two working smart-pixel chips, in a single-poly, 1.6-$\mu$m, *n*-well CMOS technology. In addition to their optical input, these chips exhibit much better area and speed/power figures than previous CNN implementations [11], [12].

Section II describes some general aspects of smart-pixel chips, and Section III outlines the proposed computation algorithm. Sections IV and V discuss the sensory and processing circuitry, respectively, and the experimental prototypes are described in Section VI.

## II. SMART-PIXEL CHIPS

In this paper, *pixel* denotes the elementary sensory unit used to detect pointwise light signals. These sensory units are realized in CMOS technology using any compatible junction device to generate a current whose value is an increasing function of the light intensity [3], [13], [14]. The acquisition of *two-dimensional* scenes requires pixels arranged onto regular *grids*, as shown in Fig. 1. Each pixel in this sensory plane generates a current $I_c$ which codifies a corresponding point of the input image, where the index $c \equiv (i, j)$ indicates the pixel at the $i$th row and $j$th column on the grid and varies over the whole grid domain $\mathcal{GD}(c \in \mathcal{GD})$. Thus, the whole image is captured into a matrix of currents $[I_c]$.
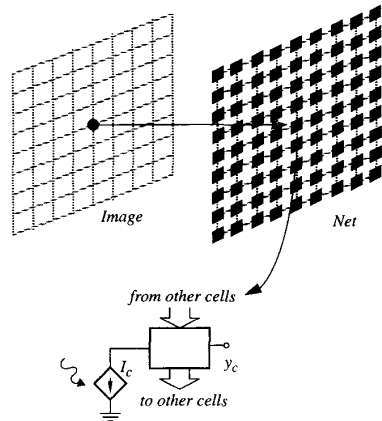
Fig. 1. Illustrating the core architecture of smart-pixel chips.

Fig. 1 illustrates the architecture of smart-pixel chips: each unit (also called *smart-pixel* or *cell* ) *senses* a point of the input image and *interacts* with the other units in the arrangement to perform parallel-processing tasks on the input current matrix $[I_c]$.

Smart-pixel chips are of strong practical interest for pattern recognition problems, to detect features of the input image. For example, Fig. 2 illustrates the task of detection of connected components (DCC), which consists of counting the number of connected pieces encountered by scanning an input image in a given direction [15]. Pattern recognition can be realized by processing the data obtained after performing this task in the directions shown in Fig. 2 [16], [17]. This data is contained in a few rows and columns at the grid borders. In addition to their usage for preprocessing tasks, smart-pixel chips are also useful as stand-alone units for nonintensive computation tasks such as halftoning [18], motion detection [19]–[21], range-finding [3], etc.

### III. THE CNN PARALLEL PROCESSING PARADIGM

As Fig. 1 illustrates, smart-pixel CNN chips consist of regular arrangements of *identical* units, each including a *photosensor* and an analog computing *cell*. Such an entity transforms the input image $[I_c]$ into an output matrix $[y_c]$ via a *dynamic* process of interactions among the computing cells. The distinctive feature of the CNN paradigm is that these interactions are *local*, limited for each cell to a reduced set of neighbors, located within a distance $r$ in the grid. In particular, there is a wide catalog of image processing tasks available for networks where parameter $r$ (called *neighborhood radius*) is unity—very appealing for VLSI implementations because connection among units is made by abutment, requiring no extra routing.

The dynamic computation process of CNN's, as proposed in [1], involves three variables per cell: (a) cell *state:* $x_c(t)$, which conveys cell energy information as a function of time; (b) cell *output:* $y_c(t)$, obtained from the cell state via a soft-limiter *piecewise-linear* transformation,

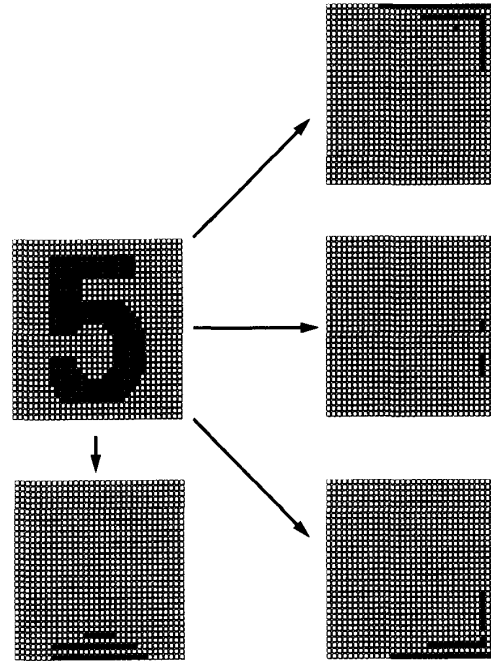$$y_c = f(x_c) \equiv \tfrac{1}{2}(|x_c + 1| - |x_c - 1|) \tag{1}$$



Fig. 2. Connected component detection in four different directions.

drawn in Fig. 3(a); and (c) cell *external-input:* $u_c$. Processing itself is governed by a set of coupled nonlinear differential equations, one per cell. We use equations that differ from those originally proposed by Chua-Yang [1], and which enable the optimization of the speed/power ratio and area occupation of VLSI CNN chips. The proposed equations are given by [9], [10]:

$$\tau \frac{dx_c}{dt} = -g[x_c(t)] + D_c$$
$$+ \sum_{d \in N_r(c)} \{A_{cd}y_d(t) + B_{cd}u_d\}$$
$$\forall c \in \mathcal{GD} \tag{2}$$

where $g(\cdot)$ is a nonlinear dissipative term defined as,

$$g(x_c) = \begin{cases} m(x_c + 1) - 1 & x_c < -1 \\ x_c & \text{otherwise} \\ m(x_c - 1) + 1 & x_c > 1 \end{cases} \tag{3}$$

where $m > 1$ is a parameter of the model. Function $g(\cdot)$ is drawn in Fig. 3(b). Summations in (2) extend over the *neighborhood* of the $c$th cell, denoted by $N_r(c)$, which contains adjacent cells located within a distance $r$ in the grid, and includes cell $c$ itself.

