A Processor Element for a Mixed Signal Cellular Processor Array Vision Chip

Stephen J. Carey, *Member, IEEE,* Alexey Lopich, *Member, IEEE*, and Piotr Dudek, *Senior Member, IEEE*

School of Electrical and Electronic Engineering, The University of Manchester, M13 9PL, UK
Email: p.dudek@manchester.ac.uk

Abstract — **A combined analogue and digital processing element for a pixel-parallel vision chip has been designed in 0.18μm CMOS technology. In addition to 7 analogue registers, each pixel incorporates 14 bits of digital memory. In the analogue domain its processing capabilities include addition, subtraction and squaring, with digital domain NOT and OR operators also available. The processing element has dimensions of 32x32μm and is designed to operate at 10MHz. A test chip has been fabricated.**

I. INTRODUCTION

THIS paper discusses a processing element (PE) that can be incorporated into VLSI vision chips. When assembled as an array of processor elements, the elements execute the same instruction in parallel upon local data (i.e. an SIMD computer). This approach has successfully been applied to image processing by many developers using analogue [1,2], digital [3] and mixed signal approaches [4]. Coupled with a chip controller, these ICs can be exploited to run sophisticated vision applications. These applications have included algorithms for edge and motion detection, and segmentation of objects. By pre-processing on the focal plane, raw image readout is not necessary; only processed images are read to external devices.

In the proposed PE, a mixed analogue and digital array of storage and sub-processor elements (in addition to a photodiode) is included allowing more efficient processing and storage of data dependant on data-type. This is the strategy advanced for the MIPA4k [4]. The PE described here maintains the architecture of a general purpose processor, allowing its usage to cover many different applications, while the pitch of the processing elements (PEs) should allow an IC with greater resolution at lower cost to be fabricated. The PE utilises switched current storage (as successfully developed in the SCAMP range of vision ICs [1]) for analogue storage and 3-transistor digital DRAM memories as deployed in ASPA vision chips [5]. The PE is designed for fabrication on a 180nm 6-metal, 1-poly process, which has brought significant challenges to the analogue design.

Advances in cellular processor array SIMD ICs have progressed from the IC developed by Bernard et al [6]. Our SCAMP3 [7] analogue vision chip (0.35μm) realised 128x128 arrays with cell area of 2435μm^2. The PE described here (0.18μm) has an area of 977um^2 with some previous work in processor arrays shown in Table I.

TABLE I. PROCESSING ELEMENT AREA FOR SIMD CELLULAR PROCESSOR ARRAYS

Work	Process	Cell dimensions/μm
PVLSAR2.2 [6]	2μm	80.5 x 100
NSIP[8]	0.8μm	60 x 60
ACE16K [2]	0.35μm	73.3 x 75.7
SCAMP3[7]	0.35μm	49.35 x 49.35
MIPA4k[4]	0.18μm	72 x 61
ASPA[5]	0.18μm	51 x 54
This work	*0.18μm*	*31.26 x 31.26*

Our PE design (in common with that reported by other vision chip designers [2]) was limited by the connectivity available to the PE, which, even with 6 metal layers, restricted the functions that could be included, with routing around the photodiode across all metal layers presenting the largest obstacle. While developments in 3D technology are progressing which should overcome vision-chip routing and low fill-factor issues [9], these technologies are not part of the mainstream and come at a cost. Hence, we are still left with the compromise of PE functionality vs number of PEs for a given die size. With several years of algorithmic and system control development behind SCAMP3 [e.g. 10, 11], the key features that have been demonstrated are: very low power operation, the low error of algorithmic operations, and long register lifetime. With the move to a smaller process node and smaller PE size, design focus is on the retention of these key features.

In the design, we aim to re-balance the memory, processing and data capture facilities within each PE to the algorithms that we have developed so far with SCAMP3. This has entailed reducing the number of analogue registers in favour of digital registers, adding diffusion and squaring systems and increasing the relative size and sensitivity of the photodiode.

