CMOS DESIGN OF CELLULAR APAPs and FPAPAPs: AN OVERVIEW

A. RODRÍGUEZ-VÁZQUEZ[a]

Instituto de Microelectrónica de Sevilla-CNM-CSIC. Avda. Reina Mercedes s/n
41012 Sevilla (SPAIN). Tel.: +34 955056666, Fax: +34 955056686.
E-mail: rcarmona@imse.cnm.es

CNN-based analogic visual microprocessors have similarities with the so-called Single Instruction Multiple Data systems [4], although they work directly on analog signal representations obtained through embedded optical sensors and hence do need neither a front-end sensory plane nor analog-to-digital converters. The architecture of these visual microprocessors is illustrated in Fig. 1 through two prototype chips, namely: ACE4K [2] and ACE16K [5]. In both cases, as in other related chips [3 6 7 8], the architecture includes a core array of interconnected elementary processing units, surrounded by a global circuitry. This latter circuitry is intended for:

- Control and timing.
- Addressing and buffering of the core cells.
- Input/output.
- Storage of user-selectable instructions (programs) to control the sequence of operations of the processing core.
- Storage of user-selectable analogic programming parameter configurations (templates).

On the other hand, the core of interconnected processing units embeds different functions on a common silicon substrate (see Fig. 2 for illustration purposes), namely:

- 2-D sensing.
- 2-D analog/digital array processing concurrent with the signal sensing.
- 2-D spatio-temporal processing determined by local, receptive-field-like programmable interconnections.
- 2-D memory banks for concurrent on-line uploading and downloading of short-term analog and digital data.

Several analogic visual microprocessor chips in different CMOS technologies have been reported during the last few years. Table 1 presents a summary of some of the most relevant data of those implementations with at least 20×20 pixels. Some columns correspond to chips intended for black and white input images, while others are for chips which accept gray scale input images. As with any other analog processing circuit, figures of merit about performance must contemplate accuracy and area occupation in addition to speed and power consumption. Any comparison must refer to the number of operations per second and to the accuracy. The data in the table highlights the following:

- There is a trade-off between area occupation (cell density) and accuracy, on the one hand, and speed of operation and power consumption, on the other. This trade-off is typical of analog integrated circuits [9].

a. This overview is mostly based on the work of Servando Espejo, Rafael Domínguez-Castro, Ricardo Carmona and Gustavo Liñán, and has been supported by ESPRIT V Project IST-1999-19007 and by ONR/NICOP Grant N00014-00-10429 (POAC), and the Spanish CICYT Project TIC-1999-0826.

