# A General-Purpose CMOS Vision Chip with a Processor-Per-Pixel SIMD Array

Piotr Dudek and Peter J. Hicks

*Department of Electrical Engineering and Electronics*
*University of Manchester Institute of Science and Technology (UMIST)*
*P.O. Box 88, Manchester M60 1QD, United Kingdom*
*E-mail: p.dudek@umist.ac.uk, p.j.hicks@umist.ac.uk*

## Abstract

*The paper discusses the architecture and implementation of a new SIMD focal-plane processor array integrated circuit. The chip employs switched-current "analogue microprocessors" as processing nodes in a digital-like massively parallel computer architecture. Using analogue processing elements allows the achievement of real-time image processing speeds with high efficiency in terms of silicon area and power dissipation. The prototype 21×21 SCAMP vision chip is fabricated in a 0.6µm CMOS technology and achieves a cell size of 98.6µm×98.6µm. The approach is compared with state-of-the-art vision chips build using digital SIMD arrays and CNN-based processors. Experimental results are presented.*

## 1. Introduction

To meet computational demands of computer vision algorithms, particularly if cost, size, and power dissipation of the system are important, it is often beneficial to perform some image processing directly on the focal plane, using a smart-sensor device. Some simple low-level image processing tasks can be implemented using dedicated analogue circuits embedded within each pixel of the image sensor array [1,2]. By assigning a processor to each pixel of an image the inherent fine-grain parallelism of low-level image processing tasks can be fully exploited. The application-specific hardware solutions, however, lack the flexibility of a software-programmable computer, with its ability to implement a variety of complicated algorithms using relatively simple hardware. The main difficulty in implementing a software-programmable pixel-per-processor vision chip is the very limited area available for each processor in the array. Some of the chips reported in the literature use single-bit processors [3,4], however, due to their limited capabilities (a few bits of memory per processor) they can be hardly considered "general-purpose". There have been also attempts at achieving vision chips with more complex bit-serial processors [5]. An interesting alternative to digital processors is provided by "analogic" processors [6,7], derived form the CNN (Cellular Neural Network) architecture by augmenting it with analogue and digital memories. Our approach combines a conventional digital architecture with analogue circuitry through the "analogue microprocessor" concept [8]. This yields a particularly good compromise between cell area and functionality, performance and power dissipation. In this paper the architecture and design of our vision chip [9] is discussed and compared with other approaches.

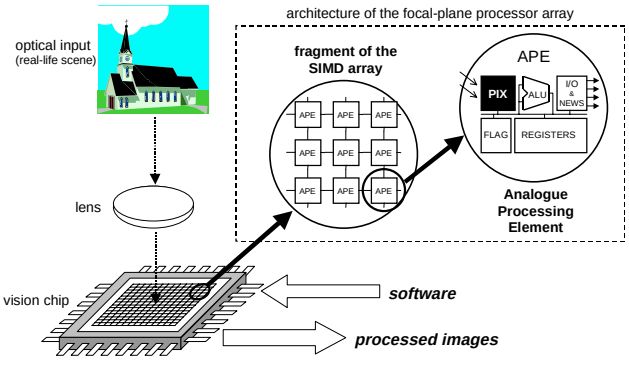



*Fig.1. A programmable SIMD focal-plane processor array*

## 2. SCAMP Architecture

The general concept of our chip, named SCAMP (SIMD Current-Mode Analogue Matrix Processor), is presented in Fig.1. The processing core is a mesh-connected array of processors, which are called APEs (analogue processing elements). This name reflects the fact, that data is represented and manipulated inside the APEs as analogue samples. However, the operation of the system is equivalent to that of von Neumann's cellular automata and similar to many digital massively parallel computers. The APEs execute identical instructions on their local data in an SIMD (Single Instruction Multiple Data) fashion. As the processor array size corresponds to the image size, and instructions are performed on an entire array at once, it is convenient to represent the architecture as consisting of several register-planes (see Fig.2). Each register-plane **A-K** can hold a grey-level image or another array variable. Transfer instructions (for example **A←B**) represent the transfer of an image from one plane to the other. Similarly, arithmetic operations (e.g. **A←B+C**) perform pixel-wise arithmetic operations on the data planes. The

