Mixed Signal SIMD Cellular Processor Array Vision Chip Operating at 30,000 fps

Stephen J. Carey, David R.W. Barr, Bin Wang, Alexey Lopich and Piotr Dudek

School of Electrical and Electronic Engineering
The University of Manchester, Manchester, M13 7PL, United Kingdom
e-mail: p.dudek@manchester.ac.uk

***Abstract* — A prototype vision chip has been designed that incorporates a 20 x 64 array of processing elements on a 31µm pitch. Each processor element includes 14 bits of digital memory in addition to 7 analogue registers. Digital operands include NOR and NOT with operations of diffusion, subtraction, inversion and squaring available in the analogue domain. The cells of the array can be configured as an asynchronous propagation network allowing operations such as flood filling to occur with times of ~1µs across the array. Exploiting this feature allows the chip to recognise the difference between closed and open shapes at 30,000 frames per second. The chip is fabricated in 0.18µm CMOS technology.**

***Index Terms*—Vision chip, smart sensor, asynchronous image processing, cellular processor.**

I. Introduction

Vision chips have the potential to provide improved performance in a variety of machine vision applications. Under different configurations, a vision chip can be made to execute image processing algorithms with very low power consumption or, alternatively, operate at very high speeds. Both of these positive attributes arise from the execution of the algorithm upon the focal plane. Information is transferred between near-neighbour processing elements (PEs), with information extracted from the images communicated between the chip and the system controller. Overall, a reduced dataflow (relative to conventional systems) is accomplished in respect of both low power and high-speed applications. Many vision chips, in common with the chip described here, operate as SIMD computational devices. Such devices have been presented widely [1-8] in both digital, analogue and mixed signal form.

This paper discusses the IC produced from the PE design described earlier [9], and reports tests of the array. In this IC design, each cell of the array incorporates a mixed analogue and digital array of storage and processing elements (in addition to a photodiode) allowing more efficient processing and storage of data dependant on data-type. While the chip utilises analogue processing and memories, each PE is a sampled data system (hybrid analogue/digital) in common with digital processors rather than a continuous time processor. This test IC is intended to be the precursor of a larger array and is controlled by an external instruction issuing device. All instructions issued are executed upon all PEs of the array simultaneously (as per the SIMD processor paradigm) but use local data within the PE (or that of their near neighbours).

The IC described maintains the strategy of the general purpose vision chip, allowing its usage to cover many different applications, while the pitch of the PEs allows an IC with greater resolution at lower cost to be fabricated in comparison to its predecessor SCAMP-3 [1]. The IC utilises switched current memories for analogue storage, with additional storage provided by 3-transistor digital DRAM memories as advanced in ASPA vision chips [8]. The sub-systems that could be incorporated in the PEs of the array were limited by the connectivity that could be provided to the PEs. However, relative to SCAMP-3 with 3 metal layers available, the 6 metal layers of the 0.18µm process has afforded considerable additional features and functionality.

Applications of vision chips rest on the functions of motion, edge detection, tracking and shape detection; these functions require a combination of analogue and digital operations. In the design of this IC, we have shifted the ratio of analogue to digital memory, reducing the number of space-consuming analogue registers while increasing the digital resources available. Furthermore, the digital memories can be used to configure an asynchronous propagation network [10], allowing global operations on data, in addition to performing logical local operations. Also added to this chip is a controllable diffusion network allowing shaped spatial diffusions to occur in a single instruction, and a squarer offering one-quadrant operational capability.

