A New Current Mode Programmable
Cellular Neural Network

L. Ravezzi¹, G.F. Dalla Betta², G. Setti³

¹DIM-Università di Trento, Via Mesiano 77, 38050 Trento, Italy
²IRST-Divisione Microsistemi, Loc. Pante di Povo, 38050 Trento, Italy
³DEIS, Università di Bologna, Viale Risorgimento 2, 40136 Bologna, Italy

ABSTRACT We report on the design of a new full-analog current-mode CNN in a 1.2 µm CMOS technology, whose cell core is characterized by an intrinsic capability of weights control, low power consumption and small area occupation. Circuit simulations allowed the design approach to be validated and the electrical performance of the CNN to be predicted; moreover, it is shown that the proposed CNN can be successfully adopted for several applications in image processing. A preliminary CNN test-chip, consisting of a 8×1 array for CCD and shadow detection, is currently being fabricated at IRST (Trento, Italy) in a 2.5 µm CMOS technology.

1 Introduction

Recently, many theoretical studies in the Cellular Neural Network (CNN) field tend to focus on networks made of cells having more and more complex dynamics and coupled by increasingly complex interactions. Despite these useful generalizations, the interest about the original, simple Chua and Yang [1] model has not decreased, both from the theoretical and the practical point of view. In fact, on one hand, the relatively simple cell structure and the local interactions allow to deeply understand the way a CNN works and to highlight some of its potentials and limitations to process information (typically images) [2][3]. On the other hand, the use of first neighboring cells connections only, the architecture simplicity of both processing units and interconnections elements, and the space invariant property make CNN's particularly well suited for an analog VLSI implementation.

Nevertheless, to achieve an acceptable resolution with standard design procedures, a large cell grid size is required, which results in considerable area occupation, low production yield and high power consumption [4]. From this point of view, current-mode VLSI implementations are particularly attractive [5][6][7]. In fact, these circuits often employ minimum size transistors, thus allowing for a high modularity in the cell structure, simplifying the hardware design. As further advantage, since image acquiring devices (e.g. charge coupled devices) are usually characterized by signal output current, they can be directly connected to current-mode CNN's without needing current-voltage converters.

In this work we present the design of a new CMOS full-analog current-mode CNN which can be adopted for several applications in image processing. In particular, the CMOS circuit proposed for the cell core is characterized by intrinsic weights programmability, low power consumption and small area occupation. The paper is organized as follows. Section 2 describes the basic building blocks of the cell core, Section 3 reports on the design and simulations of an entire cell, while in Section 4 the simulation results for a complete CNN performing different applications are presented and discussed. Moreover we also report the layout of 8×1 test CNN, which is currently being fabricated at IRST facility, as well as the results of several performed Montecarlo simulations.

2 Design of the cell core

For the proposed current-mode CNN we refer to the so-called Full Signal Range (FSR) model [6][7], which is described by

\[ \frac{dx^e}{dt} = -g(x(t)) + \sum \limits_{d \in \mathbb{N}_+(c)} A^e_d \psi_d(t) + B^e d^e + D^e, \] (1)
where $x^c$, $u^c$ and $y^c = 1/2(|x^c| + 1 - |x^c| - 1)$ are the cell state, input and output variable and where

$$g(x^c(t)) = \begin{cases} 
  m(x^c - 1) + 1 & x^c > 1 \\
  x^c & |x^c| \leq 1 \\
  m(x^c + 1) - 1 & x^c < -1. 
\end{cases} \quad (2)$$

In the following, we will only consider the limiting case $m \to \infty$. Therefore, the cell state and output variable are identical and the nonlinear output function can be discarded, so that FSR based CNN's are characterized by an overall hardware simplification with respect to networks based on the original Chua-Yang model [7]. The block diagram for a FSR-CNN cell can be represented as shown in Fig. 1, where, with standard notation, $A_{ij}^c$, $B_{ij}^c$ and $D^c$ indicate the $A$- and $B$-template coefficients and the offset term, respectively, and where $N_r(c)$ is the cell neighborhood.

![Figure 1: Block diagram for a FSR model based CNN cell](image)

If a circuit implementation of this structure is the goal, particular attention must be given to the integrator block, which also includes the necessary nonlinear transformation in order to simplify the cell structure. The new proposed circuit for implementing this block has been obtained by suitably adding the necessary saturation effect to a bidirectional current mirror based integrator, as shown in Fig. 2.