II. ARCHITECTURE

The basic architecture of the PE is shown in Figure 1. The PE has dimensions of 31.26 x 31.26μm and includes 14 bits of memory (13 DRAM; 1 SRAM), 7 analogue registers, an analogue comparator, a diffusion network, a current squarer and a photodiode. The PE is addressed externally by means of a 61-bit instruction code word (ICW), with additional control supplied from 12 analogue inputs.
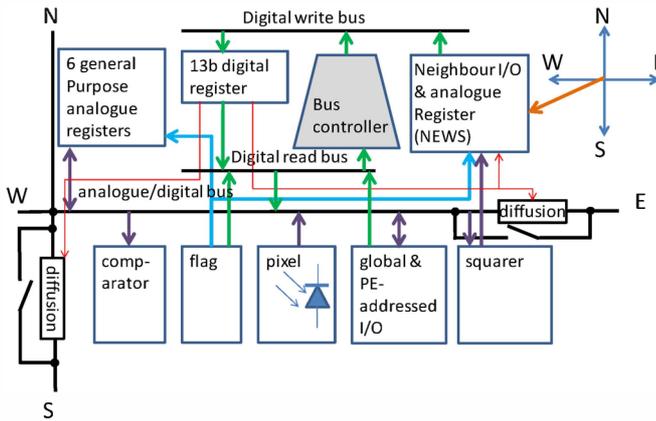
Figure 1. Architecture of the processing element

The PE, containing 174 transistors, includes digital and analogue sub-systems with the ability for either sub-system to access a common digital/analogue bus allowing exchange of data within the cell or to its orthogonal neighbours. This data bus can be written to directly by all the analogue sub-systems (except the squarer) and by the digital system. If an analogue signal is attached to the bus, the data is represented by current; if a digital signal is attached, data is a voltage.

The PE is designed to operate at 10MHz with an analogue supply voltage of 1.5V and digital supply of 1.8V.

A. Digital Sub-system

The main purposes of the digital block are to provide storage, execution-conditional mechanisms for analogue memory transfers, random access addressing capability and basic logic operations on binary data such as masks and markers. It also facilitates global asynchronous trigger-wave propagations [12] across the entire array, thus enabling efficient execution of complex operations such as object reconstruction, hole filling, watershed transformation and many others.
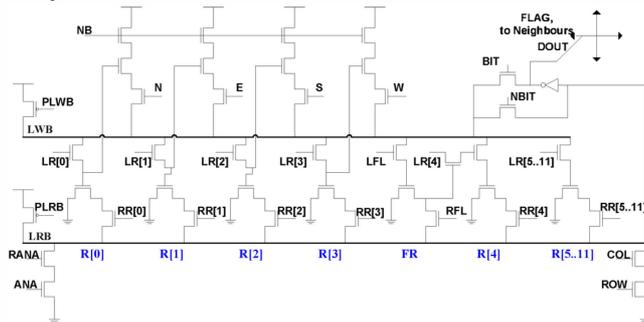

Figure 2. Digital memory in the processing cell.

The digital sub circuit incorporates twelve binary registers R[0…11], the flag register, basic logic gates and an interface for neighbour communication and addressing (Figure 2). For the purpose of area optimization, local memory is based on 3T dynamic latches. Logical data-path operations on local or neighbourhood data are performed on local data buses. The inverted data from the memory is read (by switching read signals *RR[0…11]* to logic '1') to the pre-charged Local Read Bus (LRB). If several registers are read simultaneously the

LRB operates as a logic NOR gate. The result is then transferred (directly or through an inverter) to the Local Write Bus (LWB), to which the memory inputs are connected. If any of the memory load signals, *LR[0..11]* are logic '1', then the data from the LWB is transferred to the corresponding memory element. Digital output *DOUT* is connected to the corresponding inputs (*N,E,S,W*) in four local neighbours. Neighbours' inputs are multiplexed to the LWB, so that logic OR can be executed during neighbour transfers. By multiplying signals pre-charging the LWB (*PLWB*) and the LRB (*PLRB*) with the φ_1 phase (Figure 4) of the clock and read and write instructions (*RR[0…11]* and *LR[0…11]*) with the φ_2 phase, all memory data transfers and logic operations are executed within one clock cycle.

In order to enable conditional execution and global asynchronous processing, the load operation for register R[4] is controlled by the local flag register (FR). The same flag register controls the propagation space for global asynchronous trigger wave propagations. In order to execute such operations, the initial marker has to be stored in register R[4]. After that, by continuously reading and writing (locally controlled by FR) to R[4] with appropriate neighbours selected, the marker propagates asynchronously from cell to cell within the propagation space. The mask, defining from which neighbours triggering is allowed, is stored in the local memory (R[0..3]), thus enabling local constraining of the propagation network topology. The result of the propagation is stored in register R[4]. Communication with the analogue sub-system as well as pixel addressing is performed identically to memory operations by discharging the LRB, with the result appearing at *DOUT*.

B. Analogue Sub-system

The analogue sub-system has much in common with the architecture of SCAMP3 [7]. The major differences are in a reduction in the number of analogue registers (from 9 to 7), the addition of a variable strength diffusion network and addition of a current squarer.