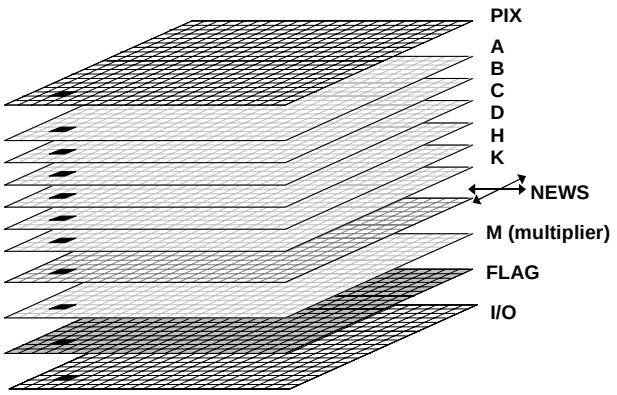



*Fig.2. The SCAMP architecture (a single APE is marked)*

array supports inversion and summation of any number of arguments in a single instruction, executed in a single clock cycle. Multiplication (scaling) is performed via a special-purpose multiplier register **M**.

Communication between four nearest neighbours in the array is facilitated via a special-purpose **NEWS** register. The array also supports random-access input and output. Additionally, entire row, column or entire array can be addressed for read-out, resulting in a global summation operation. This feature is very useful for monitoring the state of the entire array and also greatly simplifies the design of global algorithms, such as histograming.

Image acquisition is supported via a special-purpose register-plane **PIX**. The value held in this register-plane corresponds to the state of the photodetector array, which works in an integration mode. Non-destructive read-out ensures that multiple exposure times are possible, which can be used to extend the dynamic range of the image sensor.

As in the majority of SIMD processor arrays, local autonomy is supported by the activity flag register. This register can be set or reset depending on the result of a comparison operation. Only those APEs which have the **FLAG** register set perform broadcast instructions.

## 3. VLSI Implementation

The architecture outlined above requires a processing element of significant complexity. It is very important, if focal-plane device with a reasonable resolution is considered, to minimise the silicon area of the processing cell. We have achieved high integration level by employing switched-current analogue processing elements, based upon the "analogue microprocessor" concept.

The circuitry of the APE is described in more detail elsewhere [8,9], here we will only provide a brief overview, contrasting the analogue approach with the digital one. Firstly, it has to be noticed, that the general architecture of the APE (see inset in Fig.1) is akin to that of a digital processing element. Each APE includes registers, arithmetic-logic unit (ALU), I/O port, activity flag register and photodetector, all connected via an internal data bus. As a result of consecutive instructions data is transferred between the registers, or manipulated inside the ALU, in a universal Turing machine fashion. However, in the APE, data is represented not by 1's and 0's of a digital computer, but by analogue samples. Registers are built as switched-current memory cells; data processing is also performed in current-mode on analogue samples. This yields immediate advantages in terms of performance, silicon area and power dissipation. Firstly, only a single capacitor is required to store an analogue variable (whereas a digital system requires $N$ capacitors to store an $N$-bit integer). Secondly, only one wire is required for the analogue bus and other signal communication paths (whereas an $N$-bit digital processor requires $N$-wire buses, while a bit-serial processor requires more wires to address each bit of the storage space separately and multiple clock cycles for data transfer operations). Thirdly, addition in a current mode system is performed directly on the analogue bus (current summation in a node) with no need for explicit circuitry. Moreover, in contrast to bit-serial processors, addition is performed in a single clock-cycle. This also includes a useful operation of register-plane summation, where values of selected registers from all the APEs in the array are added together. Furthermore, the inversion operation is inherent in the current-memory cell and performed at no cost with each storage operation. Finally, which is important for a focal-plane device, the analogue processor can interface directly to the analogue signal from the photodetector, with no need for an A/D conversion.

