# Hardware solutions supporting a multiscale approach in early vision

G.M. Bisio[a], M. Confalone[a], L. Raffo[b], S.P. Sabatini[a]

[a] PSPC Research Group-DIBE University of Genova Via all'Opera Pia 11A - 16145 Genova
[b] DIEE-Dept. of Electrical and Electronic Engineering University of Cagliari
Piazza d'armi - 09123 Cagliari
E-mail: pspc@dibe.unige.it

**Abstract -** Recurrent networks are efficient computational structures for embedded analog filtering operations in smart vision sensors. Still, their use requires adequate control of various network parameters to achieve flexibility without incurring instability. In this paper, we analyze the behaviour of bounded recurrent networks, and we find how to modify interconnections at boundary nodes to avoid artifacts in the profiles of the resulting convolution kernels, particularly when the kernels span over a wide number of pixels. This solution can be implemented at circuital level just by varying the ratio of current mirrors at boundary cells, without modifying their inner arrangement.

## 1 Introduction

Analog VLSI vision chips are expected to provide very efficient solutions for a broad range of early vision tasks [1]. However, even when some specific restricted functionality is addressed (e.g., stereo depth maps for autonomous robot navigation, detection of discontinuities in the optical flow for surveillance in a dynamic environment), major advances have to be made to have solutions for real-world applications. One of this concerns processing flexibility of the hardware core; particularly appealing are those solution which achieve flexibility still through an analog approach.

Object in the world, and details in images are meaningful over a certain range of scale. Therefore, robust processing of visual data requires a *multiscale* strategy. Typically, early vision tasks can be formulated as interactions with data through some operators. It is thus necessary to be able to control one or more parameters of the operator to represent the signal at different scales.

In this perspective, current analog VLSI vision sensors suffer, to various extent, of the following drawbacks: (i) uncapability of analyzing image data at different scales, since the size of the operator is frozen in the circuital architecture; (ii) low resolution, since the pixel size of active imagers tends to increase with the complexity of the task to be performed (even worse if one introduces some degree of flexibility); (iii) more critical design when stressing architectural solutions towards large scale operators.

In this paper, we present a solution to these problems for a network of processing elements (lattice network) with tunable interconnection strength [2]. Specifically, we address the problem of controlling the accuracy of multiscale operators at the boundaries of this kind of recurrent network. This will be achieved giving proper specificity to the interconnection strength of processing elements located at the boundary of the imager.

## 2 Bounded Lattice Network

### 2.1 Computational model

We consider a discrete 1-D form of spatially continous linear operators specified in differential and/or integral form mapped on a network of interacting nodes (i.e. lattice network), with 2nd-order degree of locality. Every node of the lattice network is excited by the visual signal $s(n)$ and is connected to the nearest and the next nearest neighbor nodes [6]. The functionality of the lattice network can be described by the following 2nd-order linear difference equation:

$$\begin{aligned} a_{-2}e(n-2) + a_{-1}e(n-1) + a_0 e(n) & + \\ +a_1 e(n+1) + a_2 e(n+2) & = s(n) \quad (1) \end{aligned}$$

where $e(n)$ is the output, $s(n)$ the input, $a_i$ and $a_{-i}$ are coefficients describing the interactions among the nodes of the 1-D lattice network. The 1-D lattice network is characterized by a stable impulse response $h(n)$ of the following type:

$$\begin{aligned} h(n) = Ce^{-|n|\lambda}((1+e^{-2\lambda})sin(\omega_0)cos(\omega_0 n) + \\ +(1-e^{-2\lambda})cos(\omega_0)sin(\omega_0|n|)) \quad (2) \end{aligned}$$

where $\omega_0$ is the spatial frequency, $\lambda$ is related to the spatial extension, and $C$ is a normalization constant. In Fig. 1 the output values for three different pairs of the interaction parameters are shown. Their values are chosen to obtain filters of different spatial frequencies, but with the same number of lobes, to keep constant the relative bandwidth.