![Figure 2: New proposed integrator with saturation](image)

To understand how the desired output current saturation effect can be achieved, refer to the circuits of Fig. 3(a) and indicate as resistive stage and inverting stage the circuits obtained by connecting this block with the one of Fig. 3(b) and Fig. 3(c), respectively (namely the two stages of a bidirectional current mirror). Consider now for instance a resistive stage: when the current drained by $M_{p3}$ (or $M_{n3}$) is lower than the current supplied by $I_{b1}$ ($I_{b2}$), $M_{p2}$ ($M_{n2}$) is on and, thanks to its diode connection, it keeps node A (node B) at a nearly constant voltage. The current drained by $M_{p3}$ ($M_{n3}$), however, cannot exceed the current provided by $I_{b1}$ ($I_{b2}$), thus realizing the nonlinear function. In this way, the desired integrator with saturation can be realized by connecting a resistive stage and an inverting stage and by adding a capacitor $C_e$, as shown in Fig. 2, where the ideal current sources $I_{b1}$ and $I_{b2}$ of Fig. 3(a) have been realized by simple current mirrors. It is worthwhile stressing that correct circuit operation can be achieved by using minimum size transistors only.

By varying the current forced by $M_{p4}$ ($M_{n4}$) in Fig. 2, a corresponding variation in the gain of the bidirectional current mirror $M_{p3} - M_{p6}$, $M_{n3} - M_{n6}$ can be obtained. Hence, the multipliers shown in the block scheme of Fig. 1 can be realized with simple inverting stages, thus easily achieving the template coefficient programmability. Therefore, since the drain current of $M_{p4} - M_{n4}$ determines the corresponding template element value, in the following we will simply refer to it as current weight). Note also that, in Fig. 2, the weight control-voltages on the gate of transistors $M_{p1}$, $M_{n1}$ are generated by transistors
Figure 3: Schematic diagram of the building blocks of a bidirectional mirror with saturation

$M_{p7}-M_{n7}$ (inside the dashed box). Depending on power consumption or area occupation requirements, these transistors can either be isolated from the cells, or placed locally in every cell. The former solution requires to distribute the control voltages $V_{p1}$ and $V_{p2}$ all over the network, thus increasing the global wiring complexity, while the latter needs a lower number of interconnections but adds a current path between power supply and ground in each cell, thus increasing the power consumption. However, it is worthwhile noting that the aspect ratio of $M_{p7}$ ($M_{n7}$) can be lower than that of $M_{p4}$ ($M_{n4}$), thus reducing the power consumption of the cell.

The current gain of the integrator of Fig. 2 in the middle point of its transfer characteristic can be calculated by means of a small signals analysis and is given by:

$$A_i = \frac{i_{out}}{i_{in}} = \frac{g_{mn6} + g_{mp6}}{r_{n3} + r_{p3}}$$

(3)

In order to obtain a piecewise-linear characteristic for the integrator, a linear dependence of $A_i$ on the selected weight would be required. This cannot be achieved because of the square-function dependence of the transconductance of $M_{n6}$ and $M_{p6}$, operating in saturation, on the current weights (namely the drain currents of $M_{n4}$ and $M_{p4}$). Fig. 4(a) shows the ideal and real characteristics of $A_i$ as a function of the current weight $I_Q$: as can be seen, there is only one matching value ($I_Q$) for the two curves. The real sigmoidal characteristics corresponding to the two cases $I_Q < I'_Q$, and $I_Q > I'_Q$ are shown in Fig. 4(b) and Fig. 4(c), respectively. As the difference between the ideal and real characteristic is lower for $I_Q < I'_Q$ (see Fig. 4(a)), the minimum error will be reached by choosing $I'_Q$ equal to the highest weight allowed. Worthwhile to stress that, despite of the differences between the obtained sigmoidal input-output characteristic and the piecewise linear one, in all the considered applications the network qualitative behaviour do not appreciably differ from the ideal one, similarly to what happens for CNN’s described by the original model [3].

The time constant corresponding to the central point of the sigmoidal characteristic, disregarding the parasitic capacitances, is:

$$\tau = \frac{1}{\omega_r} = \frac{r_{n3}r_{p3}C_x}{r_{p3} + r_{n3}} = \frac{2C_x}{g_{mp3} + g_{mm3}}$$

(4)
Thus, if $C_2$ is kept constant, an increase in the speed of the network can be achieved by means of either an increase of the aspect ratio of $M_{23}$ and $M_{33}$ (see Fig. 2) or an increase of the current forced through the resistive stage of the integrator. Hence, a reasonable trade-off among speed response, area occupation and power consumption of the network has to be found.

3 Cell implementation

The basic building blocks described in the previous section have been developed with reference to a well established 1.2um CMOS technology and considering a power supply voltage of ±2.5V. The cell has been designed to allow for the complete programmability of the $A$- and $B$-template coefficients and the offset term. In order for most practical applications to be performed, a programmable weight values range from ±0.125 to ±8 with analog selection has been considered. The simulated sigmoidal characteristics for different negative weights are reported in Fig. 5.