Processing tasks performed by CNN's are determined by the *convergence* of (2) to binary $(y_c = \pm 1, \forall c \in \mathcal{GD})$ equilibrium states following the transient initialized by $[x_c(0)]$, driven by $[u_c]$, and under the boundary conditions imposed by cells at the net border. Depending on the application, the current $I_c$ generated at each cell's photosensor is used as initial value of the state variable $x_c(0)$ or as external input
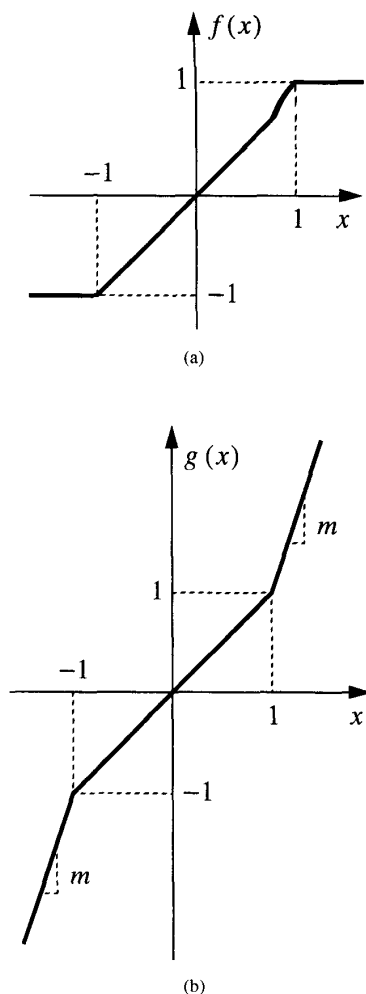
(a)



(b)

Fig. 3.   CNN cell nonlinearities. (a) Output nonlinearity; (b) dissipative term.

TABLE I
SOME CNN TEMPLATES

| Application | A | B | D |
|---|---|---|---|
| Noise Filtering [2] | $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ | $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ | $0$ |
| Hole Filling [28] | $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ | $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ | $-1$ |
| Convex Corners Extraction [2] | $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ | $\begin{bmatrix} -1/4 & -1/4 & -1/4 \\ -1/4 & 2 & -1/4 \\ -1/4 & -1/4 & -1/4 \end{bmatrix}$ | $-3$ |
| Borders Extraction [2] | $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ | $\begin{bmatrix} -1/4 & -1/4 & -1/4 \\ -1/4 & 2 & -1/4 \\ -1/4 & -1/4 & -1/4 \end{bmatrix}$ | $-2$ |
| Connected Component Detection [15] | $\begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$ | $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ | $0$ |
| Shadow Creation [29] | $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}$ | $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ | $0$ |

$u_c$. In the later case, the initial states are usually set to a constant value. The outcome of the task depends on parameters $B_{cd}$, $A_{cd}$, and $D_c$ of (2), called *control, feedback*, and *offset* parameters, respectively, and on the boundary conditions. The control and feedback parameters can be arranged into matrices, which provide a pictorial view of the interactions within each cell's neighborhood. For uniform networks these matrices are invariant throughout the grid domain—they are *templates*. The functionality of uniform CNN's is determined by its control, $B$, and feedback, $A$, template matrices, and its offset parameter, $D$. For illustration purposes, Table I summarizes the templates used for some significant preprocessing tasks.

To guarantee correct operation of smart-pixel CNN chips, an important mathematical issue is to determine conditions of the template parameters that yield convergence of the output matrix $[y_c]$ to binary states for any input. Such a mathematical analysis for the model proposed in this paper, given by (2) and (3), is out of this paper's scope and has been reported elsewhere [9] for any $m \geq 1$. Our circuits use the particular

case $m \rightarrow \infty$, in which the nonlinear dissipative term forces the state variable $x_c$ to remain within the interval $[-1, 1]$. Consequently, $x_c(t) = y_c(t)$, and the implementation of the nonlinear operator in (1) is not required.

## IV. SENSORY CIRCUITRY

### A. Photosensors

The simplest photosensitive devices for CMOS $n$-well technologies are reverse-biased *photodiodes*, formed either directly between $n_+$-diffusion and substrate [3] or between well and substrate [13]. Current level for both devices is an increasing function of the junction area. In particular, we have measured currents up to 20 nA for well-substrate photodiodes with well area of $100 \times 100\,\mu\text{m}^2$, in a 1.6-$\mu$m single-poly technology, under environmental laboratory lighting. This current level increases significantly using a vertical CMOS-BJT as photosensor. Fig. 4(a) shows a conceptual layout and cross-section for this device, whose current is approximately proportional to the area of the well/substrate junction, $A_w$ in the figure. Current generated by this device is $\beta + 1$ times larger than that of a photodiode with the same well area

$$I_T \sim (\beta + 1)I_W \propto A_W \tag{4}$$

where $I_T$ denotes the phototransistor current, $I_W$ is the corresponding current for the well-diode, and $\beta$ is the transistor current-gain; measured $\beta$ for this technology is $37.7 \pm 0.8$, basically independent of transistor geometry [22].

We have measured currents up to $430 \pm 30$ nA (under normal laboratory illumination) for phototransistors with passivated well area of $60 \times 60\,\mu\text{m}^2$. Consequently, and since current and well area are approximately linearly related, it extrapolates current levels of 20 nA for minimum area devices ($13.6 \times 13.6\,\mu\text{m}^2$)—needed for increased density smart-pixel chips. However, for some critical tasks [7] these levels provided by minimum photosensors may not be large enough to guarantee the matching level required by the signal processing circuitry,
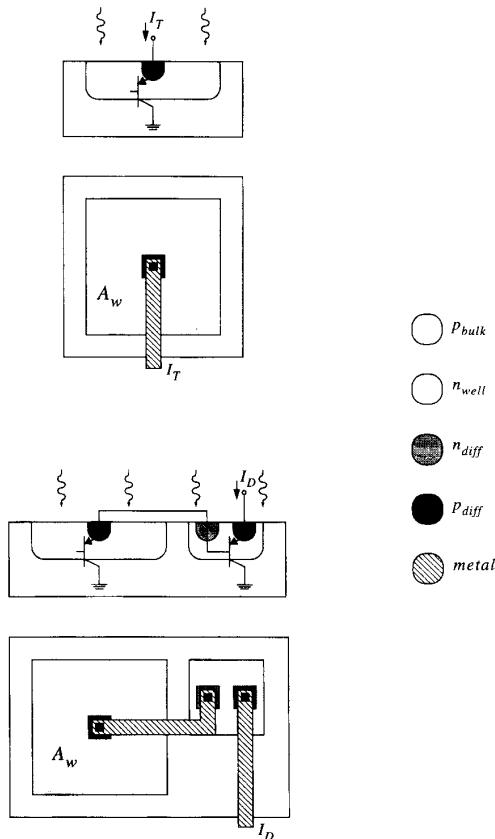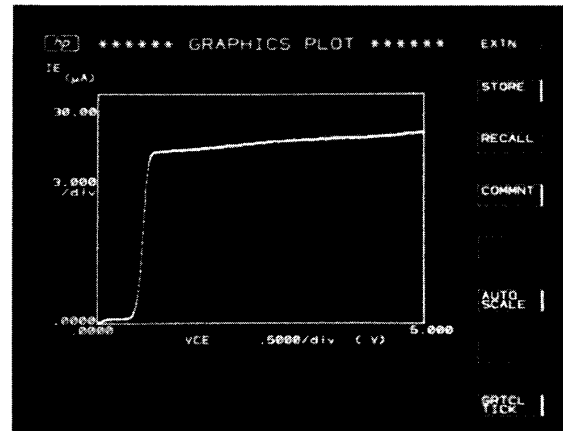
Fig. 4. (a) CMOS compatible vertical p-n-p transistor. (b) Darlington configuration of two vertical p-n-p transistors.