These factors contribute to the fact, that APEs can be made more compact than equivalent digital processing elements. On the other hand, algorithmic programmability ensures that our chip is more versatile than any other analogue vision chip, with the notable exception of recently developed "analogic" processors [6,7] based on the CNN-UM concept. The CNN-UM is a SIMD/CNN hybrid - each processing cell of the CNN array contains additional local memory and can act as a cellular automaton. The SCAMP approach can be thus considered a degenerate case of the CNN-UM processor, inasmuch as it does not include the CNN core. In spite of this, it can perform all of the tasks that are performed by the CNN-UM while achieving arguably better balance between versatility, performance, cell area and power dissipation. One reason for this is that the CNN requires a large number of synapses (multipliers) in each cell. Even if they are implemented as single-transistors they still can occupy significant silicon area, due to accuracy requirements. Moreover, there is an inherent computational overhead when executing some practical image processing algorithms on the CNN-UM. The basic "instruction" in a CNN-UM system is a solution of a spatial-temporal non-linear differential equation, which is performed over the entire array within a few microseconds. This offers enormous computing power, if we consider how many "operations per second" are

required to simulate this system on a standard computer. However, a repeated solution of spatial-temporal non-linear differential equations is not necessarily the most efficient way to perform typical tasks such as edge detection, convolution, manipulation of binary images, etc. A simple sequence of arithmetic/logic operations and neighbour transfers is often all that is required to perform these tasks and so they can be efficiently implemented using more conventional algorithms on a typical SIMD array. Furthermore, many CNN algorithms require non-linear templates, a feature difficult to implement on a VLSI chip. Finally, the accuracy issue is very significant in practical implementations of a dense CNN array, which suffers from mismatch effects. This is also important in our chip, but less of a problem since the APE uses current-copier techniques, which are inherently immune to the mismatch problem. (Multiplication, performed using scaled current-mirrors is affected by mismatch, but the accuracy can be increased using register-based multiplication [10]). Also, fixed pattern noise effects originating from mismatch between transistors in the photodetector and input circuits can be effectively suppressed in the SCAMP chip using software-based correlated double sampling [9].

## 4. Experimental Results

A prototype SCAMP chip (Fig.3) was fabricated in a standard digital CMOS 0.6μm technology from AMS. The 10mm$^2$ chip comprises a 21×21 array of APEs, as well as random-access I/O logic, on-chip digital to analogue converter, and control logic. An external controller is required to store a program and provide a sequence of instructions to the SCAMP array. These instructions are decoded and distributed to the APEs using separate drivers for each row and column of the array, which makes it easy to scale-up the design to a larger array size. Each APE contains 128 transistors in a 98.6μm×98.6μm silicon area.

The photodiode area is equal to 820μm$^2$, which yields a fill factor of 8.4%. With 1000 lux illumination level full-contrast images are obtained at 25 frames/second. The measured fixed pattern noise of the imager, with correlated double sampling, is equal to 1% rms.

The APEs work with clock frequencies up to 2.5MHz, which yields a peak performance of over 1.1 GIPS (Giga Instructions Per Second) per 21×21 chip. The chip uses 3.3V (analogue) and 5V (digital) power supply. Peak power dissipation is below 40mW per chip, however it can be much reduced depending on the frame rate and algorithm being performed.

### 4.1 Accuracy

The design of the analogue circuitry of the APE involves trade-offs between size, power dissipation and accuracy of processing. It has to be noted that unlike digital processors, where the accuracy of operations is limited by the chosen word length, analogue processors have their accuracy limited by errors and noise inherent in the analogue circuitry. The magnitude of the signal-dependent error of a register transfer operation in the APE was measured to be equal to approximately 0.5% of the maximum signal level. Each transfer also contributes a noise of 0.11% rms. The rate of decay of analogue values stored in the registers, due to leakage currents, is equal to 0.19% per ms, at 125 lux. Although the accumulative effects of errors degrade the performance below the equivalent 7-bits accuracy suggested by the above figures, nevertheless for many low-level image processing algorithms this accuracy level is sufficient.

