ACCURATE DESIGN OF ANALOG CNN IN CMOS DIGITAL TECHNOLOGIES

A. Rodríguez-Vázquez, R. Domínguez-Castro and J.L. Huertas

Centro Nacional de Microelectrónica
Edificio CICA, Avda. Reina Mercedes s/n, 41012-Sevilla, Spain

ABSTRACT

We explore the design of CNN by using sampled-data analog current-mode techniques which does neither require capacitors nor resistors but just MOS transistors. This feature made the proposed technique well suited for implementation in conventional VLSI MOS technologies. A set of building blocks is presented and their performance validated by device-level simulation results. Also, guidelines are given concerning the choice of the circuit parameters for optimum operation.

1. Introduction

In spite of the prevalent dominance of digital techniques, analog circuits have continued to be recognized as the ideal solution for those applications where either *real-time* processing or *low-power* consumption is required. The current status of analog VLSI technology makes it possible to overcome many of the problems in the implementation of analog circuits. In particular, there exists a strong renovated interest in the design of new analog computational models based on some aspects of biological neural nets [Mead89]. It has been encouraged for recent proposals demonstrating the application of these *artificial neural networks* in fields as optimization, associative memory, pattern recognition, etc. [Tank86, Lipp87 Mead89a, Mead89b].

Most of the up to now proposed analog neural network implementations have been focused on fully interconnected architectures [Mead89b, Morg90], specifically in the Hopfield's neural net model [Tank86]. Recently a new circuit architecture, called *Cellular Neural Network*, has been proposed [Chua88a] whose basic computational units, called cells, are only connected to their corresponding nearest neighbors in the net. This *local connectivity* property is very convenient to simplify routing in VLSI implementations. Also, Cellular Neural Networks (CNN) show

potential applications in areas as image processing and pattern recognition [Chua88b], which combined to the local connectivity feature motivate research in the implementation of this kind of architectures.

A typical *cell* contains linear capacitors, linear resistors, linear and nonlinear controlled sources and independent sources [Chua88a]. In [Yang90] a VLSI circuit has been proposed where the nonlinear controlled sources are implemented by exploiting the large signal differential current to voltage characteristics of a MOS differential pair. Replication and selective weighting of the differential current provided by the pair is achieved by means of batteries of current mirrors.

There are several drawbacks in this implementation approach. On the one hand, non-conventional specialized technologies are required as a consequence of the use of both resistors and capacitors. This also yields large area occupancy since large time constants have to be used to reduce the influence of parasitics [Yang90]. On the other, *programmability* and *reconfigurability*, which are fundamental features for neural systems cannot be easily included. Finally, accomodating the dynamic ranges of the different voltages and currents in the circuits to ensure operation according to the model may not be an easy task. For instance, it may be difficult to conjugate the requirement for large variation ranges in the voltage representing the state of each cell [Chua88a] with the use of cascode current mirrors. Cascode mirrors seem on the other hand to be unavoidable to reduce the influence of parasitic output impedance on the cell state node.

In this presentation we discuss a VLSI implementation of analog CNN where only MOS transistors are required. This is very appealing for conventional VLSI MOS technologies where available resistors and capacitors are bulky and inaccurate. Besides, all processing in the proposed approach is made on currents which is very appealing from the point of view of ensuring proper dynamic ranges when using scaled down technologies.

2. VLSI Analog CNN Architecture

Let us assume an *MxN* CNN and consider the cell on the *ith* row and the *jth* column, *C(i, j)*. A block diagram for the analog VLSI implementation of this cell is shown in Fig.1a, where the variables *u*, *x* and *y* (specialized by suffices indicating the cells rows and columns) have been used to denote the *inputs*, *states* and *outputs* of the cells, respectively. The following state equation can be written down from this block diagram:

$$\tau \frac{dx_{ij}}{dt} = -A_o x_{ij}(t) + \left| \sum_{\substack{c(k,l) \\ \in \\ Nr(i,j)}} A(i,j;k,l)\, y_{kl}(t) + \sum_{\substack{c(k,l) \\ \in \\ Nr(i,j)}} B(i,j;k,l)\, \mu_{kl} \right| + D_{ij} \qquad \text{(1a)}$$