(a)



(b)

Fig. 5. Measured output characteristics of a Darlington phototransistor with $A_W = 60 \times 60 \ \mu m^2$: (a) under constant environment illumination; (b) effect of gradual reduction of illumination during the sweep of $V_{CE}$.

thus requiring some amplification. Simplest strategies use either larger wells or cascaded current amplifiers—very costly in terms of area occupation and, for the latter, inaccurate. Instead, we use an additional vertical BJT to achieve Darlington amplification by a factor of $\beta + 1$, with practically no area overhead. Fig. 4(b) shows the conceptual layout and cross-section for this Darlington phototransistor. Current for this device is
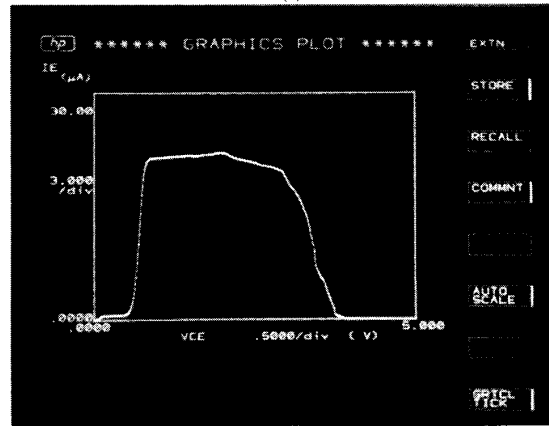
$$I_D \sim (\beta + 1)I_T \sim (\beta + 1)^2 I_W. \quad (5)$$

while its area occupation is scarcely increased by that of a minimum-size vertical BJT. Measurements with $A_w$ in Fig. 4(b) equal to $60 \times 60 \ \mu m^2$ result in currents up to $18 \pm 2 \ \mu A$.

Fig. 5(a) shows the *output* characteristic measured from a Darlington phototransistor with $A_w = 60 \times 60 \ \mu m^2$ under constant environmental illumination (bright-current), while Fig. 5(b) shows the result obtained when the environment light is gradually reduced to complete darkness. Dark-current was $215 \pm 10$ pA, which means that the bright-to-dark current-range is close to 100 dB for environmental laboratory illumination. The same range is observed for a single p-n-p device, while simple photodiodes yield about 80 dB. Although these results are optimistic in the sense that in real images there will be

no completely dark areas, the bright-to-dark current-ratios measured provide a wide enough range for data acquisition. The amplification of the Darlington structure provides a sufficient current level even if device area is substantially decreased.

*B. Autozero Strategy*

Although photosensors produce unidirectional current flow, double-rail signals are easily obtained by bias-shifting, as shown in Fig. 6(a). Current source $I_{TH}$ sets the zero-level of the double-rail signal. To guarantee good contrast, its value should be set somewhere between the maximum and the minimum light-induced currents among all photosensors. If lighting conditions for all possible input scenes are uniform and known a priori, $I_{TH}$ can be set to a fixed value. In a more general case where the chip must handle scenes with different lighting conditions, some kind of auto-zero strategy must be devised to generate $I_{TH}$ approximately equal to the average of the photosensor currents over the whole array.
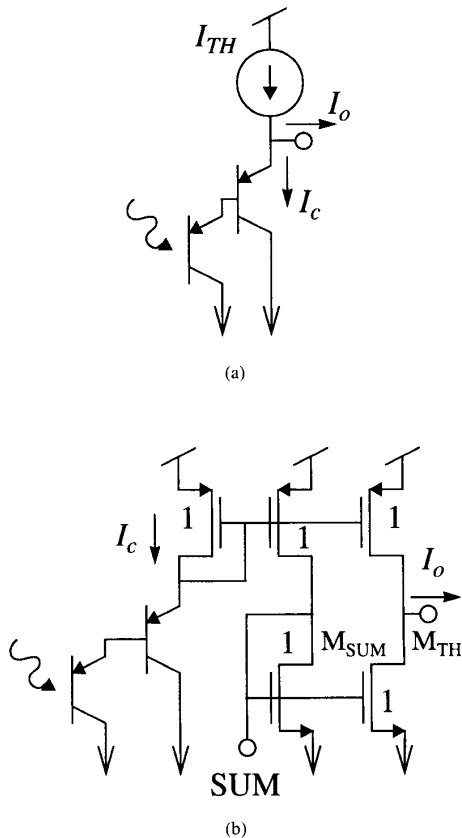
(a)



(b)

Fig. 6. Threshold circuitry for photosensors. (a) Bias-shifting for double-rail current output; (b) auto-zero circuitry.

A simple, yet convenient, auto-zero strategy uses four extra transistors at each sensor. Fig. 6(b) shows the schematic of a sensor including the auto-zero circuitry. All $p$-channel transistors have equal size; the same applies to $n$-channel transistors. The low-impedance node labelled SUM is a global node, common to all pixels. Note that the current $I_c$ at the $c$th photosensor is replicated twice. One of the replicas interfaces the processing circuitry, while the other is rooted to the global-node SUM, and aggregated to the remaining sensor currents. Thus, calculation of the current $I_{TH}$ through transistor $M_{TH}$ obtains the following

$$I_{TH} = \frac{N g_{mn}}{N g_{op} + N g_{mn}} \frac{1}{N} \sum_{c \in \mathcal{GD}} I_c \qquad (6)$$

where $g_{op}$ is the output conductance of the $p$-channel transistor, $g_{mn}$ is the transconductance of the $n$-channel transistor, and $N$ the number of pixels. For simplicity (6) assumes equal transconductances and conductances for all pixels. The first factor in (6) reflects the current division performed at node SUM, while the second corresponds to the gain of the mirror formed by the parallel combination of the $M_{SUM}$ transistors and $M_{TH}$. Assuming $g_{mn} \gg g_{op}$, (6) gives $I_{TH}$ equal to the average of the photosensor currents, and the light-threshold is automatically adjusted to the average illumination.
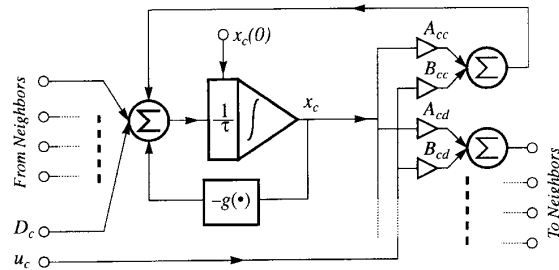


Fig. 7. Conceptual block diagram for the processing circuitry of a CNN smart pixel.

