

IMPLEMENTATION OF SIMD VISION CHIP WITH 128x128 ARRAY OF ANALOGUE PROCESSING ELEMENTS

Piotr Dudek

School of Electrical and Electronic Engineering, University of Manchester
PO Box 88, Manchester M60 1QD, United Kingdom
p.dudek@manchester.ac.uk

ABSTRACT

This paper presents the latest implementation of the SIMD Current-mode Analogue Matrix Processor architecture. The SCAMP-3 vision chip has been fabricated in a 0.35 μ m CMOS technology and comprises a 128x128 general-purpose programmable processor-per-pixel array. The architecture of the chip is overviewed and implementation issues are considered. The circuit design of the analogue register is presented, the layout of the Analogue Processing Element is discussed and the design of control-signal drivers and readout circuitry is overviewed.

1. INTRODUCTION

Programmable vision chips, which combine image sensors and pixel-parallel processors, can offer significant advantages in many computer vision applications, providing image pre-processing capability with low power consumption and high computational performance. Inherently parallel low-level image processing algorithms map naturally to fine-grained pixel-parallel processor arrays, while the ability to perform computations right next to the sensors reduces the power consumption and I/O bandwidth requirements. However, the requirement of physically co-locating photosensor array and processor array, in a processor-per-pixel manner, introduces severe constraints in terms of the physical design of the device, circuit area, and power consumption. This is especially true if a low-cost requirement implies the use of a standard CMOS technology. To meet the implementation constraints, mixed-mode and analogue circuits are usually employed in the design of the processor. It has been demonstrated [1-3] that use of analogue processors can offer significant benefits in terms of cost, processing speed and power consumption, when limited accuracy of processing is acceptable. The implementation of a large analogue VLSI circuit, however, is a challenging task. The trade-offs between cell area, power dissipation and processing speed have to be suitably resolved. A full-custom design is obviously required, and the constraints on the physical geometry of the circuit topology are an important consideration affecting the circuit and system design. This implies a co-design of processor architecture, circuitry and layout. Furthermore, to achieve a robust implementation, the issues of noise and mismatch have to

be carefully controlled. Again, the architecture and physical design have to be considered together.

In this paper the design of a 128x128 pixel-per-processor vision chip (SCAMP-3, shown in Figure 1) is presented. The chip has been fabricated in a 0.35 μ m CMOS technology and comprises over 1.8 million transistors (most of which are working in analogue mode). The design is based on the design used in our previous 39x48 array chip, which has been reported in [4]. Some details of the processing element design and readout architecture were presented in [5] and [6]. In this paper, several aspects of the chip implementation are elaborated. In particular, detailed schematic diagram of the analogue register circuit is presented, the layout of the processing element is discussed and the control/readout circuitry is overviewed.

2. CHIP ARCHITECTURE

The architecture of the chip is depicted in Figure 2. The pixel-parallel image processing capability is provided by a 128x128 array of analogue processing elements (APEs). Each APE is a simple processor, which includes nine registers (eight general-purpose ones and one used for exchanging data with nearest neighbours), a comparator with an activity-flag latch, and I/O circuits. Each APE also contains a photodetector circuit, so that a 128x128 image sensor array is embedded within the processor array. The array operates in SIMD