where N_r (i, j) represents the *r-neighborhood* of the cell $C(i, j)$ according to the definition in [Chua88a], A_o, $A(i, j; k, l)$, $B(i, j; k, l)$ and D_{ij} are real parameters and $y_{kl}(t)$ can be obtained from the corresponding state variable by the following nonlinear transformation:

$$y_{kl} = \tfrac{1}{2}(|x_{kl}+1| - |x_{kl}-1|) \tag{1b}$$

which is depicted in Fig.1b.

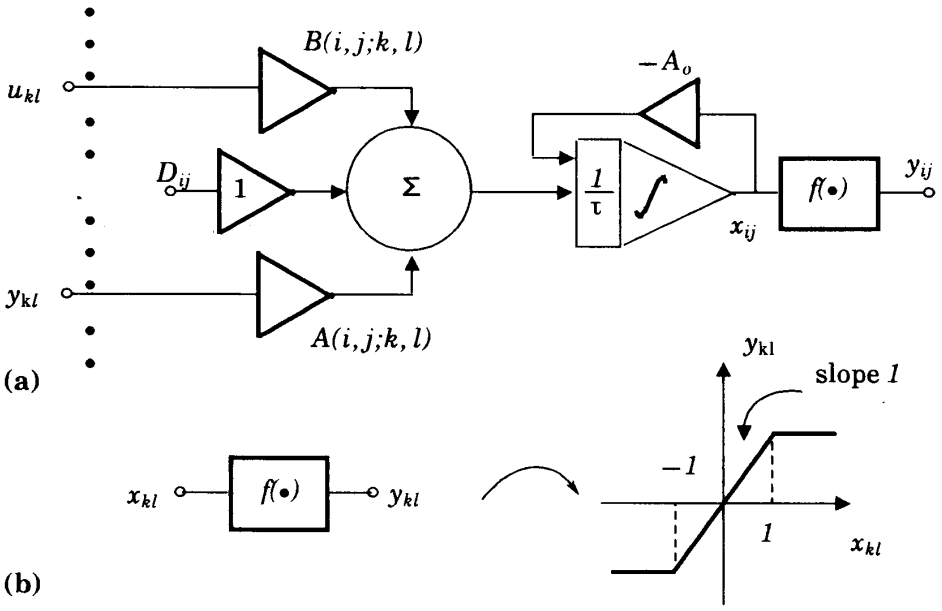

Figure 1: Block Diagram for cell C(i, j) of an analog CNN

From Fig,1 we can see that each cell has inputs coming both from the driving signals, u_{kl}, and from the output signals of the cells included in its neighborhood. Also, according to Fig.1, the only operators required for the implementation of CNN are *weighted-summers*, *lossy integrators* and *saturation nonlinearities*. There are two basic approaches for the implementation of these operators in analog VLSI: 1) using OTA-C techniques. 2) using switched-capacitors. Both approaches allow exploitation of KCL for the implementations of the summer operation, which is very convenient to simplify the circuits. Resorting to switched-capacitors offers the advantage of a better controlled dynamics, the price to be paid for being a reduction in the operation speed. On the other hand, and although OTA-C techniques are intrinsically faster, their dynamics may be severely degraded by the influence of unavoidable parasitics, thereby imposing the need of making a difficult tradeoff between stability and speed of operation. Both approaches share

however the drawback of requiring capacitors (in some cases also resistors) for the lossy integrators and hence are not well suited for conventional VLSI technologies. Here we propose to use an alternative design technique exhibiting the well-controlled dynamics feature of switched-capacitors but requiring only MOS transistors for its practical implementation. It is an analog sampled-data current-mode approach based on the following discrete-time version of (1a):

$$x_{ij}(n+1) = x_{ij}(n) + \frac{1}{\tau}\left\{-A_o x_{ij}(n) + \left[\sum_{\substack{c(k,l)\\ \in \\ Nr(i,j)}} A(i,j;k,l)\, y_{kl}(n) + \sum_{\substack{c(k,l)\\ \in \\ Nr(i,j)}} B(i,j;k,l)\, u_{kl}\right] + D_{ij}\right\} \qquad (2)$$

where n is the discrete-time variable ($t = nT_c$, T_c being the clock period and $n = 0,1,2,3...$) and $y_{kl}(n)$ is given as a function of $x_{kl}(n)$ according to (1b).