As an example application of this IC, we include a description of a program that distinguishes between open and closed shapes at 30,000fps, demonstrating the utility of focal-plane processing. As earlier applications have also

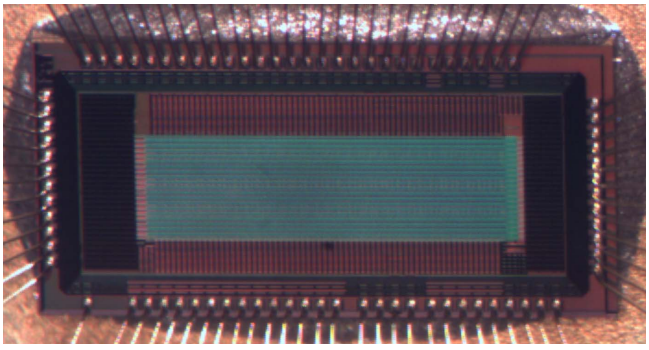

Figure 1. Photograph of the IC (1P6M 0.18µm CMOS process)

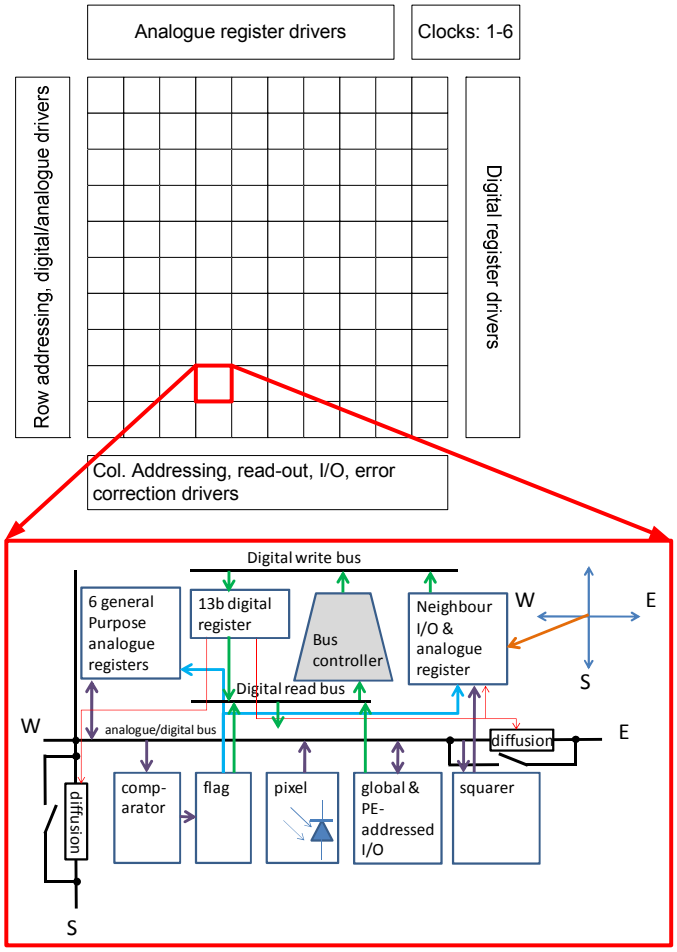

Figure 2. Architecture of the IC

demonstrated [11], frames are not read-out at this speed, only the results of the processing operations are transferred to the program controller.

II. Architecture and Performance

The photograph and basic architecture of the chip are shown in Figures 1 and 2 respectively. The PE has dimensions of 31.26 x 31.26µm and includes 14 bits of memory (13 DRAM; 1 SRAM), 7 analogue registers, a local activity register (flag), an analogue comparator, a current squarer and a photodiode (two types, N-well diode and N+/P-subst, are explored - split equally across the array). The PE divides into 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.

The PE is controlled by means of a 61-bit instruction code word (ICW), with additional control supplied from 12 analogue inputs. The long ICW in turn requires massive PE connectivity with row and column drivers spread across all 4 sides of the IC.

The array is designed to operate at 10MHz with an analogue supply voltage of 1.5V and digital supply voltage of 1.8V. The IC has a total quiescent current of 17µA.