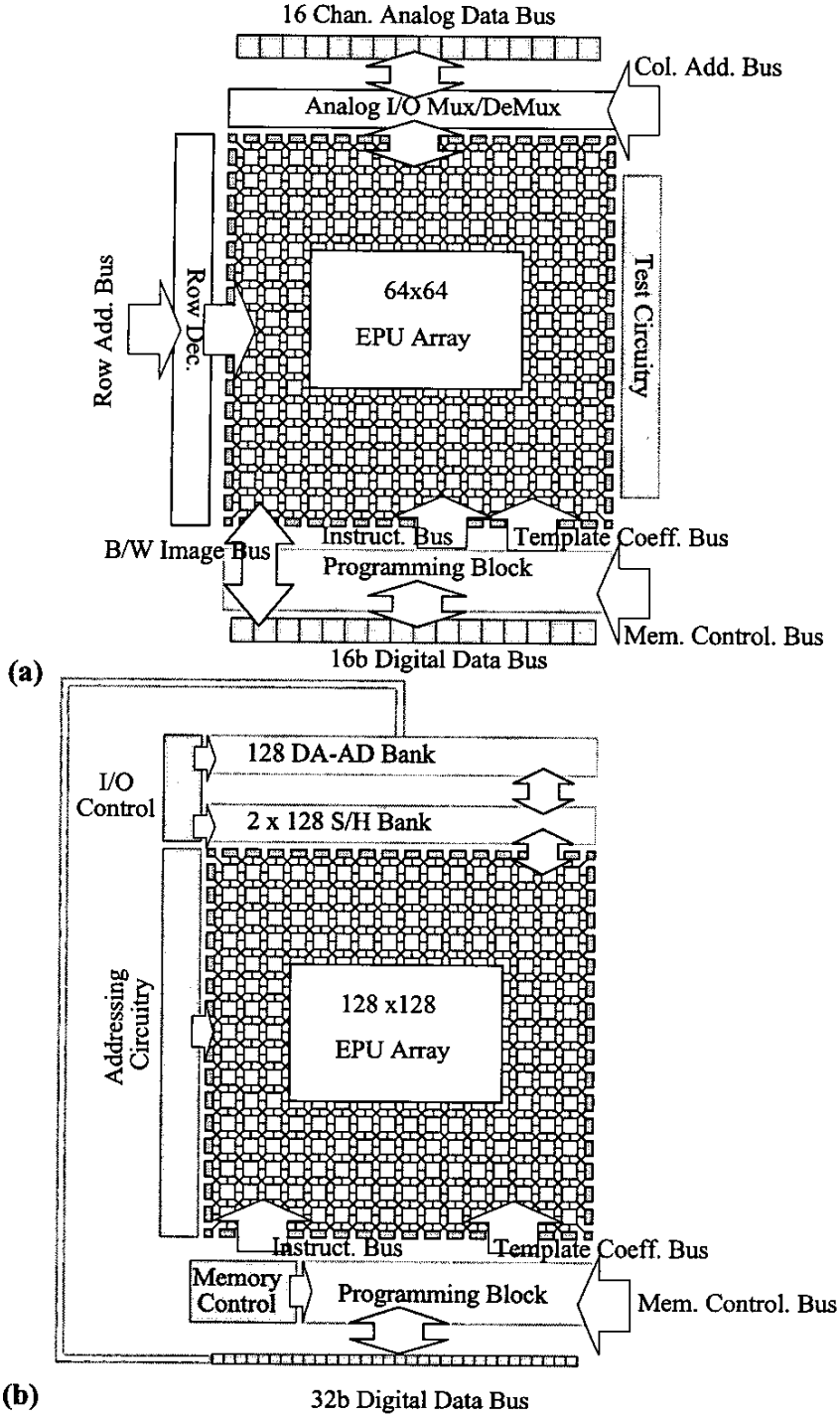

Figure 1. Architectures of Analogic Visual Microprocessor Chips: (a) ACE4K [2], (b) ACE16K [5].