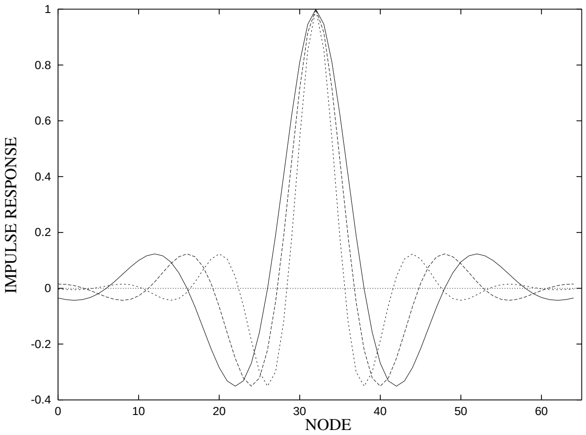



Figure 1: Normalized iso-lobe impulse responses ($\omega_0/\lambda = 3$) of the network for three pairs of ($-a_1/a_0$,$a_2/a_0$) interaction coefficients ratios. The selected points are $(0.675, 0.177)$, $(0.678, 0.182)$, $(0.686, 0.195)$, corresponding to the following pairs of filter parameters $(\omega_0, \lambda)$: $(\pi/10, \pi/30)$, $(\pi/8, \pi/24)$, $(\pi/6, \pi/18)$.

## 2.2 The termination cells

The previous analysis are based on the necessary assumptions that the system was linear and space invariant, i.e. the network must be infinite. However, one could argue that if the stimulus is applied far enough from the boundaries the network behaviour would be similar to that of the infinite case. This is either a great limitation in the use of the network or, in general, a wrong assumption as clearly indicated with an example in [5]. Nevertheless it is possible to compensate the boundary effects making a sort of *spatial matching* so that the physical finite network behaves the same way of the infinite theoretical one.

Let us first consider a network of infinite extent but finite stimulus distribution (for each $m \geq n+1$ $s(m) = 0$) as depicted in Fig. 2.

When $m \geq n+1$ the node equations (see Eq. 1) are homogeneous. We can solve such subsystem of equations, and, for each $m \geq n+1$, we can determine $e(m)$ as a linear combination of $e(n)$ and $e(n-1)$ for the boundary values. In particular for $m = n+1$ and $m = n+2$ we have:

$$\begin{aligned} e(n+1) &= Pe(n) + Qe(n-1) \\ e(n+2) &= Re(n) + Se(n-1) \end{aligned} \tag{3}$$

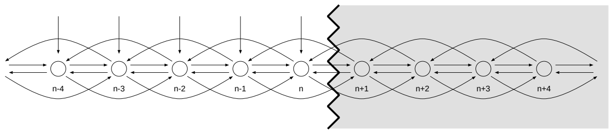



Figure 2: Bounded 1D 2nd-order network. The obscured part on the right is introduced to simulate the effect of the boundary.

This result can be interpreted the other way around, if we observe that the obscured part of the network $(m \geq n+1)$ influences the rest of the network only through the status of nodes $n+1$ and $n+2$. Consequently, we can reformulate the equation for node $n$ and $n-1$ removing all the boundary effects:

$$\begin{aligned} & a_2e(n-2) + (a_1(1+Q) + a_2S)e(n-1) + \\ + \quad & (a_0 + Pa_1 + Ra_2)e(n) = b_0s(n) \end{aligned} \tag{4}$$

$$\begin{aligned} a_2e(n-3) \quad + \quad & a_1e(n-2) + (a_0 + Qa_2)e(n-1) + \\ + \quad & (a_1 + a_2P)e(n) = b_0s(n-1) \end{aligned} \tag{5}$$

To find the values of the coefficients $P$, $Q$, $R$, and $S$ we consider a generic input distribution, determine the corresponding output, and identify the coefficients in Eq. 3:

$$\begin{aligned} P &= 2cos(\omega)e^{-\lambda} \\ Q &= -e^{-2\lambda} \\ R &= P^2 + Q \\ S &= PQ \end{aligned} \tag{6}$$

Summarizing, we can compensate the effects of boundary modifying the values of the interaction parameters of the two nodes close to the boundary, according to the Eq.s 4 and 5. Fig. 3 shows the effects of the termination cells in the case of the widest impulse response of Fig. 1; the network is excited at short distance from the boundary, under three conditions: (i) truncated network, (ii) unbounded network (theory), (iii) properly terminated network. There is a perfect overlapping of the responses in the case (ii) and (iii).

## 2.3 Circuital model of the termination cells

Let us follow a current mode approach and consider as the input of a single cell a current proportional to the incident light at that point (e.g. photosensor current); the signals propagated among the cells are currents too, so we can rewrite Eq. 1 as follows:

$$\begin{aligned} I_e(n) \quad = \quad & I_s(n) + GI_e(n-1) + GI_e(n+1) - \\ - \quad & KI_e(n-2) - KI_e(n+2) \end{aligned} \tag{7}$$