## V. PROCESSING CIRCUITRY

### A. Basic Circuit Building Blocks

Fig. 7 is a block diagram for the processing circuitry of the $c$th unit in a smart-pixel CNN chip, according to (2). This figure shows a core integrator with nonlinear losses and an output structure to generate weighted replicas of the $c$th input $u_c$ and state $x_c$, for transmission to the neighbor smart-pixels. The integrator is driven by weighted replicas of the input and state signals of the smart pixels in the neighborhood $N_r(c)$, plus an offset term, obtaining the following signal to drive the core integrator

$$J_c(t) = D_c + \sum_{d \in N_r(c)} \{A_{cd} x_d(t) + B_{cd} u_d\}. \qquad (7)$$

Current-mode provides a convenient choice to realize the processing circuitry of smart-pixel CNN's. On one hand, it enables direct interface with the sensors, whose outputs are currents. On the other, current summation at the integrator input node is directly achieved by routing signals to a common node. Finally, analog operators involved in Fig. 7 (weighted-replication, integration, and limitation) are realized by simple current mirror circuits.

Fig. 8(a) realizes the core integrator. Input current $J_c^*(t)$ is an unnormalized version of $J_c(t)$ in (7), with normalization factor $I_Q$: $J_c^*(t) = I_Q J_c(t)$. Output current $x_c^*(t)$ is the corresponding unnormalized version of $x_c(t)$. The parallel combination of the diode-connected input transistor $M_1$ and capacitor $C$ yields a time constant $\tau = C/g_m$, where $g_m$ is the transconductance parameter of $M_1$. On the other hand, note that current $x_c^*$ cannot swing beyond the values of the current sources which drive the common output node of transistors $M_2$ and $M_3$—meaning that $|x_c^*| < I_Q$. Thus, analysis of this circuit results in:

$$\tau \frac{dx_c^*}{dt} = J_c^* - I_Q g\left(\frac{x_c^*}{I_Q}\right) \qquad (8)$$

as required to realize (2), and where $g(\cdot)$ is the function defined in (3) with $m \to \infty$. In practice $\tau$ does not remain constant, but varies with input current level. However, most processing tasks tolerate this variation with no degradation of the network functionality [9].

Fig. 8(b) shows a circuit to realize the output structure of Fig. 7 from voltage $V_R$, and using the basic current mirror
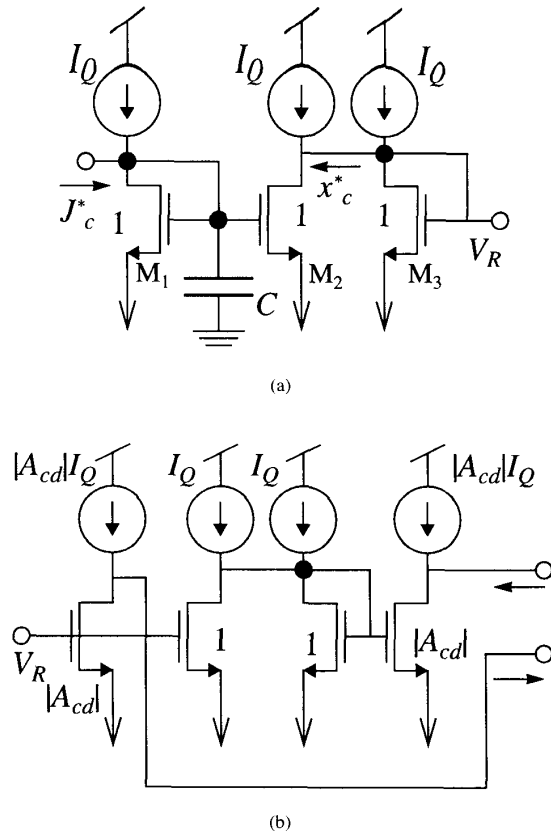
(a)



(b)

Fig. 8. Current-mode circuit blocks for the processing circuitry of CNN smart pixels. (a) Core integrator; (b) output structure.

principle of weighted replication [23]. Note that Fig. 8(b) contains two different substructures to cover each possible sign of the weight $A_{cd}$. Positive weights are obtained using a single output transistor whose geometry factor is $|A_{cd}|$ times that of transistor $M_1$. Thus, a current $A_{ce}x_c^*$ is sourced to the output node. Negative weights require an additional current mirror with unity weight for sign inversion.

### B. Some Circuit Design Issues

The following is a brief comment of dominant nonidealities encountered in the practical implementation of smart-pixel CNN chips and associated circuits.

*1) Current Gain Error:* A major source of error is the finite ratio of the input conductance $g_{in}$ to the output conductance $g_o$ of the current mirrors, which causes current gain error due to spurious current division. It is especially significant at the input node of the integrator, where the gain error $\epsilon$, is given approximately by [9]:

$$\epsilon \approx (N+1)\frac{g_o}{g_{in}} \qquad (9)$$

where $N$ denotes the number of mirrors driving this node—up to 18 for templates with no zero entries on a rectangular grid net with unity neighborhood parameter. For improved $g_o/g_{in}$ figures with short channel devices, cascode mirrors,

regulated mirrors, or a combination of both must be used [23]. In particular, analysis shows that the cascode mirror of Fig. 9(a) obtains values of $g_o/g_{in}$ several orders of magnitude lower than that for single mirrors, with smaller area occupation. Chips reported in Section VI are realized using these mirrors, and sized to handle the whole input current range with minimum distortion and smallest possible devices. For mirrors biased by a current $I_Q$, we obtain the following sizing equations