Dynamics of the model in (1a) has been analyzed in [Chua88a] where several theorems are given providing the foundation for the design of CNNs. Equation (2) exhibits similar dynamic properties as (1a). In order to illustrate these properties and in particular to take a glimpse at the effects on the choice of τ we are showing now results for 4×4 networks and using the following cloning template [Chua88a]:

$$\begin{matrix} 0.0 & 1.0 & 0.0 \\ 1.0 & 2.0 & 1.0 \\ 0.0 & 1.0 & 0.0 \end{matrix}$$

Let us consider the initial states of Fig.2a. Assume first $\tau = 10$. After *50* iterations the states of Fig.2b are reached which, under a maximum error of *0.5%*, coincide to the final states shown in [Chua88a] for the same example. Assume now $\tau = 5$, The states of Fig.2b are obtained after *20* iterations. A maximum deviation of *1%* can now be measured. Assume finally $\tau = 1$. After just two iterations the states of Fig.2d are reached, which exactly coincides to the theoretical ones. Making $\tau = 1/2$ is the limit ensuring stability for this example. In the most general case the limiting value for stability is given by $\tau > A_o/2$ with $A_o < A(i,j;i,j)$. It is important to point out that parasitics in the herein proposed implementation do not introduce alterations in the nominal dynamics given by (2). Thus, τ can be selected in practical designs to the value ensuring maximum operation speed.

3. Building Blocks for Current-Mode CNNs

The basic building block for the circuits to be presented herein is the *current-mirror*. More specifically, our circuits are based on the proper exploitation of the elementary current-mirror applications illustrated in Fig.3. Fig.3a. illustrates a

0.8	0.8	0.8	-0.8
0.8	0.8	0.8	-0.8
0.8	0.8	0.8	-0.8
-0.8	-0.8	-0.8	-0.8

(a)

3.98	4.98	2.99	-1.99
4.98	5.98	3.98	-2.99
2.99	3.98	1.99	-2.99
-1.99	-2.99	-2.99	-3.98

(b)

3.96	4.95	2.97	-1.98
4.95	5.94	3.96	-2.97
2.97	3.96	1.98	-2.97
-1.98	-2.97	-2.97	-3.96

(c)

4	5	3	-2
4	6	4	-3
3	4	2	-3
-2	-3	-3	-4

(d)