A. Digital Sub-system

The digital system allows the short-term storage of 13 bits

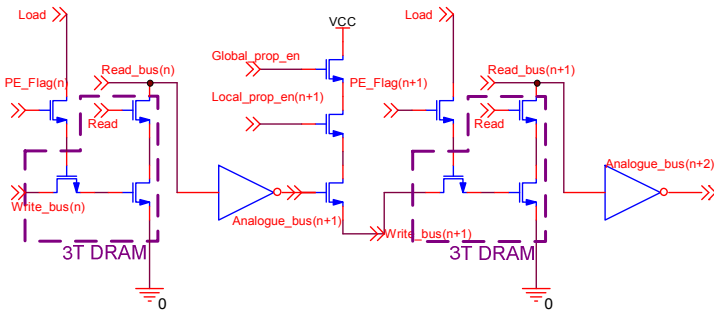

Figure 3. Memories configured as an asynchronous propagation network.

of digital data in 3T dynamic memory. Due to leakage of the write transistor, a refresh is required every 50µs.

The memory cells can be configured as an asynchronous propagation network, shown (in 1-D form for clarity) in Figure 3. This requires that a memory (with flag controlled write) is simultaneously written to and read from. The "read bus" of one cell is set-up (under local and global control) to continuously write to the "write bus" of the corresponding memory cell within the adjacent PE. This creates a 2-D logic chain that reaches across the entire array. Propagation is initiated by marking one or several start pixels. The operation executes the effective instruction: *fill up to where flag = "0"*.

Data propagates from cell to cell within the network in a time of 16ns, or uni-directionally across the entire array in 1µs.

B. Performance Analogue Sub-system

This has much in common with the architecture of SCAMP-3 [1]. 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 a current squarer [12]. The fill factor is 6.5% for the N+/P-subst. diode and is 5.7% for the N-well diode.

1) Analogue Memories

The analogue memories are based on the S^2I cell [13]. The standard cell was modified to include an error correction circuit [14] to allow a reduction in cell size while maintaining low signal-dependent error (signal-independent errors can be compensated for). The purpose is to compensate for the fixed charge imposed by clock feedthrough on the memory

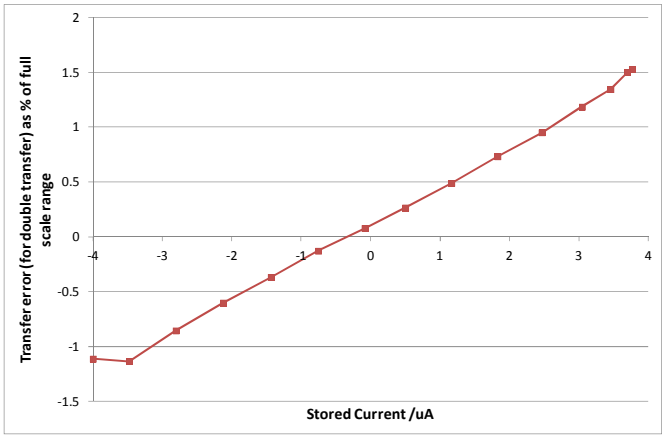

Figure 4. Register transfer error after 2 register transfers

transistors.

On measurement of error after two transfers with the error correction circuit disabled, it was found that the peak-to-peak signal dependent error across the full scale register range (-3.5µA to 3.5µA) was 2.5% (see Figure 4). Results for two transfers are presented since a single transfer is inverting; after two transfers, the result should be identical to that of the input. The simulation from the design phase, with parasitic capacitance included, was for an error of 1.2%. Using the error correction circuit with a drive potential of 1.4V, this error could be shifted by 1.8% of full scale; for a follow-on chip this will clearly only require a minor design change to provide very low register errors. Also, of particular importance for algorithms requiring motion detection, is the degradation of analogue registers over time. This was found to be dependent on stored value as represented in Figure 5.