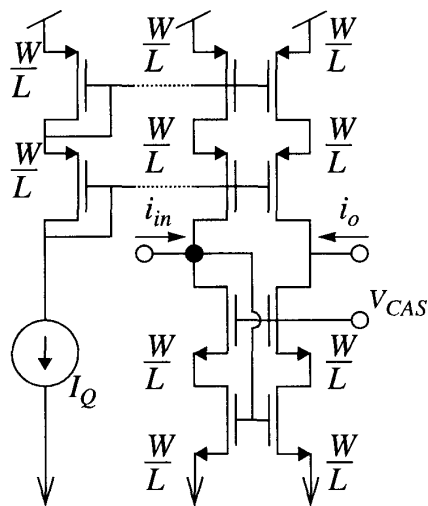$$W_n = \frac{8I_Q}{k_n V_{Tn}^2}L_n; \qquad V_{CAS} \approx V_{SS} + 2V_{Tn} \qquad (10)$$

where $k_n = \mu C_{ox}/2, V_{Tn}$ is the threshold voltage, and $V_{CAS}$ is the cascode voltage, which can be generated as shown in Fig. 9(b). We assume the same geometries $W_n$ and $L_n$ for all $n$-channel transistors in the cascode mirror. $W$ values for larger currents (associated to weighted replication) are calculated by imposing the constraint that all transistors have equal current density. Alternatively, for a given aspect ratio $W_n/L_n$, (10) establishes a bound for the maximum bias current of the device.

*2) Mismatch error; area, power and reliability:* Transistor geometry ratios, static gain error due to nonnull $g_o/g_{in}$, and power dissipation increase with $I_Q$. Hence, a bias current as small as possible should be chosen. The issue is to identify the minimum feasible rail current value. A lowest limit is certainly established by leakage, which in our case is increased by light effects. However, a more restrictive bound exists due to MOS transistor mismatch and Early voltage $(V_A)$ degradation with channel length.
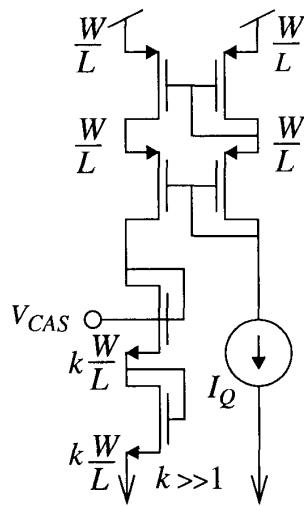
Mismatch is produced mainly by variations of $V_T$ and $\beta = \mu C_{ox} W/L$, whose standard deviations $\sigma(V_T)$ and $\sigma(\beta)/\beta$ for devices with equal layout show a component inversely proportional to the square root of the channel area, and another proportional to the distance between devices [24]. However, in the technology used and for transistor pairs closer than about 2.5 mm, the distance-dependent component is negligible for devices with channel area of less than $100\,\mu m^2$ [24]—larger than the values obtained using (10) for bias current below $\sim 50\,\mu A$ and channel lengths of $3.2\,\mu m$. Lower channel lengths have not been considered for several reasons, like short-channel effects, early-voltage degradation, and increased mismatch effects due to the associated low channel areas. In addition, lower transistor geometries do not result in appreciable area reductions due to the minimum contact size ($4\,\mu m$ with surrounding metal and diffusion in the technology used).

Another important consideration is that for a given $\sigma(V_T)$ and $\sigma(\beta)/\beta$, the ratio $\sigma(I)/I$ in MOS transistors operating in strong inversion and after pinch-off has an inverse dependency with $v_{gs} - V_T$. This means that once geometries have been set to achieve acceptable mismatch levels, bias current cannot be decreased too far below the bound given by (10), since this would produce a low $v_{gs}$ voltage at the bias point, with the corresponding large $\sigma(I)/I$. Hence, mismatch considerations establish bounds for both minimum area and power trends.

*3) Light effects on the processing circuitry:* Optical image acquisition forces the processing circuitry to be exposed to

(a)



(b)

Fig. 9. CMOS bias-shifted mirrors and biasing devices. (a) Cascode mirror and reference circuitry, (b) cascode voltage generation.



Fig. 10. Microphotograph of the DCC prototype.

## VI. EXPERIMENTAL RESULTS

### A. 16 × 16 *DCC Prototype*

The following measurements were taken from a 16 × 16 smart-pixel CNN chip intended for horizontal connected component detection [15] (see Fig. 2)—a basic preprocessing step for pattern recognition. Fig. 10 shows a microphotograph of the prototype, which in addition to the smart-pixel array contains boundary cells, output buffers, bias stages, and some digital control circuitry for the output image downloading process. The dimensions of the core array are $1890 \times 1530 \,\mu\text{m}^2$, and its power dissipation is 27 mW. The total chip dimensions, including the bonding pads, are $2480 \times 2500 \,\mu\text{m}^2$, with a total power dissipation of 42 mW and a total of 24 pins.

Fig. 11 shows the schematic and layout of one elementary unit. Unit dimensions are $118 \times 96 \,\mu\text{m}^2$, which include the sensor and associated regulation circuitry (~30% of the area), the processing circuitry, an additional current replication for output evaluation, and all required routing (cells are connected to each other by abutment). Sensor is realized with two minimum-size p-n-p devices in a Darlington configuration, to produce a bright-current under laboratory lighting of about $1 \,\mu\text{A}$, large enough for the matching requirements of this application. Cascode structures are used for both current sources and mirrors, and except for control switches, all MOS transistors have $W = 4 \,\mu\text{m}$ and $L = 3.2 \,\mu\text{m}$. Power dissipation with a 5 V supply and under environmental light in the laboratory is $105 \,\mu\text{W}$ per cell, unitary current being $I_Q = 2 \,\mu\text{A}$.

We have obtained 100% success (out of 30 trials) for full device level Montecarlo simulation of this chip. These Montecarlo simulations are based on the expected variations of the threshold voltages $V_{T0}$ and the large signal transconductance $\beta$ (body effect parameter $\gamma$ influences only the cascode transistors). Global biasing voltages are used for current reference generation, and bias stages are included in the simulation. Dispersion due to mismatch among transistors of different current sources did not produce critical results. Thus, global biasing is a fair approach for this application.