Figure 2: (a) Initial states, (b)-(d) Final states for different values of τ

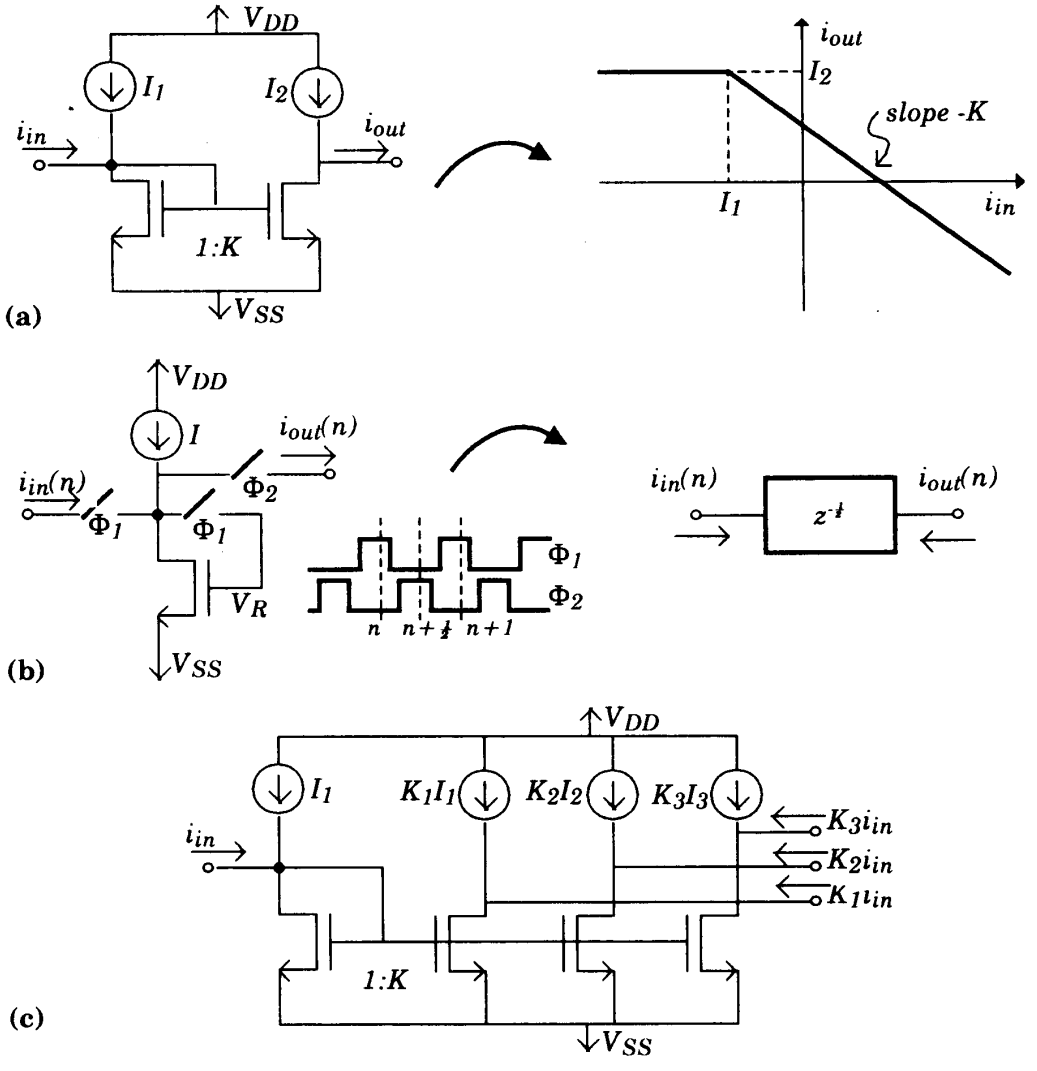

Figure 3: Elementary applications of current mirrors

principle for *nonlinear function generation* by means of current-mirrors. The constant current sources shown at the top of the circuit schematics can be obtained from a reference current by using p-channel mirrors. A convex nonlinear i_{out} vs. i_{in} characteristics result for this circuit. Concave characteristics can be realized in a similar way. Combining convex and concave characteristics any continuous piecewise-linear function can be synthesized. Fig.3b illustrates the implementation of a *delay operator* by using a dynamic current mirror. Two non-overlapping clock signals are required to control the analog switches. While ϕ_1 is high, a reference voltage, V_R, is created and stored in the gate parasitic capacitor. Then, when ϕ_1 becomes low and ϕ_2 high a delayed replica of the input current is obtained as determined by the stored reference voltage. Finally, Fig.3c illustrates the principle of *weighted replication* by means of multiple current-mirrors. A common voltage reference, V_R, is created from the input current and applied to the gate of different apropriately sized transistors.