![Figure 5: Sigmoidal characteristics for negative weights (a) and a particular for small value weights (b)](image)

The absolute value of a weight can be changed by using an inverting stage, while, to obtain opposite sign weights, a switch which allows to choose between the output variable and the complementary one is necessary (see Fig. 6). Therefore, for each weight, two control signals are needed, one for the absolute value and one for the sign. In summary, the cell core contains the integrator and the current inverter (with unitary gain) realized with a further inverting stage and a resistive stage. The schematic of the cell with the core block, the block implementing the self-feedback coefficients and those needed to supply weighted currents to the neighboring cells is reported in Fig. 6.

![Figure 6: Schematic of a cell with some block realizing the feedback template coefficients](image)

The CMOS pass-transistor shown in Fig. 6 is required for the initialization of the cell state variable. In particular the feedback path is open during the initialization phase and the initial current is forced in the state node of the integrator resistive stage. On the contrary, the processing phase is started by closing the feedback path. The initial current is stored within a built-in memory present in each cell (not shown), which consists of a 4-bits dynamic memory with D/A converter. In particular, the stored initial current can vary between $-1 \mu A$ to $+1 \mu A$ with a step of about 150nA.

Finally, the control template coefficients are implemented with a circuit similar to that of Fig. 6, which, however, does not contain the self-feedback and the CMOS pass-transistor (in this case, in fact, the core block input current, stored in the built-in memory, corresponds to the input variable).
Fig. 7(a) shows the switching transients of a cell having a state capacitor of $1\mu F$ with $\pm 2\mu A$ steps. The fall time and rise time are about $1\mu sec$. The power consumption of core block and the self-feedback stage with a coefficient equal to 2 is about $60\mu W$.

![Figure 7: Switching transient of a cell with positive and negative step](image)

Figure 7: Switching transient of a cell with positive and negative step

Other simulations proved the functionality of the same cell with a $\pm 1.65 V$ supply voltage; in this case the power consumption of the core block and the self-feedback stage (with weight equal to 2) decreases to $37\mu W$ (a reduction of 38%), whereas the switching time with the same steps and capacitor as before is about $1.2\mu sec$ (increase of 20%).

4 Applications

A $8 \times 8$ fully-programmable CNN was designed and simulated at device-level in order to validate the proposed cell design for different applications. The $A$- and $B$-templates, offset terms $Dc$, initial conditions $x_c(0)$ and input values $u_c$ used for the different applications are the same as in [8]. Circuit simulations proved the functionality of the network for noise removal, hole filling (Fig. 8(a)), and edge-detection applications (Fig. 8(b)). The correct operation of a $8 \times 1$ array for connected component detection (CCD) was also evaluated (see Fig. 8(c)). Table 1 reports the transient times and power consumptions of the whole network for the considered applications. The parasitic capacitances extracted from the layout were also taken into account for these simulations.

![Figure 8: (a) Hole filling, (b) edge detection, and (c) CCD operations](image)

Table 1: Simulated performance of the designed CNN for different applications

<table>
<thead>
<tr>
<th></th>
<th>Noise Removal</th>
<th>Hole Filling</th>
<th>Edge Detection</th>
<th>CCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power consumption (mW)</td>
<td>5</td>
<td>11</td>
<td>11</td>
<td>0.56</td>
</tr>
<tr>
<td>Transient times $\mu sec$</td>
<td>2</td>
<td>3.9</td>
<td>2.2</td>
<td>4</td>
</tr>
</tbody>
</table>

In order to further validate the proposed design approach with low development costs and time delay, a preliminary CNN test-chip, consisting of a $8 \times 1$ array for CCD and shadow detection, has been designed and is currently being fabricated at IRST (Trento, Italy) with a 2.5 $\mu m$ CMOS technology. The layout of this latter structure containing 9 cells ($8 \times 1$ array + 1 single cell) as well as the I/O and control circuitry, for an overall area occupation of $2.3mm \times 0.7mm$, is reported in Fig. 9.

The robustness of the network against process-related parameter fluctuations was carefully tested by means of Montecarlo simulations at device-level. In particular, we found that the designed CNN is insensitive to a uniform deviation up to 15% with respect to nominal values of oxide thickness and threshold voltage. Noticeably, the CNN operates correctly even supposing a randomly-distributed variation up to 3% of the same parameters. As an example, Fig. 10(a) shows the final value of the state current for the cells composing the previously mentioned CNN for CCD operation resulting from a
Monte Carlo simulation. As can be seen, the current values are only slightly different from the ideal ones (1 \mu A and -1 \mu A, corresponding to "black" and "white", respectively), hence the network still works properly (compare Fig. 10(a) to Fig. 8(c)). The same technological parameters fluctuation also affects the CNN dynamic performance; in fact, as shown in Fig. 10(b), an increase up to 50% can occur for the transient time of cell 3, that, due to the domino propagation mechanism between cells [8], corresponds to the worst case for the particular choice of the initial condition shown in Fig. 8(c).

Acknowledgements The authors are grateful to Prof. G. Masetti and Prof. S. Graffi of the University of Bologna and Prof. G. Soncini of the University of Trento for many useful suggestions and discussions.

References