1) Analogue Memories

The analogue memories are based on the S^2I cell [13]. Locally-autonomous write operations are possible dependent on local FLAG state; if low, the parasitic gate capacitance of Mw1r, Mw1 and Mw2 keep the gate potential near ground during a write cycle, inhibiting the write operation and retaining the previous state of the register. The standard cell was modified (see Figure 3; timing in Figure 4) from our previous work [1] to include an error correction capacitor [14] to allow a modest reduction in cell size while maintaining low signal dependent error.

The error correction input compensates for the large clock feedthrough due to the switch transistors with the smaller size of the gate-source capacitance of the memory. This error is caused by a largely identical charge Δq being imposed on the memory transistor's gate (already charged with Q_{mem}) by the clock every time the register is written to. Due to the non-linear V_{gs}-I_{ds} characteristic of a MOS transistor, this fixed charge imposes a *variable* current error dependent on the

transistor's current V_{gs} value. Adding this charge back to the MOSFET can reduce this current error. Simulation results are in Figure 5. A fixed global EC level of 1V was used for the simulation over all input currents. The simulations indicate an error of <0.25% (of full scale, i.e. 8μA) should be achievable across the full register range for a double register transfer; i.e. executing the two sequential commands with pseudocode:

$$B = -A + \varepsilon_1$$
$$A = -B + \varepsilon_2$$

The double transfer error is given by $\varepsilon_2 - \varepsilon_1$, with inversion error defined as ε_1, with both these error metrics being a function of the initial A value. For the purposes of running applications on vision chips, the variation of these errors (i.e. $(\varepsilon_2(A_{hi}) - \varepsilon_1(A_{hi}))_{max} - (\varepsilon_2(A_{lo}) - \varepsilon_1(A_{lo}))_{min}$) over the operating range is of more importance than the absolute value of the error, since offset errors can be compensated for, while signal dependent errors can not.

The nominal biasing current for these cells is 4μA. The cost of the error correction is an extra track per register. For the analogue memory transistors, we use standard 1.8V MOS devices.

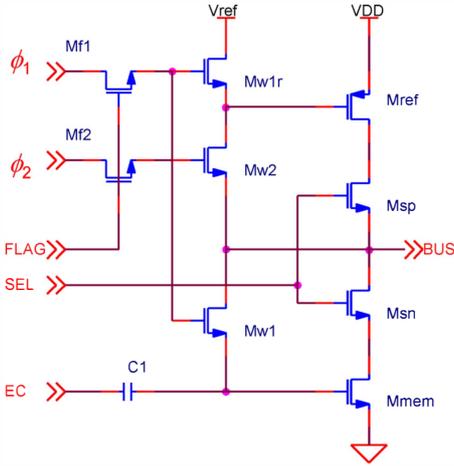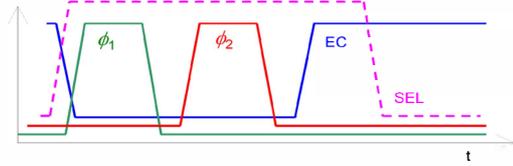Figure 3. Analogue register schematic

Figure 4. Timing of write cycle for the modified S²I current memory (signals are offset in y-axis for clarity; ϕ_1, ϕ_2 and SEL switch from 0V to 1.8V; EC is switched from 0V to a global potential between 0.7V and 1.8V)

1) Diffusion network

The diffusion network allows the execution of low pass spatial filters in a single instruction cycle. The network can be locally broken (in the horizontal and/or vertical directions) by outputs from the local digital register bank. Hence, operations such as local averaging can be performed, with scope restricted by boundary conditions within PEs in an array. The network consists of two N-type transistors configured to act as resistors, and controlled by means of an analogue bias voltage. A combination of analogue and digital registers can act as sources or sinks to the network.
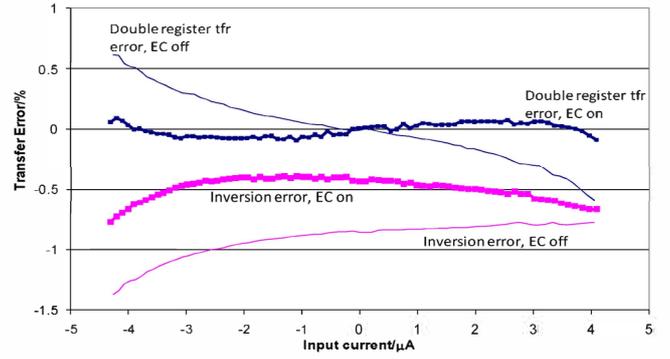
Figure 5. Simulation of register transfer error with current transferred with, and without error correction (EC)

2) Photosensor

In our test chip, two types of photosensor are used. Half the PEs have an N-well photodiode, the others have an N+ in P-subst. diode. All photodiodes use 3.3V diffusions in an effort to improve depletion width depth. The photodiode can be conditionally reset dependent on the local flag state. The fill factor has been marginally increased to 6.5% (from 5.6% for SCAMP3) for the N+/P-subst. diode and is 5.7% for the N-well diode.