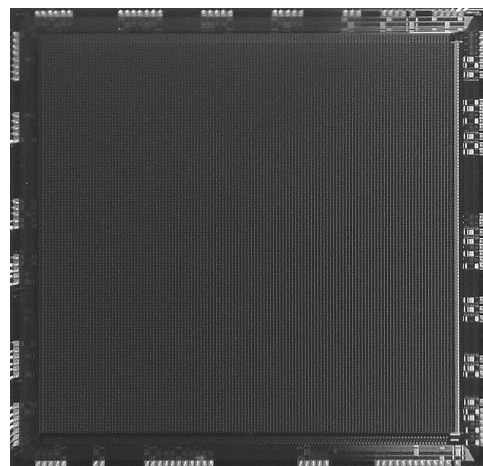


Fig. 1. The SCAMP-3 chip

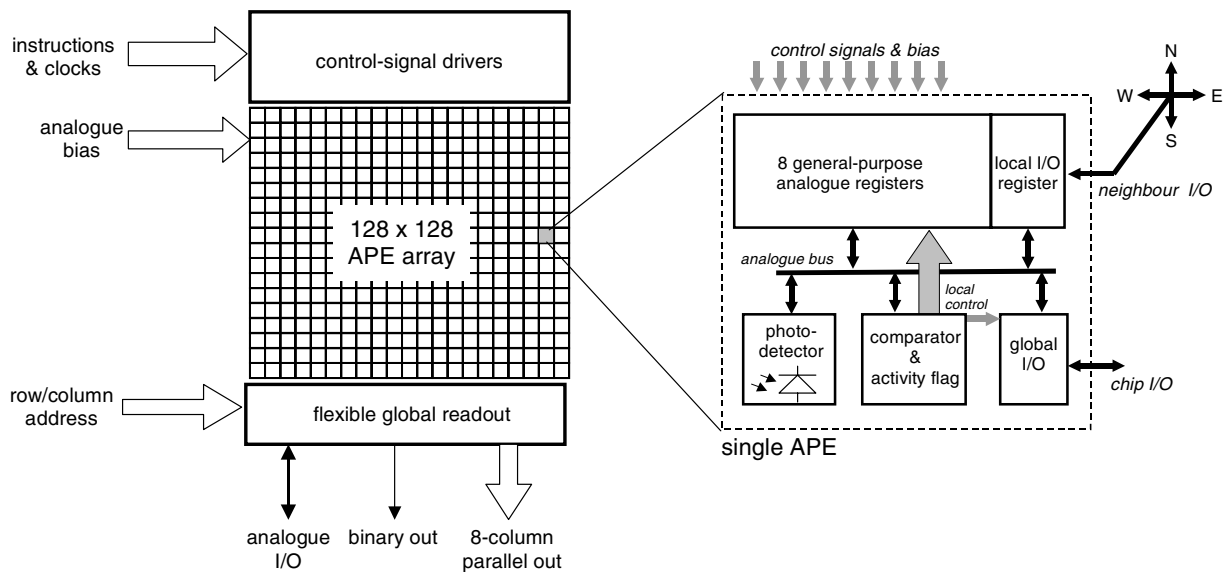


Fig.2. Architecture of the SCAMP-3 vision chip and its Analogue Processing Element (APE)

(Single Instruction Multiple Data) mode, the micro-instructions are issued by a single (external) controller and distributed to all APEs in the array via drivers located at the periphery of the array. The processing results are read-out from the array via flexible global readout circuitry, which enables pixel-addressing for analogue, binary and column parallel readout as well as global summation and logic OR operations.

3. IMPLEMENTATION

The most critical aspect of the overall chip design is the design of the processing element. This is where the constraints on the cell size and power consumption have to be carefully weighted against desired functionality, performance, and accuracy levels. Fortunately, small silicon area available to a single processor implies a simple cell structure, which makes a co-design of architecture, circuitry and layout of the device a manageable task. It has to be said, that when designing a general purpose processor, the trade-offs between speed, power, area and accuracy are resolved in a somewhat arbitrary way, guided to some extent by the envisaged applications: in this design we decided to minimize the size of the APE, so that reasonably high-resolution array (128×128) was feasible on a 50 mm² chip. At the same time, low power consumption was a priority (peak power below 250 mW per chip), and accuracy had to be kept at an acceptable level to perform low-level image processing. As a compromise, a moderate speed of operation, 1.25 MIPS (million instructions per second) per cell was achieved. Nevertheless, the massively parallel operation results in 20 GIPS (giga-instructions per second) per chip, which compares quite favourably with conventional digital signal processors. This speed is

sufficient for performing, for example, 1250 operations per pixel at 1000 frames per second, which should be sufficient for majority of computer vision applications. On the other hand, when executing simpler algorithms, e.g. working at 20 frames per second with 200 instructions per pixel, the power consumption of the chip could be reduced to below one milliwatt. Therefore, the chip should be particularly suited to low-power requirements of battery-powered systems.

The APEs are implemented as “analogue micro-processors” [1], they execute software instructions, but achieve this using analogue circuits, and store data in a form of analogue sampled signals. Switched-current signal processing techniques are employed to achieve “ALU-free” design, i.e. no dedicated hardware exists to implement arithmetic operations. Instead, the basic operations of addition, inversion and division are executed directly in the registers/analogue bus system [5]. This results in significant silicon area savings and consequently also improves the accuracy of processing. Further discussion of benefits of this approach was presented in [7].

The layout of the APE is shown in Figure 3. The APE occupies silicon area of 49.35µm×49.35µm, most of this area is occupied by registers. The overall accuracy is ensured by a mixture of hardware and software techniques. The registers were designed as S²I memory cells [8], as depicted in Figure 4. Large-area transistors M_{MEM} and M_{REF} are used to improve accuracy - using long and narrow transistors leads to a nominal biasing current of 1.7 µA with good voltage swing. The analogue power supply voltage is 2.5 V. The cells are laid out to minimize the capacitive coupling onto