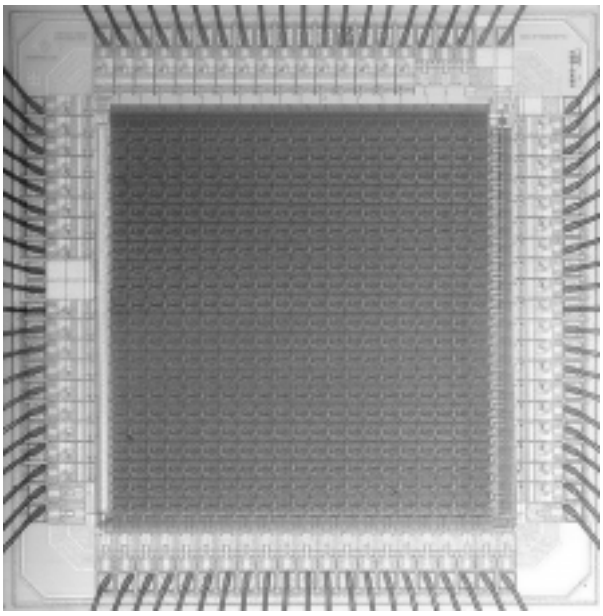
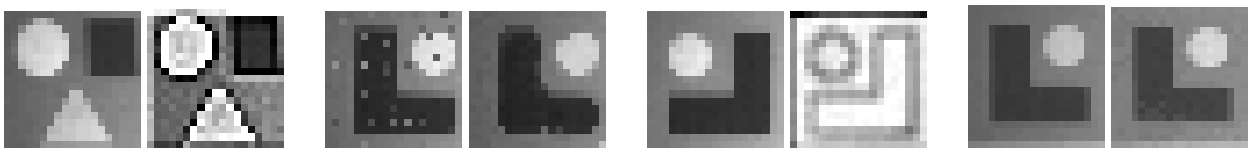


*Fig.3. Chip microphotograph*

### 4.2 Performance & Comparisons

The software-programmable architecture of our vision chip allows the implementation of a variety of low-level image processing tasks. We have successfully implemented and tested a number of algorithms, including convolution, linear and non-linear filtering, edge detection, segmentation, motion detection and estimation, histograming and histogram modifications, mathematical morphology and even Conway's Game of Life [10]. Some examples are presented in Figure 4. The execution times for several low-level image processing algorithms are listed in Table I. The fabricated 21×21 chip is a small-size proof-of-concept device. If the design were scaled to a 0.35μm technology, this would allow an integration of a full-size 256×256 processor array on a 250mm$^2$ chip. Such chip, clocked at 8MHz, would perform over 500 GIPS – a processing power that is two orders of magnitude higher than that of present-day microprocessors.

The maximum power dissipation is equal to 85μW per APE. However, as there is no DC current in an idle APE, power dissipation is much reduced when the time of processing is short compared with the frame rate. So, for



(a) (b) (c) (d)

*Fig.4. Image processing examples, Left: acquired image. Right: results of focal-plane processing on SCAMP chip: (a) sharpening, (b) median filter, (c) Sobel edge detection, (d) pixel-parallel 5-bit A/D D/A conversion chain.*

Table I. Time of execution of several algorithms on the SCAMP chip (not including read-out time).

| algorithm | execution time |
|---|---|
| Smooth using 3×3 convolution template | 5.6 μs |
| Sharpen using 3×3 convolution template | 6.0 μs |
| Edge detection with Sobel templates | 11.6 μs |
| Median Filter in 3×3 neighbourhood | 61.6 μs |
| Histogram with 64 bins | 205.6 μs |
| Motion estimation (21×21 global block search matching in x direction, with max. displacement ±3 pixels) | 46.4 μs |
| A/D converter (5-bit conversion, ramp) | 130.8 μs |
| D/A converter (5-bit conversion) | 11.2 μs |

example, while performing real-time edge detection at a frame rate of 25 frames/second we obtain power dissipation of 13nW per APE. The power dissipation figure can be therefore lower for our chip, than it is for some application-specific analogue vision chips, working in continuous time. Moreover, as the algorithmic program execution implies time-multiplexing of hardware resources, the APE area is not so much larger than the pixel area of many special-purpose vision chips, which implement algorithms in hardware [1,2].