3) Squarer

The incorporation of a squarer circuit adds a useful processing function allowing, for example, Euclidean distance calculation. The squarer can also be utilised as a multiplier by determining $(x+y)^2 - x^2 - y^2$ ($= 2xy$). The squarer implemented [15] utilises three PMOS transistors in saturation mode (Figure 6). It can operate on currents sunk to the sub-system from the common analogue bus, and stores the result in the NEWS register (since the system can not both read and write to the bus at the same time). With reference to Figure 6, it outputs a current given by equation 1.

$$I_{sq} = \frac{I_{in}^2}{4I_{Msq}} \qquad (1)$$

Design constraints [16] require that:

$$I_{Msq} \geq \frac{I_{in(max)}}{2} \qquad (2)$$

If I_{Msq} is set at the minimum level, to maximise signal output, $I_{sq(max)} = I_{in(max)}/2$. The bias potentials for the squarer are set globally.
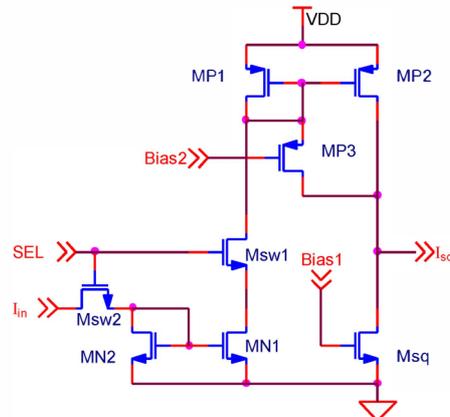
Figure 6. Squarer schematic

Simulation results for the squarer, with an input current loaded from a register, result stored in the NEWS register, then transferred back to the original register are shown in Figure 7.
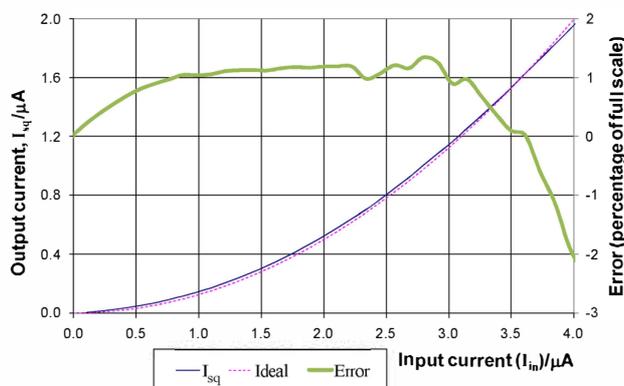

Figure 7. Simulated response of squarer sub-system

III. IMPLEMENTATION

A 1.5x3mm test IC has been fabricated with 20x64 PEs. In the test IC two types of photodiode are investigated, an N-well and N+/P-subst. diodes, split equally across the array. The PE cell measures 32x32μm. The distribution of the functional blocks in the PE layout is shown in Figure 8. A die photograph is shown in Figure 9. As noted earlier, the PE design is routing resource limited, with significant silicon area occupied with routing. Metal 1 and 2 were used for intra-cell routing, metal 3 was partially used for power routing, with the rest of metal 3 and metal 4, 5 used for cell control (ICWs). Metal 6 was used for power supplies. The routing resource requirement of a PE consists of 78 digital and 19 analogue wires.
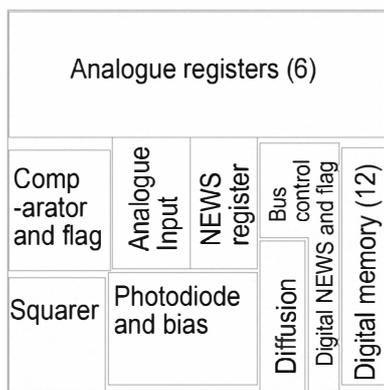