Fig. 12(a) illustrates the chip measurement setup and Fig. 12(b) shows five input images (left column) and the

light, which results in an increase of the leakage currents at the reverse-biased substrate-diffusion junctions. Unitary bias currents must be sufficiently large in order to neglect this effect. Also, MOS threshold voltage depends on light intensity, increasing the mismatch effect on current mirrors and sources. Current mirror transistors are commonly placed nearby, and hence light-intensity gradients have a reduced effect. On the contrary, current sources in different cells, biased by common global voltages, can exhibit larger dispersions. The tolerance of a particular application to variations in the unitary bias current must be evaluated in general, and local references should be used when required.
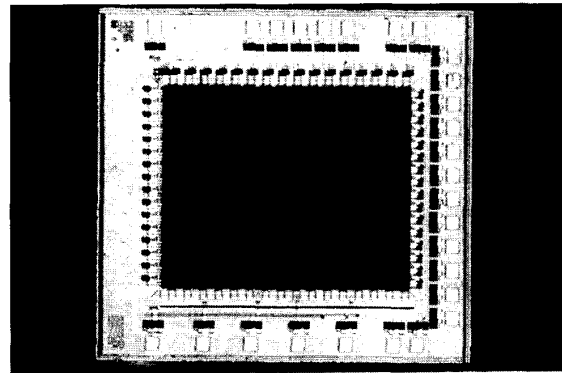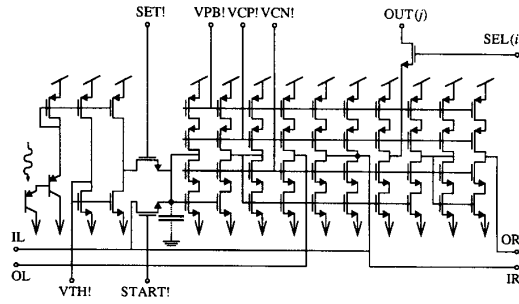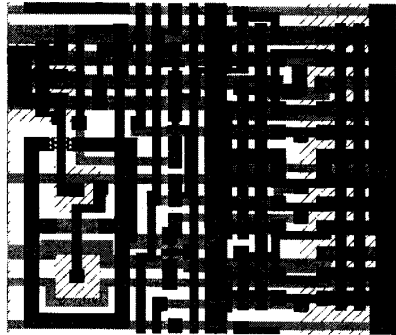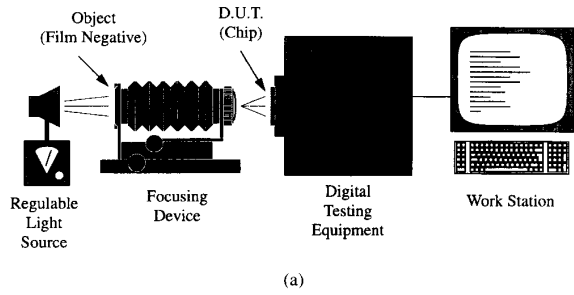
(a)



(b)

Fig. 11. (a) Schematic (refer to text for dimensions) and (b) layout of elementary unit of DCC prototype.

measured output images (right column). The prototype was exhaustively tested with 1200 input images. Fig. 12(c) contains the output waveforms observed from the cells in a particular row of the array during a processing example. The input pixels are displayed at the left side column, while the output ones are at the right. The signals display the measured transient evolution of the output of the cells in the row. Measured convergence time is $1.6\,\mu s$. Output image downloading requires $8\,\mu s$, using a 2 MHz digital clock frequency for the serial downloading process. Circuit operation remains correct, with no speed degradation, if the voltage supply is reduced from the nominal 5 V down to 2.7 V. This is another positive consequence of using current-mode techniques.
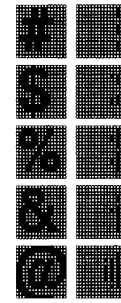
### B. 16 × 16 Radon Transform Prototype

This prototype performs the Radon Transform [25] of 16 × 16 pixels input images. This chip accepts electrical, as well as optical, input. The processing circuitry is based on a modified version of (2), where time has been discretized, and the nonlinearity is hard
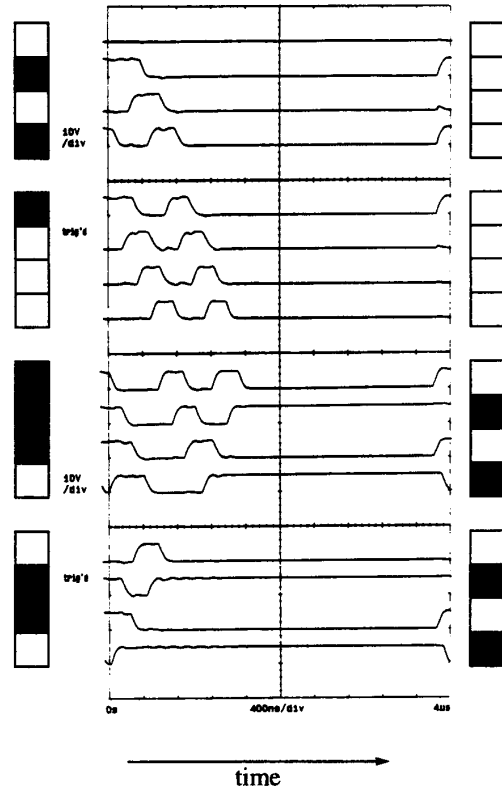
$$x_c(n+1) = \begin{cases} 1, & \text{for } D_c + \sum_{d \in N_r(c)} \{A_{cd}x_d(n) + B_{cd}u_d\} > 0. \\ -1, & \text{otherwise} \end{cases}$$

(11)



(a)



(b)



time

(c)

Fig. 12. (a) Measurement set up, (b) five input images and the output measured from DCC, and (c) measured transient response on a row of cells.
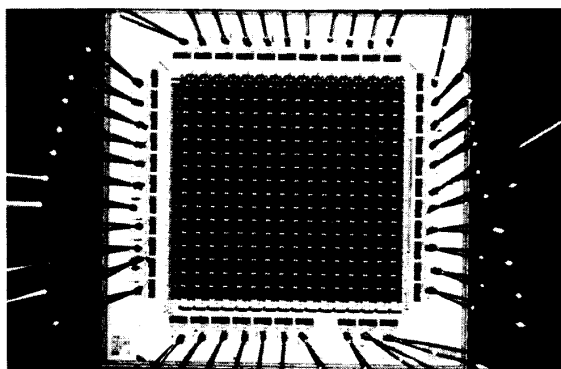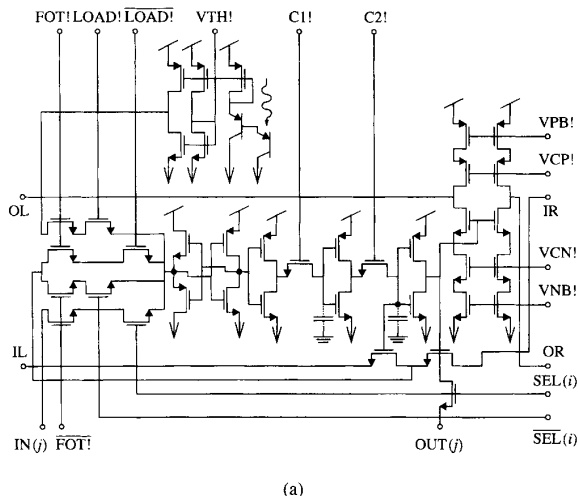
(a)



(b)

Fig. 13. (a) Schematic of elementary unit and (b) microphotograph of the Radon Transform prototype.