The efficiency of our approach, in terms of processing speed, power dissipation and cell area can be compared with other programmable vision chips. Consider a state-of-the-art digital SIMD vision chip [5]. Its functionality is similar to that of the SCAMP chip. This digital chip performs edge detection and smoothing in 5.6μs and 7.7μs respectively, similar performance to our chip (although the quoted algorithms use only 4-bit numbers and simplified 4-neighbour templates), but the peak power dissipation (2.4mW per processing element) is 28 times larger. Although the bit-serial digital processing elements contain less memory than the APE (25-bits, which allows storage of only four 6-bit variables) the equivalent cell area (150μm × 150μm in 0.35μm CMOS) is over six times larger than that of the APE.

Considering further comparisons, it has to be noted, that some single-bit digital SIMD vision chips with limited memory [3,4] can achieve smaller cell area – however, they have very limited functionality as compared with the SCAMP chip. Similarly, the CNN-UM vision chip described in [6] is intended to process binary images only. The latest CNN-UM vision chip [7], however, can process grey-scale images. It contains processing nodes with 4 analogue and 4 binary memories (i.e. less local memory than the APE) but the cell area of 120μm×102.2μm in 0.5μm CMOS and power dissipation of 250μW per cell are still higher than these of the APE.

## 5. Conclusions

A general-purpose programmable vision chip that allows real-time focal-plane processing of grey-scale images has been presented. The SCAMP chip is an SIMD processor array with an analogue data-path. It attempts to combine, in the most efficient way, the flexibility of a software-programmable digital computer and high processing speed, low power dissipation and small cell area that can be achieved using analogue circuits.

## References


[1] S.Y.Lin *et.al.* "Neuromorphic vision processing system", in Electronics Letters, vol.33, no.12, pp.1039-1040, June 1997

[2] C.M.Higgins *et.al.* "Pulse-Based 2-D Motion Sensors", in IEEE Trans. on Circuits and Systems-II:Analog and Digital Signal Processing, vol. 46, no. 6, pp. 677-687, June 1998

[3] J.E.Eklund *et.al.* "VLSI Implementation of a Focal Plane Image Processor – A Realisation of the Near-Sensor Image Processing Concept", in IEEE Trans. on VLSI Systems, vol.4, no.3, pp.322-335, Sept.1996

[4] F. Paillet, D. Mercier, and T.M.Bernard, "Making the most of 15k$\lambda^2$ silicon area for a digital retina", Proc. SPIE, Vol. 3410, AFPAEC'98

[5] M.Ishikawa, K.Ogawa, T.Komuro, I.Ishii, "A CMOS Vision Chip with SIMD Processing Element Array for 1ms Image Processing", Proc. Conf. ISSCC'99, TP 12.2, 1999.

[6] R. Domínguez-Castro *et.al.*, "A 0.8-μm CMOS Two-Dimensional Programmable Mixed-Signal Focal-Plane Array Processor...", in IEEE Journal of Solid-State Circuits, vol.32, no.7, pp.1013-1025, July 1997

[7] G.Liñan *et. al.* "The CNNUC3: an analog I/O 64x64 CNN universal machine chip prototype with 7-bit analog accuracy", Proc. Conf. CNNA'2000 , pp. 201 -206

[8] P.Dudek and P.J.Hicks, "A CMOS general-purpose sampled-data analog processing element", IEEE Trans. Circuits and Systems-II:Analog and Digital Signal Processing, vol. 47, no. 5, May 2000, pp. 467-473

[9] P.Dudek and P.J.Hicks, "An SIMD focal plane analogue processor array", Proc Conf. ISCAS'2001, May 6-9 2001

[10] P.Dudek, "A programmable focal-plane analogue processor array", Ph.D. Thesis, UMIST, Manchester, May 2000.