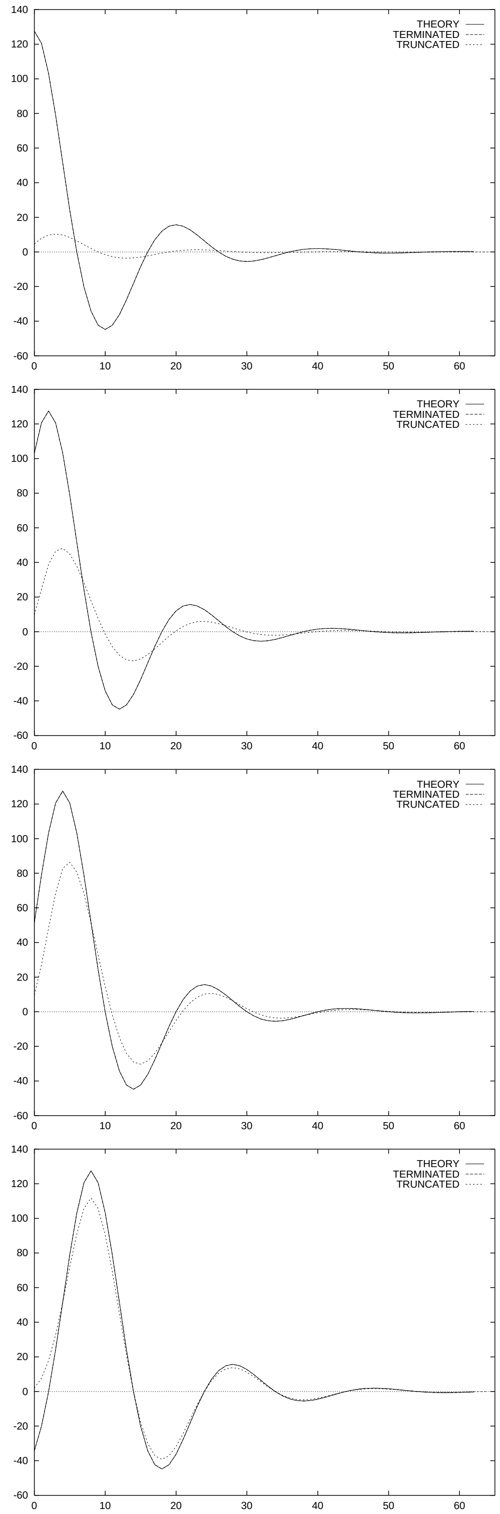



Figure 3: The impulse response of the network excited at node 0, 2, 4 and 8, from the boundary, under three conditions: (i) truncated network, (ii) unbounded network (theory), (iii) properly terminated network. There is an absolute coincidence for the responses in case (ii) and (iii).

$$e(n) = \frac{I_e(n)}{I_{norm}}, \qquad s(n) = \frac{I_s(n)}{I_{norm}}$$

where $I_s(n)$ is the normalized input current (i.e. stimulus), $I_e(n)$ is the normalized output current (i.e. excitation), $G = -a_1/a_0$ and $K = a_2/a_0$. According with Eq. 7 a single cell can be interpreted as a simple node in which the currents (i.e. the contributions) from the neighbour cells and the input current sum with each other (Kirchhoff current law): at the same time the cell has two output currents $GI_e(n)$ toward the nearest nodes and two (negative) output currents $KI_e(n)$ toward the next nearest nodes. The circuit structure of a node is based on the use of Current Controlled Current Sources (CCCSs) with current gain $G$ and $K$ determined according to Eq. 1 for the inner nodes and Eq.s 4- 5 for the termination cells. Fig. 4 shows the internal structure of the termination cells. The first termination cell (node $n$) has only two lattice interaction parameters:

$$\begin{aligned} G_{n,n-1} &= \frac{a_1(1+Q)+a_2 S}{a_0 + Pa_1 + Ra_2} \\ K_{n,n-2} &= \frac{a_2}{a_0 + Pa_1 + Ra_2} \end{aligned} \tag{8}$$