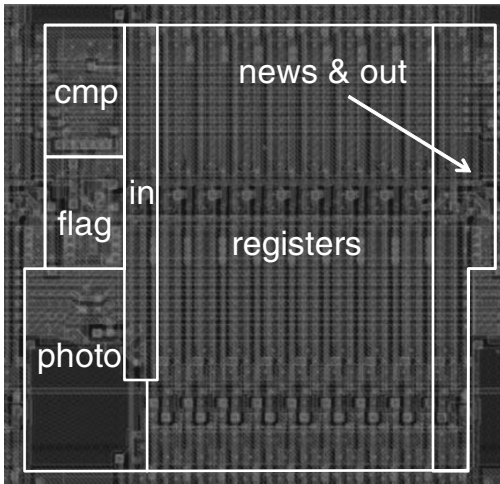


Fig.3. The APE layout. The cell size is $49.35\mu\text{m}\times 49.35\mu\text{m}$

sensitive nodes from control signals (which are routed over the register area), and adjacent registers. As can be seen in Figure 3, each register extends for the full height of the APE cell. The register is connected to the analogue bus using transistors M_{SN} and M_{SP} . Consequently, the biasing current is switched off when register is not used, this minimizes power consumption. The storage is enabled by switches M_{W1} and M_{W1R} (operating during phase 1) and M_{W2} (operating during phase 2), according to the S^2I technique [8]. Conditional opening of these switches (controlled by a state of the local activity flag register) is implemented using transistors M_{F1} and M_{F2} .

The register associated with the neighbour communication is laid out on the right of the register bank (see Figure 3) to minimize the differences in errors when storing data in this register, as compared with other registers. It connects to analogue buses of the neighbours via four transistor switches (in North, East, West and South direction). The input circuit, placed to the left of the register bank, uses a circuit similar to the register cell, with transistors M_{MEM} and M_{REF} , and thus provides layout-level matching of the surrounding of the first register in the register bank.

The current comparator is implemented using simple voltage differential amplifier, with gate of one transistor in the pair connected to the analogue bus and the gate of the second transistor biased by a dc voltage (this voltage is equal to the nominal analogue bus voltage on phase 2 of the storage cycle, its value is close to V_{REF}). The comparator is followed by a latch, which serves as the activity flag register.

Photodetector circuit uses an n-type diffusion diode and can operate in integration mode, with close-to-linear characteristic, or in continuous-time log-compression mode. The pixel fill-factor is approximately 5.6%.

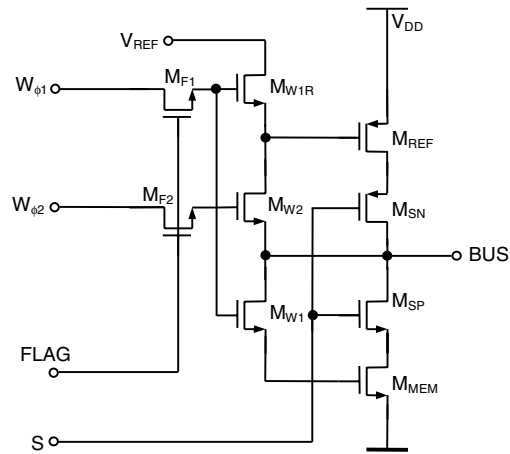


Fig.4. Schematic diagram of the register cell.

Control signals, analogue bias voltages and power supply lines have been routed over the APE area. The power supply and ground planes should ideally be used, but these have not been possible in the 3-layer metal technology. Instead, a power supply grid has been implemented. The registers have been shielded to prevent signal coupling from the digital control signals. The metal layers have been also used as an optical shield, so that only the photodiode area was exposed to light.

Apart from a careful design, to minimize errors related to clock-feedthrough, charge injection and output conductance of the S^2I cells and fixed-pattern noise errors caused by mismatch, a number of techniques (e.g. correlated double sampling, cancellation of signal-independent errors, algorithmic compensation of division error) can be implemented in software to improve the overall accuracy of processing.

The simple architecture, algorithmic error cancellation techniques, and careful circuit design result in a robust operation of the processor, which has been confirmed by our experiences with the previous 39×48 chip. Control signals are digital, except one analogue voltage that provides input value to the APE. No adjustments of instruction parameters are required – the programs which are designed in a simulator can be directly transferred to the hardware and execute as expected. The accuracy is, of course, limited, but we have successfully implemented linear and non-linear filters (e.g. median filter, gray-level morphology), orientation-selective filters, edge detectors, motion detectors, active contours, hole filling, skeletons, models of autowave propagation in excitable medium, etc. The measurement results from the 128×128 chip are not yet available, but the experimental results from the 39×48 chip, with identical APE design, have been reported in [4] and it

expected that the overall performance of this implementation will be similar to that of the previous design.

4. CONTROL AND READOUT

Each APE requires 40 control signals and 10 analogue voltages (bias and power supply). Digital control signals are derived from instruction-code-words that are provided to the chip from an external controller. The relatively low frequency of control signals (1.25 MHz) simplifies the design of the signal distribution network. Control signals are routed horizontally and vertically, with one buffer/driver for each row and column of the array. All analogue bias voltages are driving high-impedance nodes, providing either fixed dc bias to gates of transistors, or at most they are required to provide relatively small charge supplement at low frequency. The bias voltages are provided externally and routed globally.