The building blocks for our current-mode CNN circuits use the ideas above together with proper circuit techniques to enhance accuracy. We are using the current mirrors of Fig.4. Fig.4a shows the schematics for an enhanced cascode n-channel switched current mirror. The amplifier is used for the purpose of, on the one

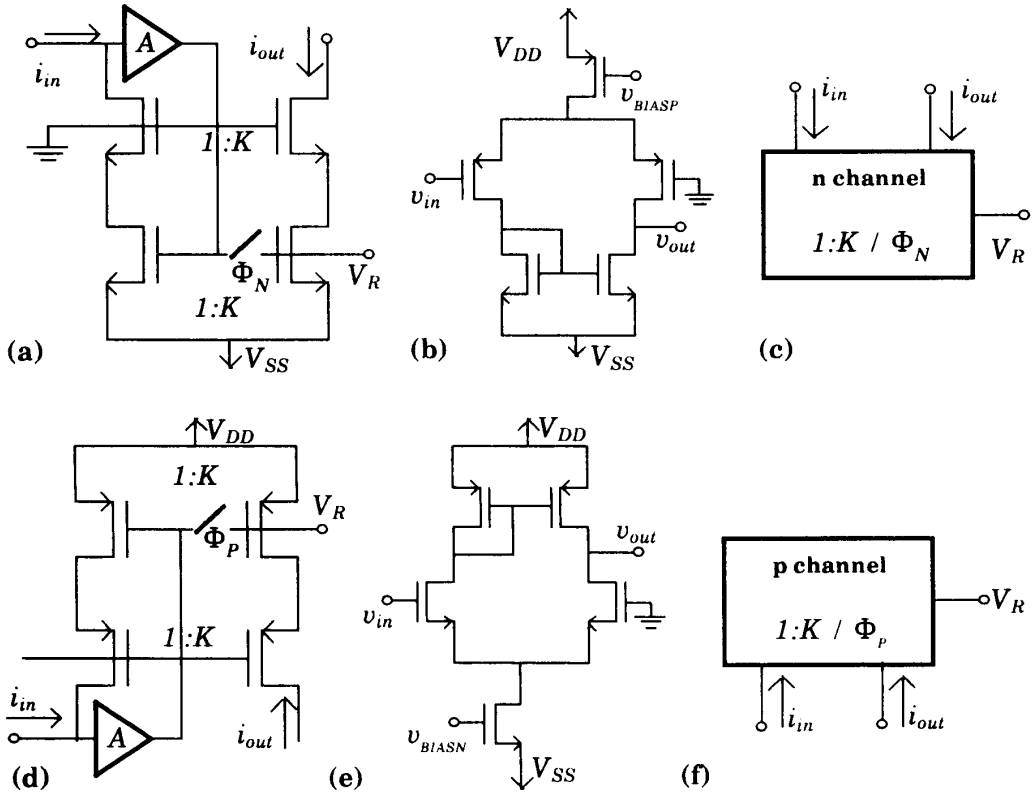

Figure 4: Current mirrors for the CNN building blocks

hand, reducing the mirror input impedance and, on the other, isolating the gate capacitor of the bottom transistors from the input signal. The amplifier schematics is shown in Fig.4b. The switch in Fig.4a can be implemented by a single NMOS transistor. Fig.4c shows a symbol for the switched n-channel mirror. Figs.4d-f show the corresponding schematics and symbol for the p-channel mirror. For both mirrors, the output current can be replicated by using the reference voltage V_R. Fig.5 shows the schematics of the nonlinear operator using the blocks of Fig.4. For this operator switching of the mirrors is not required. The switch transistor of Fig.4a can hence be substituted by a single wire. Fig.6 shows the schematics for the integrator block. Using these blocks and multiple output mirrors it is possible to implement the cell block diagram of Fig.1. By replication of this basic cell any arbitrary CNN can be then implemented.

The performance of the proposed blocks has been validated by extensive simulations using transistor parameters for a 3μm technology and the program HSPICE [Meta88]. By the way of example, Fig.7 shows the simulated characteristics for the nonlinear block. The blocks have been optimized to operate at clock frequencies up to 500Khz. Larger frequencies are expected when using scaled down technologies. Simulations have been also performed for a 4×4 network and considering different templates, the simulated performance being for all the cases in accordance to what is predicted by the model.

References

[Chua88a] L. O.Chua and L. Yang: "Cellular Neural Networks: Theory". *IEEE Trans. Circuit and Systems*, Vol.35, pp 1257-1272, October 1988.