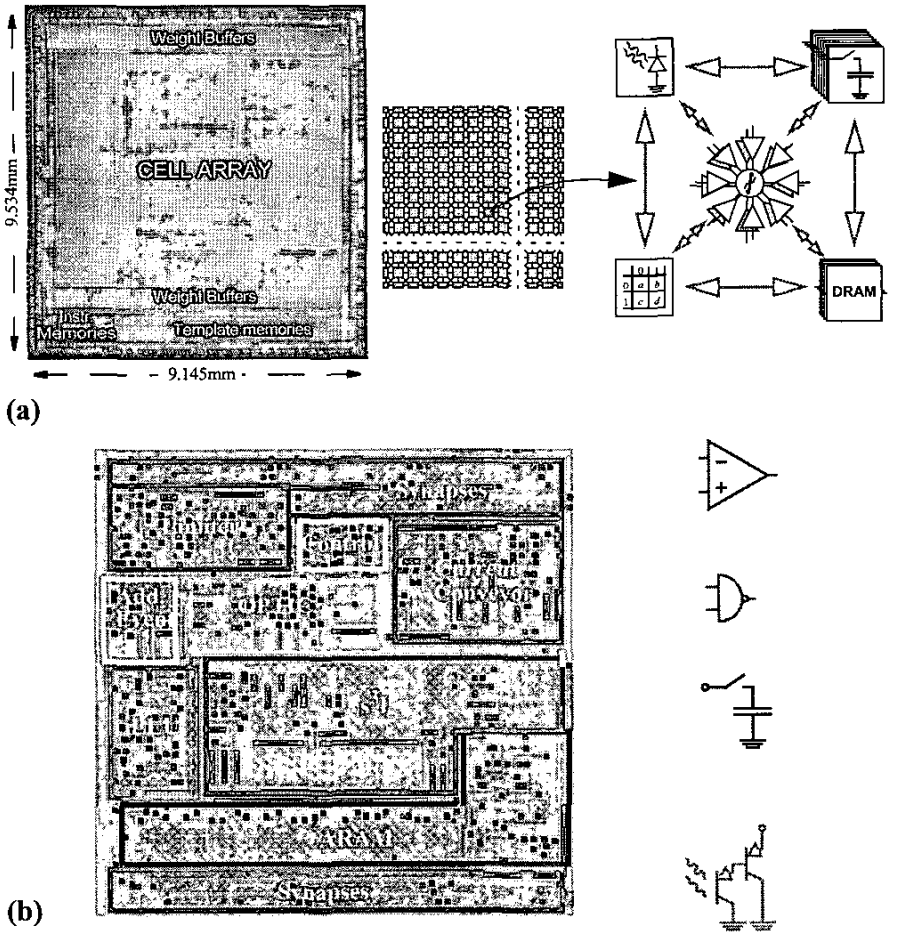

Figure 2. Illustrating the embedding of different functional features at the core processing array of visual microprocessors. (a) Microphotograph of the ACE4K chip (left figure), and conceptual representation of the distributed functions embedded in the core array (right figure); (b) Layout of a processing unit of the ACE16K showing the areas occupied by the different functions realized concurrently by the core array.

- The evolution towards scaled-down technologies reports advantages in terms of speed and cell density. Actually, the ACE16K chip has 128×128 resolution and is capable of realizing sequences of 64 instructions; using up to 32 different templates (each template consisting of 24 8-bit-coded analog programming values) during a sequence; loading and downloading full-size gray-scale images to and from the cache memory, and having always 8 full-size images available for usage during the flow; with an internal processing time of 160ns, and providing digitally-coded output images (obtained with a battery of internal A/D converters) with a downloading time of 0.128ms.

Table 1 Summary and comparison of chip implementations

Ref.	7	6	8	2	5
Tech, μm	0.8	0.7	0.5	0.5	0.35
Array Size	20 x 22	20 x 20	48 x 48	64 x 64	128x128
Pix. Format[a]	B	A[b]	B	A[c]	A[d]
Weight. Prog.[e]	8-b	Continuous [-4,4] c	6-b	8-b	8-b
Memory per Cell	4 LLMs 1 State Capacitor	1-State 1-Input	2 In/2 Out Registers	4 LLMs 4 LAMs 1-State Cap 1-Inp. Cap	8 LAMs 2 LLU Op. 2 Flags 3 Pix. Cap.
Multipliers per cell[f]	9, 1[g]	5, 5, 1[h]	9, 9, 1	9, 9, 2	9, 1, 1, 1
Photo Sensors.	Yes	No	No	Yes	Yes Multimode Sensor
Program. Memory	8 Templ.	No	1 Templ.	32 Templ. 64 Digital Instructions	32 Templ. 4096 Dig. Instructions
τ	250ns	5μs	50ns	1.2μs CNN 280ns Conv	0.8μs CNN 160ns Conv
Cells/mm^2	27.5	16.7	295	82	180
Power (W)	~1 W	375 μW/cell	300mW (max.)	250μW/cell 1.2 W Chip	180μW/cell 4W Chip
Speed	15.8 GOPS	0.53 GOPS	0.5 TOPS	40 GOPS CNN	0.19 TOPS CNN
XPS/area	0.98 GOPS/mm^2	22 10^6 OPS/mm^2	64 GOPS/mm^2	1 GOPS/mm^2	3.6 GOPS/mm^2
XPS/Pow (OP/J)	1.58 10^{10}	3.5 10^9	1.6 10^{12}	3.95 10^{10}	8.25 10^{10}
Electr. I/O	22 Lines Binary Bus	20 Lines Analog Bus	48-b Binary Bus	16b B. Bus 16 Lines Analog Bus	32b Digital Data Bus