Figure 8. Approximate sub-system floorplan within the PE

IV. CONCLUSIONS

A general purpose processing element has been designed and a test chip fabricated. The PE designed targets achieving a 256x256 vision chip within a 100mm^2 die. It was found that the analogue components within the switched current microprocessors scale only modestly with the use of 0.18μm technology compared to previous work at 0.35μm. Experimental results from the chip are expected to be available for presentation at the conference in May 2011.
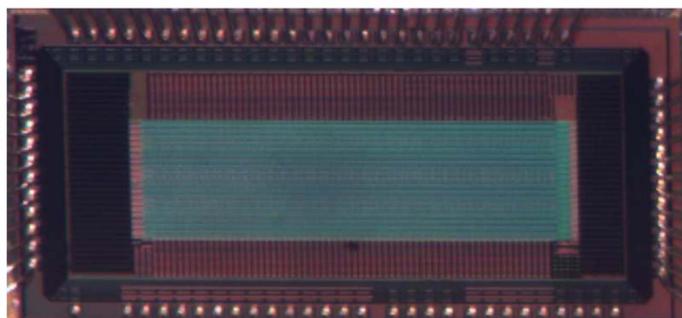

Figure 9. Die photo. The chip contains 20x64 array of PEs

REFERENCES

1. P. Dudek and P. J. Hicks, "A general-purpose processor-per-pixel analog SIMD vision chip," IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, vol. 52, pp. 13-20, 2005.
2. G. L. Cembrano, A. Rodriguez-Vazquez, R. C. Galan, F. Jimenez-Garrido, S. Espejo, and R. Dominguez-Castro, "A 1000 FPS at 128x128 vision processor with 8-bit digitized I/O", IEEE Journal of Solid-State Circuits, vol. 39, pp. 1044-1055, 2004.
3. T. Komuro, I. Ishii, M. Ishikawa, and A. Yoshida, "A digital vision chip specialized for high-speed target tracking," IEEE Transactions on Electron Devices, vol. 50, pp. 191-9, 2003.
4. J. Poikonen, M. Laiho, and A. Paasio, "MIPA4k: A 64x64 cell mixed-mode image processor array," IEEE International Symposium on Circuits and Systems (ISCAS), 2009, pp. 1927-1930.
5. A. Lopich, P.Dudek, "An 80×80 general-purpose digital vision chip in 0.18 μm CMOS technology" IEEE International Symposium on Circuits and Systems (ISCAS), 2010, pp 4257-4260.
6. F. Paillet, D. Mercier, and T. M. Bernard, "Second generation programmable artificial retina," in ASIC/SOC Conference, 1999. Proceedings. Twelfth Annual IEEE International, 1999, pp. 304-309.
7. P. Dudek and S. J. Carey, "General-purpose 128x128 SIMD processor array with integrated image sensor." Electronics Letters vol. 42, no.12, pp. 678-679, 2006.
8. J. E. Eklund, C. Svensson, and A. Astrom, "VLSI implementation of a focal plane image processor-a realization of the near-sensor image processing concept," Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, vol. 4, pp. 322-335, 1996.
9. P. Dudek, A. Lopich, and V. Gruev, "A pixel-parallel cellular processor array in a stacked three-layer 3D silicon-on-insulator technology," European Conference on Circuit Theory and Design (ECCTD), 2009, pp. 193-196.
10. D. R. W. Barr, S. J. Carey, A. Lopich, P. Dudek; (2006). "A Control System for a Cellular Processor Array". 10th International Workshop on Cellular Neural Networks and Their Applications, 2006. CNNA '06.
11. Vilarino, D. L. and P. Dudek. "Evolution of Pixel Level Snakes towards an efficient hardware implementation". IEEE International Symposium on Circuits and Systems (ISCAS), 2007.
12. P. Dudek, "An asynchronous cellular logic network for trigger-wave image processing on fine-grain massively parallel arrays." Circuits and Systems II: Express Briefs, IEEE Transactions on 53(5): 354-358. 2006
13. J. B. Hughes and K. W. Moulding, "S2I: a switched-current technique for high performance," Electronics Letters, vol. 29, pp. 1400-1401, 1993.
14. P. Dudek "A programmable focal-plane analogue processor array" Ph.D. thesis, University of Manchester Institute of Science and Technology (UMIST), May 2000.
15. K. Bult and H. Wallinga, "A class of analog CMOS circuits based on the square-law characteristic of an MOS transistor in saturation," IEEE Journal of Solid-State Circuits, vol. 22, pp. 357-365, 1987.
16. C. Y. Huang, C. Y. Chen, and B. D. Liu, "Current-mode linguistic hedge circuit for adaptive fuzzy logic controllers," Electronics Letters, vol. 31, no. 17, pp. 1517–1518, Aug. 1995.