The output interface has been designed to allow binary image readout at 80 Mpixel/second, via 8-bit column-parallel port. During binary read-out the state of the activity flag register is read. A grey-scale image read-out is designed to operate at 1 Mpixel/second (it is expected that in typical high-frame-rate applications reading-out of full grey-scale images will not be necessary). During grey-scale read-out the value of the output current from an analogue register is routed to the chip I/O pin, it is converted to a digital value using an external A/D converter.

A flexible global readout architecture [6] permits addressing groups of APEs in the array, to facilitate global operations. This is simply implemented using address decoder similar to the one shown in Figure 5. In a usual address decoder, the N -bits of the address bus are hard-wired to the inputs of the N -input 'AND' gate, corresponding to the physical address of the respective row/column, using either straight or inverted bit of the address signal A_i for each input i . If, instead of using inverted bits of the address A_i , an independent signal B_i is used, then the usual addressing is extended by the

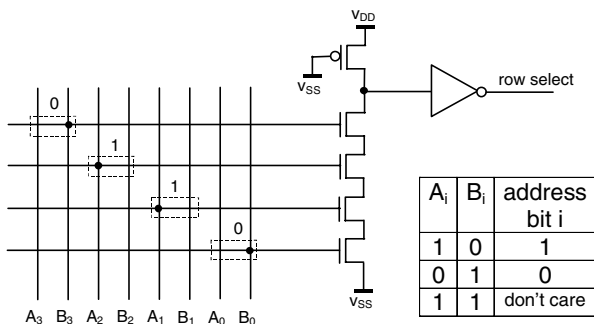


Fig.5. Address decoder. In this simplified example, a 4-bit decoder for row no.6 (i.e. "0110") is shown.

capability of an equivalent *don't care* bit in the address word, as illustrated in the table in Figure 5. Consequently, a number of rows/columns in the array can be addressed at the same time. In binary read-out mode, a logic OR operation is performed on the selected group. In analogue read-out mode, a summation of the currents from selected APEs is performed (high dynamic range of the output currents, resulting from the possibility of addressing one, many, or all 16k APEs at the same time, has to be handled by an external variable-gain amplifier). This simple scheme allows rapid calculation of global image descriptors (e.g. pixel counts, histograms), control if iterative procedures (e.g. loop until none of the pixels have changed in one iteration), multi-resolution read-out with pixel binning, extraction of object coordinates, etc.

5. CONCLUSIONS

The design of a 128x128 general-purpose pixel-per processor array vision chip has been presented in this paper. The chip has been fabricated in a 0.35 μ m technology. At the time of writing this paper the chip has not been tested yet. However, the scaled-down version of this chip, containing a 39x48 array, is operating successfully. Experimental results from the 128x128 chip are expected to be available for presentation at the conference in May 2005.

ACKNOWLEDGEMENT

This work has been supported by the EPSRC, under grant no. GR/R52688/01

REFERENCES

- [1] P.Dudek and P.J.Hicks, "A CMOS General-Purpose Sampled-Data Analogue Processing Element", in IEEE Transactions on Circuits and Systems - II: Analog and Digital Signal Processing, vol. 47, no. 5, pp. 467-473, May 2000
- [2] A.Dupret, J.O.Klein and A.Nshare, "A DSP-like analogue processing unit for smart image sensors", in Int. Journal of Circuit Theory and Applications, vol.30, pp.595-609, 2002.
- [3] G. Liñán-Cembrano et al. "A Processing Element Architecture for High-Density Focal-Plane Analog Programmable Array Processor Arrays", Proc. ISCAS 2002, vol.III, pp.341-344, May 2002.
- [4] P.Dudek, "A 39x48 General-Purpose Focal-Plane Processor Array Integrated Circuit", Proc. ISCAS 2004, vol.V, pp.449-452, Vancouver, May 2004
- [5] P.Dudek, "A Processing Element for an Analogue SIMD Vision Chip", Proc. European Conference on Circuit Theory and Design, ECCTD'03, vol.III, pp.221-224, September 2003.
- [6] P.Dudek, "A Flexible Global Readout Architecture for an Analogue SIMD Vision Chip", Proc. ISCAS 2003, Bangkok, vol.III, pp.782-785, May 2003.
- [7] P.Dudek, "Accuracy and Efficiency of Grey-level Image Filtering on VLSI Cellular Processor Arrays", Proc. CNNA 2004, pp.123-128, Budapest, July 2004
- [8] J. B. Hughes and K. W. Moulding, "S²I: A Switched-Current Technique for High Performance", in Electronics Letters, vol.29, no.16, pp.1400-1401, August 1993.