[Chua88b] L.O. Chua and L. Yang: "Cellular Neural Networks: Applications". *IEEE Trans. Circuit and Systems*, Vol.35, pp 1273-1290, October 1988.

[Lipp87] R.P. Lippmann: "An Introduction to Computation with Neural Nets". *IEEE Trans. ASSP Magazine*, Vol.4, pp 4-24, April 1987.

[Mead89a] C.A. Mead: "*Analog VLSI and Neural Systems*".Addison Wesley 1989.

[Mead89b] C.A. Mead and M. Ismail: "*Analog VLSI Implementtion of Neural Systems*". Kluwer Academic Publishers, 1989.

[Meta88] "HSPICE Users Manual". MetaSoftware Inc., 1988.

[Morg90] N. Morgan: "*Artificial Neural Networks: Electronic Implementations*". IEEE Computer Society Press, 1990.

[Krieg90] K.R. Krieg et al.: "Analog Signal Processing using Cellular Neural Networks". *Proc. IEEE ISCAS90*, pp 958-961, 1990.

[Yang90] L. Yang et al.: "VLSI Implementation of Cellular Neural Networks".*Proc. IEEE ISCAS90*, pp 2425-2427, 1990.

[Rodr90] A. Rodríguez-Vázquez et al.: "Nonlinear Switched-Capacitor "Neural" Networks for Optimization Problems". *IEEE Trans. Circuits and Systems*, Vol. 37, pp , March 1990.

[Tank86] D.A. Tank and J.J. Hopfield: "Simple Neural Optimization Networks: An A/D Converter, Signal Decision Circuit, and a Linear Programming Circuit". *IEEE Trans. Circuit and Systems*, Vol. 33, pp 533-541, May 1986.

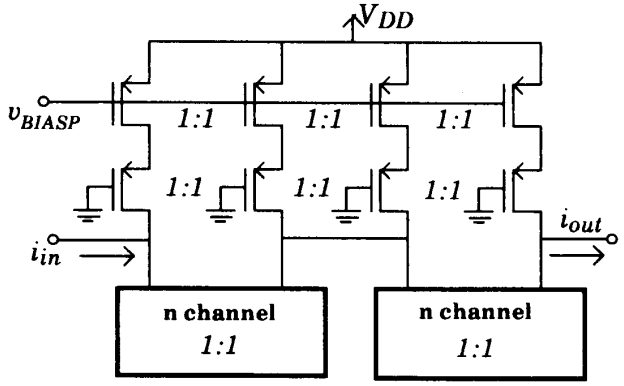

Figure 5: Nonlinear operator

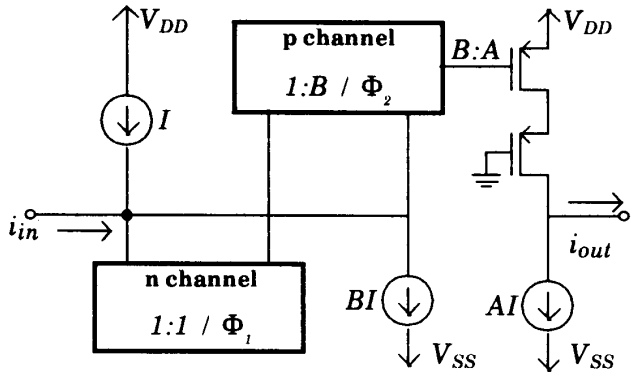

Figure 6: Integrator block

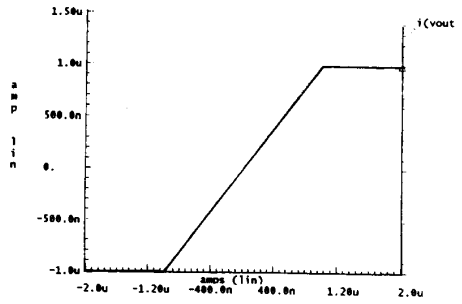

Figure 7: Simulated characteristics for the nonlinear block