Also, this application requires signal-dependent weights [26]. In particular, the weights of the contributions going from a particular cell $c$ to its neighbors depend on $x_c$. The complete set of CNN coefficients can be described using unidimensional templates as follows

$$\mathbf{A} = \begin{cases} [0\ 0\ 1], & \text{if } (x_c \geq 0) \\ [1\ 0\ 0], & \text{if } (x_c < 0) \end{cases}$$
$$\mathbf{B} = [0\ 0\ 0] \qquad D = 0 \tag{12}$$

which reflect the scaling factors applied to the contributions of a particular cell to its neighbors.

Fig. 13(a) shows a simplified schematic of a cell, which uses pass transistors to realize the delay required in (11) and a high-resolution current comparator [27] for the hard nonlinearity. The design technique and the algorithm used in this circuitry is described in detail in [9]. Fig. 13(b) shows a microphotograph of the prototype. Cell dimensions are $121\,\mu\text{m} \times 124\,\mu\text{m}$, and the power dissipated by each cell is 1 mW—significantly larger than for the DCC due to the circuitry used to implement the hard nonlinearity.
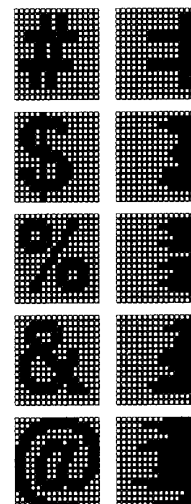


Fig. 14. Five input images and the corresponding output measured from the Radon Transform prototype.

The system contains a number of blocks located in the periphery of the cell array, like output buffers, bias stages, and digital control circuitry dedicated to the uploading and downloading processes. This additional circuitry, together with the bonding pads, result in a total system area of $2670\,\mu\text{m} \times 2680\,\mu\text{m}$, and a total system dissipation of 330 mW. The chip requires a total of 43 pins. This number is significantly higher than that of the previous prototype due to 16 input pads used for electrical input image uploading.

Using a 2 MHz digital clock frequency, image processing time is $8\,\mu\text{s}$. The serial downloading process also requires $8\,\mu\text{s}$. As an example, Fig. 14 shows five input images (left column) and the corresponding output images measured from the chip (right column). The complete test of the prototype involved 1200 images.

## VII. CONCLUSIONS

Summarizing, this paper has outlined a basic model and some design issues related to a methodology to design CNN smart-pixel chips in digital CMOS processes, and has presented measurements from two working prototypes in a 1.6-$\mu$m $n$-well CMOS technology. One calculates the number of connected pieces (DCC) of an input image in the horizontal direction, and the other evaluates the Radon Transform of an input image. The DCC chip obtains a density of $\sim$89 smart-pixels per $\text{mm}^2$ (each including sensory, regulation and processing circuitry), with a power consumption of $105\,\mu\text{W}$ per smart pixel and image processing times below $2\,\mu\text{s}$. Area and speed figures for the RT chip are similar. Although power dissipation is larger for this prototype, this can be corrected with a careful design of the current comparator [27].

As compared to previous CNN implementations, the proposed technique makes the required synergy between sensing and processing, and significantly improves area and speed/power figures. In particular, when compared to previous

chips for the same application [11], [12], the DCC chip, apart from including sensors at the cells, reduces the area consumption by a factor of 4, and improves the speed/power figure by more than one order of magnitude.

These area and power figures, and the fact that connections among pixels are made by abutment (requiring no extra routing area) enable forecasting single-die CMOS chips with $100 \times 100$ complexity and about 1 W power consumption.

These designs are mainly oriented towards preprocessing tasks which require fixed weights. We feel that there are potential application fields for these chips, provided efficient integration to massive processors is achieved. For this purpose, close cooperation between chip designers and system developers is necessary.

## ACKNOWLEDGMENT

The authors wish to thank Ricardo Carmona Galán for his work on the design of the Radon Transform prototype.

## REFERENCES

[1] L. O. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Trans. Circuits and Syst.*, vol. 35, pp. 1257–1272, Oct. 1988.
[2] L. O. Chua and L. Yang, "Cellular neural networks: Applications," *IEEE Trans. Circuits and Syst.*, vol. 35, pp. 1273–1290, Oct. 1988.
[3] A. Gruss, L. R. Carley, and T. Kanade, "Integrated sensor and range-finding analog signal processor," *IEEE J. Solid-State Circuits*, vol. 26, pp. 184–191, March 1991.
[4] E. A. Vittoz, "The design of high-performance analog circuits on digital CMOS chips," *IEEE J. Solid-State Circuits*, vol. 26, pp. 657–665, June 1985.
[5] H. Kobayashi, J. L. White, and A. A. Abidi, "An active resistor network for Gaussian filtering of images," *IEEE J. Solid-State Circuits*, vol. 26, pp. 738–748, May 1991.
[6] P. C. Yu, S. J. Decker, H. S. Lee, C. G. Sodini, and J. L. Wyatt, "CMOS resistive fuses for image smoothing and segmentation," *IEEE J. Solid-State Circuits*, vol. 27, pp. 545–553, April 1992.
[7] T. Roska and J. Nossek, Eds., "Special issue on cellular neural networks," *IEEE Trans. Circuits and Syst.-I and II*, vol. 40, March 1993.
[8] B. E. Shi and L. O. Chua, "Resistive grid image filtering: Input/output analysis via the CNN framework," *IEEE Trans. Circuits and Syst. I: Fundamental Theory and Applicat.*, vol. 39, pp. 531–548, July 1992.
[9] S. Espejo, "VLSI Design and Modeling of CNN's," Ph.D. dissertation, University of Sevilla, Spain, April 1994.
[10] A. Rodríguez-Vázquez, S. Espejo, R. Domínguez-Castro, J. L. Huertas, and E. Sánchez-Sinencio, "Current-mode techniques for the implementation of continuous-time and discrete-time cellular neural networks," *IEEE Trans. Circuits and Syst. II: Analog and Digital Signal Processing*, vol. 40, pp. 132–146, March 1993.
[11] J. M. Cruz and L. O. Chua, "A CNN chip for connected component detection," *IEEE Trans. Circuits and Syst.*, vol. 38, pp. 812–817, July 1991.
[12] H. Harrer, J. A. Nossek, and R. Steltz, "An analog implementation of discrete-time cellular neural networks," *IEEE Trans. Neural Networks*, vol. 3, pp. 466–476, May 1992.
[13] A. H. Sayles and J. P. Uyemura, "An optoelectronic CMOS memory circuit for parallel detection and storage of optical data," *IEEE J. Solid-State Circuits*, vol. 26, pp. 1110–1115, Aug. 1991.
[14] C. Jansson, P. Ingelhag, C. Svenson, and R. Forchheimer, "An addressable $256 \times 256$ photodiode image sensor array with 8-bit digital output," in *Proc. of ESSCIRC'92*, Sept. 1992, pp. 151–154.
[15] T. Matsumoto, L. O. Chua and H. Suzuki, "CNN cloning template: Connected component detector," *IEEE Trans. Circuits and Syst.*, vol. 37, pp. 633–635, May 1990.
[16] H. Suzuki, T. Matsumoto, and L. O. Chua, "A CNN handwritten character recognition," *Int. J. Circuit Theory and Applicat.*, vol. 20, pp. 601–612, New York: Wiley, Sept.–Oct. 1992.
[17] J. C. Bezdek and S. K. Pal, Eds., *Fuzzy Models For Pattern Recognition*, New York: IEEE Press, 1992.