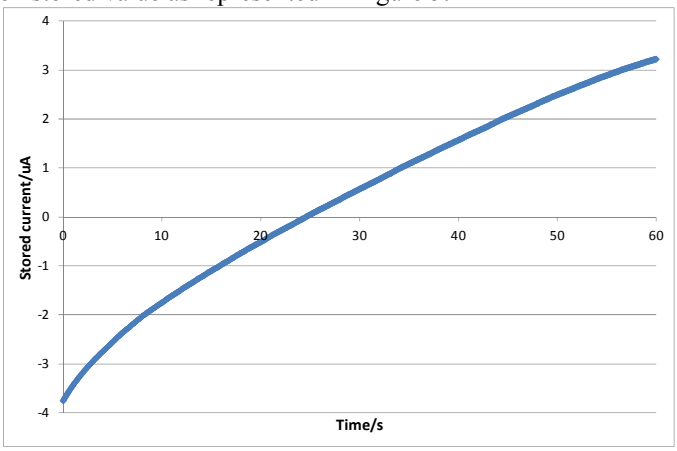

Figure 5. Decay of stored register current over time

Registers decayed at a worst-case rate of 5% of full-scale range per second (cf. 4% over 20ms in [4]). Over a typical frame interval of 50ms, this decay rate has a negligible effect.

2) Diffusion network

This consists of two N-type transistors configured to act as resistors, and controlled by means of an analogue bias voltage. Any combination of analogue registers can then act as sources or sinks to the network (shown in Figure 6). The network can be locally broken (in the horizontal and/or vertical directions) by outputs from the local digital register bank. The diffusion of a letter "U" is shown in Figure 7.

3) Imaging system

Basic imaging operates utilising differential double sampling. The current from the pixel circuit is sampled by register "A" and the pixel then reset. A second register "B" is then written with the sum of the pixel reset level added to

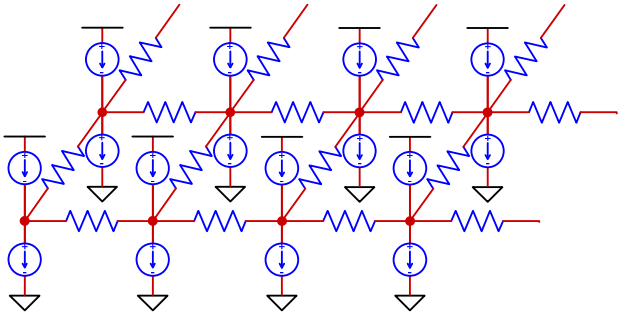

Figure 6. Diffusion network of current sources and sinks

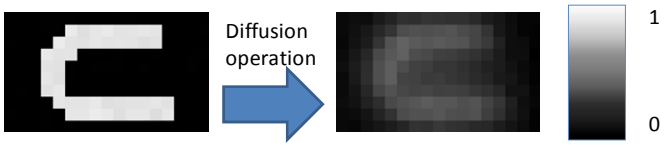

Figure 7. Spatial diffusion of a letter "U" (single clock cycle operation)

inverted register "A"; array register "B" then contains the image. The fixed pattern noise (FPN) of the imaging system (as a percentage of full-scale-range) was 1.6% rms and 1.5% rms for PE's containing N-well photodiodes and N+/p-subst photodiodes respectively. While the N-well photodiode array has slightly higher FPN, this array provides 3x higher sensitivity than the N+/P-subst photodiodes, through improved responsivity and reduced photodiode capacitance.

The light sensors show low dark current operation, to the extent that integration times of upto 1s can be utilised.

4) Squarer

The squarer enables 1-quadrant squaring operations within the PE. For evaluation, an input register was squared and both inputs and outputs recorded. The results for a selection of pixels, along with the ideal response of the system and RMS error are shown in Figure 8. Transistor mismatch is a significant contributor to variations in squarer response.

C. Data I/O and Periphery

This IC provides for global analogue and digital data write and read; PE-specific digital write is also provided, hence allowing PE-specific analogue write via a flagged global analogue write operation.

All four sides of the IC include line drivers to power the PEs and feed ICWs to control the processors. Additionally, flexible row and column decoders allow analogue data to be readout either on a per-pixel basis or summed to provide regional or global analogue outputs. With the significant increase in the ICW width (61 relative to 28 on SCAMP-3), the ICW is now registered on chip and input by means of a 16 bit bus, allowing a significant reduction in external pin-count.