a.A=Analog, B=Binary (B/W), D=Digital
b.Only B/W results are available.
c.7.7b Equivalent Accuracy.
d.8b Equivalent Accuracy.
e.It refers to the number of bits used to define weight parameters.
f.**A**, **B**, and z multipliers.
g.**A** and **B** multipliers are the same. The chip uses a time-multiplexing scheme.
h.Cross-shape neighourhood.

The capability to design cells with maximum density, speed and accuracy, and minimum area and power consumption relies basically on the exploitation of all functional features offered by the MOS transistor. This is very different from digital design, in which only the switching capability of the MOS transistor is exploited. The design of the entities which interconnect the cells (synapses) defines one of the major issues. In order to do this, different possibilities may be chosen a priori. In all cases electrical controllability is provided by default. However, the different strategies exhibit quite a different performance in the presence of systematic and random error sources, as well as a different incidence of the global signal transmission errors. Hence, careful analysis and optimization is needed to select the best approach. Such analysis and optimization are needed to achieve the cell density and accuracy levels featured by last generation chips. The background for such procedures can be found in literature [1 2 3 5 7].

References

1. T. Roska and A. Rodríguez-Vázquez (Editors), *Towards the Visual Microprocessor*. John Wiley & Sons Ltd., 2000.
2. G. Liñán, P. Földesy, S. Espejo, R. Domínguez-Castro and A. Rodríguez-Vázquez, "A 0.5μm CMOS 10^6 transistor analog programmable array processor for real-time image processing". *Proc. of the 1999 European Solid-State Circuits Conference*, pp. 358-361, September 1999.
3. R. Carmona, P. Garrido, R. Domínguez-Castro, S. Espejo and A. Rodríguez-Vázquez, "Bio-inspired analog vlsi design realizes programmable complex spatio-temporal dynamics on a single chip". *Proc. of the 2002 Conference on Design and Test in Europe*, to appear.
4. J.C. Gealow and C.G. Sodini, "A pixel-parallel image processor using logic pitch matched to dynamic memory". *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 831-839, June 1999.
5. G. Liñán, R. Domínguez-Castro, S. Espejo and A. Rodríguez-Vázquez, "ACE16K: An advanced focal-plane analog programmable array processor". *Proc. of the 2001 European Solid-State Circuits Conf.*, pp. 216-219, Villach (Austria), September 2001.
6. P. Kinget and M. Steyaert, *Analog VLSI Integration of Massive Parallel Processing Systems*. Kluver Academic Publishers, 1997.
7. R. Domínguez-Castro et al., "A 0.8μm CMOS 2-D Programmable mixed-signal focal-plane array processor with on-chip binary imaging and instruction storage". *IEEE J. Solid-State Circuits*, Vol. 32, pp. 1013-1026, 1997.
8. A. Paasio, A. Dawidziuk, K. Halonen and V. Porra, "Minimum size 0.5μm CMOS programmable 48×48 CNN test chip". *Proc. of the 1997 European Conference on Circuit Theory and Design*, pp.154-156, Budapest, September 1997.
9. A. Rodríguez-Vázquez, E. Roca, M. Delgado-Restituto, S. Espejo and R. Domínguez-Castro, "MOST-Based Design and Scaling of Synaptic Interconnections in VLSI Analog Array Processing Chips". *J. of VLSI Signal Proc. Systems for Signal, Image and Video Technology*, Vol. 23, pp. 239-266, Kluwer Academics November/December 1999.