[18] K. R. Crounse, T. Roska, and L. O. Chua, "Image halftoning with cellular neural networks," *IEEE Trans. Circuits and Syst. II: Analog and Digital Signal Processing*, vol. 40, pp. 267–283, April 1993.
[19] T. Roska, T. Boros, P. Thiran, and L. O. Chua, "Detecting simple motion using cellular neural networks," in *Proc. First IEEE Int. Workshop on Cellular Neural Networks and Their Applicat.*, Budapest, Dec. 1990, pp. 127–138.
[20] C. P. Chong, C. A. T. Salama, and K. C. Smith, "Image motion detection using analog VLSI," *IEEE J. Solid-State Circuits*, vol. 27, pp. 93–96, Jan. 1992.
[21] W. Bair, C. Koch, A. Moore, T. Horiuchi, B. Bishofberger, and J. Lazzaro, "Computing motion using analog VLSI vision chips: An experimental comparison among four approaches," in *Proc. Second Int. Conf. on Microelectron. for Neural Networks*, Munich, Germany, Oct. 1991, pp. 291, 309.
[22] B. Pérez-Verdu, F. V. Fernit texinputndez, A. Rodríguez-Vázquez, and J. L. Huertas, "Modeling and characterization of lateral BJT's in CMOS technologies," in *Proc. of the VI Spanish Congress on Int. Circuits*, 1991, pp. 75–80.
[23] C. Toumazou, F. J. Lidgey, and D. G. Haigh, Eds., *Analog IC Design: The Current-Mode Approach*, London: Peter Peregrinus, 1990.
[24] 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.
[25] C. W. Wu, L. O. Chua, and T. Roska, "A two-layer radon transform cellular neural network," *IEEE Trans. Circuits and Syst.-II*, vol. 39, pp. 488–489, July 1992.
[26] T. Roska and L. O. Chua, "Cellular neural networks with nonlinear and delay-type template elements and nonuniform grids," *Int. J. Circuit Theory and Applicat.*, vol. 20, pp. 469–481, New York: Wiley, Sept.–Oct. 1992.
[27] R. Domínguez-Castro, A. Rodríguez-Vázquez, and J. L. Huertas, "High resolution CMOS current comparators," in *Proc. 1992 European Solid-State Circutis Conf.*, Copenhagen, Denmark, Sept. 1992, pp. 242–245.
[28] T. Matsumoto, L. O. Chua, and R. Furukawa, "CNN cloning template: Hole filler," *IEEE Trans. Circuits and Syst.*, vol. 37, pp. 635–638, May 1990.
[29] T. Matsumoto, L. O. Chua, and H. Suzuki, "CNN cloning template: Shadow detector," *IEEE Trans. Circuits and Syst.*, vol. 37, pp. 1070–1073, Aug. 1990.

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

From 1989 to 1991 he was an intern at AT&T Bell Laboratories at Murray Hill, NJ, and an employee of AT&T Microelectronics of Spain. He is currently a teaching assistant at the Department of Electronics and Electromagnetism of the University of Seville, and 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, chaotic circuits and communication systems.

**Angel Rodríguez-Vázquez** (M'80) received the Licenciado en Física degree in 1977, and the Doctor en Ciencias Físicas degree in 1983, both from the University of Seville, Spain.

Since 1978 he has been with the Department of Electronics and Electromagnetism at the University of Seville where he is an associate professor. He is also at the Department of Analog Circuit Design of the Centro Nacional de Microelectrónica. His research interests are in the fields of analog/digital integrated circuit design, analog integrated neural and nonlinear networks, and modeling of analog integrated circuits.

**Rafael Domínguez-Castro** received the five-year degree in electronic physics (Licenciado en Física Electrónica) in 1987, the M.S. equivalent in microelectronics in 1989, and the Doctor en Ciencias Físicas Degree in 1993, from the University of Seville, Spain.

Since 1987 he has been with the Department of Electronics and Electromagnetism at the University of Seville, where he is currently a teaching assistant. He is also with the Department of Analog Circuit Design of the Spanish Microelectronics Center (Centro Nacional de Microelectrónica). His research interests are in analog/digital integrated circuit design, including neural and fuzzy circuits, and computer-aided design and modeling of analog integrated circuits.

**José L. Huertas** received the Licenciado en Física degree in 1969 and the Doctor en Ciencias Físicas degree in 1973, both from the University of Seville, Spain.

From 1970 to 1971 he was with the Philips International Institute, Eindhoven, the Netherlands, as a postgraduate student. Since 1971 he has been with the Department of Analog Circuit Design of the Centro Nacional de Microelectrónica. His research interests are in the fields of multivalued logic, sequential machines, analog circuit design, and nonlinear network analysis and synthesis.

**E. Sánchez-Sinencio** photograph and biography not available at time of publication.