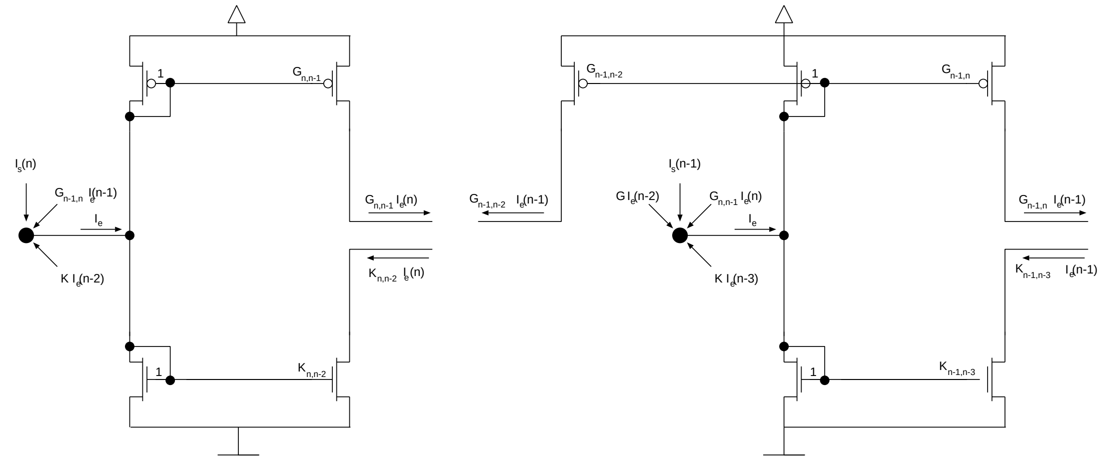
The second termination cell (node $n-1$) has three lattice interaction parameters:

$$\begin{aligned} G_{n-1,n} &= \frac{a_1 + a_2 P}{a_0 + a_2 Q} \\ G_{n-1,n-2} &= \frac{a_1}{a_0 + a_2 Q} \\ K_{n-1,n-3} &= \frac{a_2}{a_0 + a_2 Q} \end{aligned} \tag{9}$$

It is important to note that the normalization coefficients for the input signals of the two termination cells are different from those of the inner ones.

# 3 Conclusions

The architectural and circuital solutions for the implementation of visual microsystems, presented in this paper rely upon the collective computation of a lattice network where the excitation value of a node is related to the excitation of its neighbors. The resulting operators can be interpreted as iterative filters that detects specific characteristics present in the input pattern of excitation. The spatial domain on which these characteristics are detected is possibly larger than that of physical interconnections; this occurs both directly, by local interactions, and indirectly, through the propagation property of iterative filters. To overcome the deterioration of wide-extension kernels, due to the boundary effects, we introduce proper termination cells to modify the values of the interconnection parameters




(a) (b)

Figure 4: Internal circuital structure of the node $n$ (a) and $n-1$ (b): the black dot represents the node which is stimulated by the photocurrent $I_s$ and receives inhibitory ($K$'s) and excitatory ($G$'s) contributions from neighboring cells.

of the two cells close to the boundary and emulate an infinite network. Accordingly, spatial kernels of different extensions can be obtained with high accuracy on a network of limited extent. In this way, we can reach real-time computational efficiency in the measuring process on the visual space, by using a coarse-to-fine strategy to take into account that different characteristics of an image reveal themselves at different levels of resolution [4]. These operations are typical in computer vision algorithms where each image is analysed through a bank of filters of different size and tuned to different spatial frequencies (multichannel approach) [7]. With reference to such multi-scale algorithms and on the ground of the neurophysiological and psychophysical findings in natural vision [3], a VLSI system able to perform, on the same substrate, multichannel analysis of the sensed image is really desirable, and may have challenging impact on the functionality and performance of artificial vision microsystems.

**Acknowledgment** This work is partially supported by CEC-Esprit-BRA Project CORMORANT, ASI Project *Microsistemi Ottici per la Robotica Spaziale* and MURST 40% *Sistemi microelettronici e microsistemi*.

# References


[1] X. Arreguit and E.A. Vittoz. Perception systems implemented in analog VLSI for real-time applications. In *Proc. of the PerAc'94 Conference: From Perception to Action*, Lausanne, Sept.5-9, 1994, 1994.

[2] F. Bruccoleri, L. Raffo, S.P. Sabatini, and G.M. Bisio. A tunable perceptual microsystem for stereo depth estimation. In *2nd IEEE-CAS Region 8 Workshop on Analog and Mixed IC Design*, Baveno, Italy, September 1997.

[3] J.G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Amer.*, A/2:1160–1169, 1985.

[4] J.J. Koenderink. Scale-time. *Biol. Cybern.*, 58:159–162, 1988.

[5] T. Matsumoto, H. Kobayashi, and Togawa Y. Spatial versus temporal stability issues in image processing neuro chips. *IEEE Trans. Neural Net.*, 3(4):540–569, July 1992.

[6] L. Raffo. Analysis and synthesis of resistive networks for distributed visual elaborations. *Electronics Letters*, 32(8):743, April 1996.

[7] H. Wechsler. Multiscale and distributed visual representations and mappings for invariant low-level perception. In H. Wechsler, editor, *Neural Networks for Perception I*, pages 462–476. Academic Press, 1992.