III. Example Application

The completed system combines the presented vision chip, an instruction delivery and control system (synthesised upon a Spartan-3 FPGA) and a communications interface to upload algorithms and receive data that the vision chip is programmed to return. Algorithms are developed and compiled using the

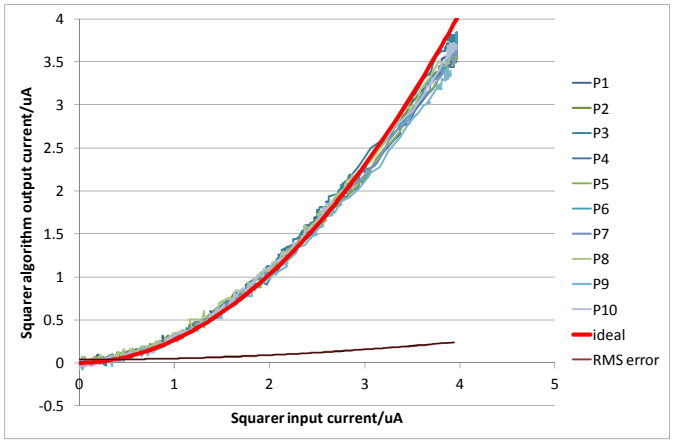

Figure 8. Operation of squarer for 10 pixels (with ideal response for comparison).

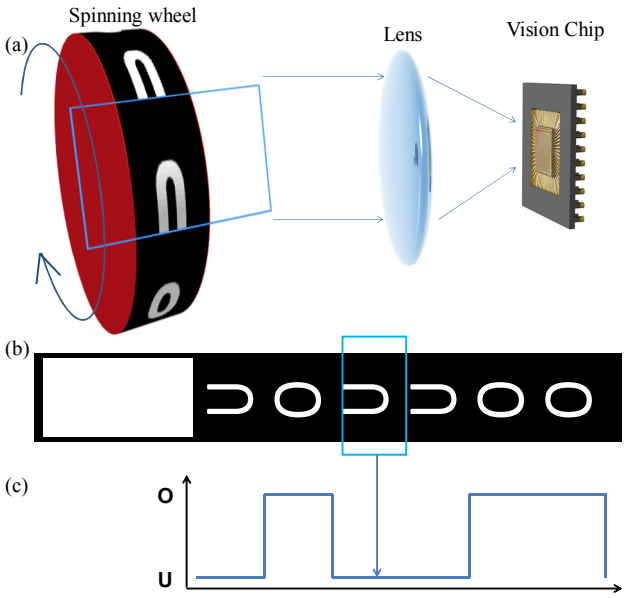

Figure 9. (a) Detection of open and closed shapes from a strip (b) with user output in (c)

APRON compiler [15]. Returned data for analysis and inspection is viewed using a custom viewer program running on a host computer. Unlike many high-speed imaging systems, we wanted to demonstrate that as well as acquiring images at high frame rates, we are also capable of processing them in real-time to determine a non-trivial metric. In this example application we wish to discriminate between open and closed binary shapes, i.e. 'O' versus 'U'. If a closed shape lies within a pre-defined area of the focal-plane, the vision chip system will return 'true' to the host, else 'false'. This occurs at a frame rate of 30,000fps.

The test apparatus (Figure 9a) consists of a wheel spinning at high speed. The circumference of the wheel is labelled with three open and three closed shapes, and an extended "filled in" area which is used as a trigger (Figure 9b). The algorithm consists of two stages. First, as the vision chip starts processing a new series of frames it waits until its entire view is white, then it waits until this is not the case. This transition is used to trigger the second stage of the algorithm which is the discrimination of open and closed shapes. A trigger is necessary to stabilise the visualisation of the returned parametric data (akin to an oscilloscope trace), shown diagrammatically in Figure 9c.

The algorithm for detecting open and closed shapes exploits the asynchronous propagation functions of the vision chip. It consists of the following steps: 1) Capture image; 2) Threshold image so background is '1', and shapes are '0', and load the digital flag register (FL); 3) Repeatedly shift in '1' from all 4 boundaries to isolate a Region of Interest (ROI) containing the spinning wheel; 4) Propagate asynchronously from the ROI boundary, effectively flood-filling up to where FL='0', storing the result in A; 5) Perform array-wide logic operation B = (!A)•FL; 6) Perform global readout, where the output result is the OR operation on all elements of B. If an object is closed, the flood fill propagation does not enter the 'hole', thus when a global "OR" readout is performed the hole still exists and the returned result is true.

The wheel on average spins at 3000rpm; with a radius of 23mm, the shapes move at 7.2m/s. After triggering, the vision chip samples images, at 30,000 fps, returning one bit of resultant data per frame (indicating a presence of a closed shape in a frame). To enhance the visualisation, the user can select a single sampling point at which an entire single frame is returned and thus can visually inspect if the frame contains an open or closed shape, and whether the waveform is in agreement.

IV. Conclusions

A prototype 20x64 vision sensor/processor chip has been tested and successfully applied to an application in high speed imaging operating at 30,000fps. This has demonstrated the utility of asynchronous propagation and data reduction of on-focal plane processing, and provided the post-silicon validation of the new PE design. The follow-up vision chip implementation, scaled-up to a larger array size, will enable high-speed and efficient implementation of image processing algorithms in practical machine vision applications.

References

1. P.Dudek and S.J.Carey, "A General-Purpose 128x128 SIMD Processor Array with Integrated Image Sensor", Electronics Letters, vol.42, no.12, pp.678-679, June 2006
2. W. Zhang, Q. Fu, N. Wu "A Programmable Vision Chip Based on Multiple Levels of Parallel Processors," IEEE Journal of Solid-State Circuits, vol. 46, pp. 2132-2147.
3. T. Komuro, S. Kagami, and M. Ishikawa, "A dynamically reconfigurable SIMD processor for a vision chip," IEEE Journal of Solid-State Circuits, vol. 39, pp. 265-268, 2004.
4. D. Ginhac, J. Dubois, M. Paindavoine, and B. Heyrman, "An SIMD programmable vision chip with high-speed focal plane image processing," EURASIP J. Embedded Syst., 2008, pp. 1-13.
5. J. Fernandez-Berni, G. Carmona, R., and G. Carranza, L., "FLIP-Q: A QCIF Resolution Focal-Plane Array for Low-Power Image Processing," IEEE Journal of Solid-State Circuits, vol. 46, pp. 669-680.
6. A. Rodriguez-Vazquez, R. Dominguez-Castro, F. Jimenez-Garrido, S. Morillas, A. Garcia, C. Utrera, M. Pardo, J. Listan, and R. Romay, ""A CMOS Vision System On-Chip with Multi-Core, Cellular Sensory-Processing Front-End" in Cellular Nanoscale Sensory Wave Computing, eds C. Baatar, W. Porod, T. Roska, pp 129-146. Springer (2010)."
7. 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.
8. A.Lopich and 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, May 2010
9. S.J. Carey, A.Lopich and P.Dudek, "A Processor Element for a Mixed Signal Cellular Processor Array Vision Chip", IEEE International Symposium on Circuits and Systems, ISCAS 2011, Rio de Janeiro, May 2011
10. P.Dudek, "An Asynchronous Cellular Logic Network for Trigger-Wave Image Processing on Fine-Grain Massively Parallel Arrays", IEEE Trans. Circ. and Syst. - II, vol. 53, no.5, pp. 354-358, May 2006
11. A. Zarandy, R. Dominguez-Castro, and S. Espejo, "Ultra-high frame rate focal plane image sensor and processor," IEEE Sensors Journal, vol. 2, pp. 559-565, 2002.
12. 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.
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. D. R. W. Barr and P. Dudek, "APRON: a cellular processor array simulation and hardware design tool," EURASIP J. Adv. Signal Process, vol. 2009, pp. 1-9, 2009.