Locating High Speed Multiple Objects using a SCAMP-5 Vision-Chip

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

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

Abstract - **Presented in this paper is a demonstration system that uses a low-power SCAMP-5 256x256 vision-chip to locate and count multiple objects moving at high speed along arbitrary trajectories. The hardware consists of a SCAMP-5 IC, its power supply system and a Xilinx Spartan3 controller. At 100,000fps, the SCAMP-5 chip can locate and readout the coordinates of a single closed-shaped object amongst clutter. At 25,000fps, the IC can readout the coordinates of 5 objects.**

Index Terms – **Vision Chip, Processor Array, SIMD, Smart Sensors**

I. INTRODUCTION

Systems that are employed to visually locate objects habitually use an imaging array (CMOS or CCD) coupled to high speed memory and a frame processor. The frames may be processed on custom firmware synthesised onto an FPGA or within a microprocessor system. Unfortunately, data rates from the imaging array can become prodigious, placing onerous demands on the connected hardware. For example, a 256x256 imaging array digitising at eight bits per pixel at 100,000fps implies a data rate of 6.5Gbytes.s^{-1}, its associated wide data buses and memory systems.

For simply high-speed event visualisation (rather than processing), focal plane storage and/or massively parallel analogue to digital converters provide solutions for a video stream to be post-processed off-line [1]. This lowers the requirements of the processing system and can achieve frame rates of >1Mfps. To collect, process and act on high speed image data, without frame-count limit, is significantly more challenging. However, use of a vision chip, outputting position information (rather than frames) at 8kHz allowed a pencil to be balanced [2]. And at 10,000fps, the ACE400 vision chip was able to distinguish between 6 different objects [3]. Both these systems reduced the data to that which is salient. A similar approach is adopted in this work.

We have developed a vision chip in 0.18µm CMOS 1P6M technology with 256x256 processing elements (PEs) based upon a recent test chip [4]. Each element incorporates 7 analogue memories and a relative plethora of digital memories (in comparison to our 0.35µm SCAMP-3 vision chip [5]), while reducing the processing element size to less than half its area. A diagram of the functional elements of the system and processor array is shown in figure 1. Within a single processing element, the process of reading or writing these memories to common buses allows sequential, simple computations that can implement complex algorithms. With many PEs operating in parallel in an SIMD (single instruction multiple data) architecture, high speed image processing is made possible.

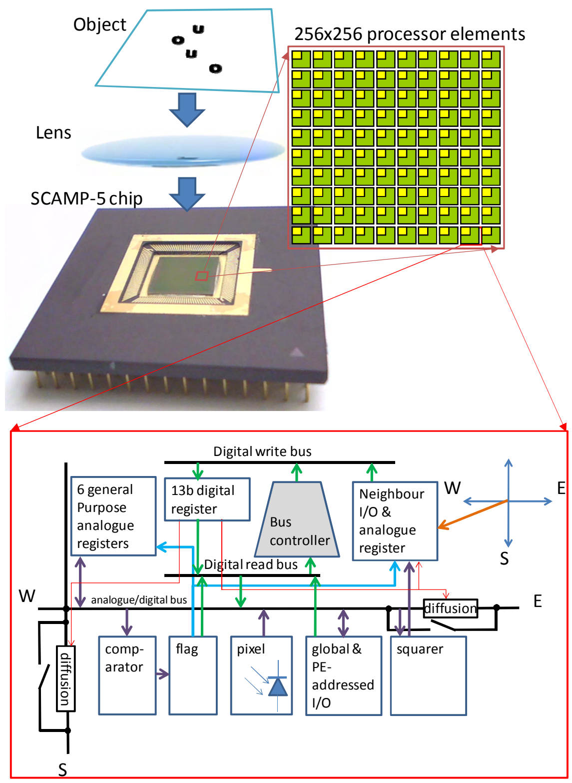

Figure 1. Architecture of a SCAMP-5 vision chip

The particular aspect of the vision chip exploited widely in this work is the use of asynchronous propagation networks to first identify closed shaped objects and then select a single object from those identified. The network is created by linking all PEs of the array through the neighbour communication system and through a digital memory cell in simultaneous read/write mode. A propagation wave will be passed through the PE if a second memory (flag) is true. If the wave never reaches an area of the IC, the pass-through

978-1-4673-0289-0/12/$31.00 ©2012 IEEE

memory cell is always false. In pseudo-code, the instruction executes "Fill from designated pixel(s) upto where Flag = 0"; a single instruction creates the wave that propagates, without further clocking, across the entire IC at 3.2ns per cell.

Present at the periphery of the IC is address extraction logic. This allows the direct readout of coordinates of object bounding boxes or of object pixel coordinates.

II. DEMONSTRATION

The system being demonstrated is a prototype SCAMP-5 vision-chip system. A spinning disc (at 15,000rpm) is adorned with a selection of open and closed shapes (fig 2). The SCAMP-5 is used to locate the position of the closed shapes, and return them to the host computer for visualisation. The system starts acquiring images after a trigger event occurs. The system then captures and processes 4096 frames, returning the coordinates of the closed shapes in each frame to the host computer. The spinning of the disc forces the closed shapes to move in a circular pattern which is visualised as a helix, with the dimensions {x, y, t} where x and y are the coordinates of the shape, and t is the time it was captured. The user can control two aspects of the system. Firstly, the rotational speed of the spinning disc. As the speed increases, the helix 'tightens', and likewise 'relaxes' as the speed is decreased. Secondly, the user can select a time at which a sample frame is returned to visually verify that the coordinate acquisition is accurate.

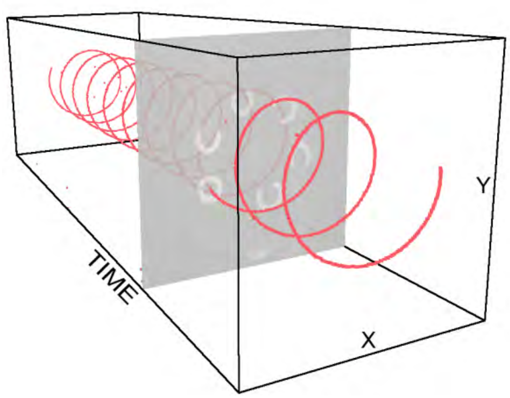

Figure 2. Visualisation of located object (actual output from system)

III. ALGORITHM FOR OBJECT LOCATION

The algorithm makes use of asynchronous propagation and address extraction operations of the SCAMP-5 processor. With reference to figure 3, the image is first captured (Fig 3(a)), and binarised (b). The high frame-rates in this demonstration require additional lighting, as the exposure time is short. Potentially many different operations could be performed to clean up the greyscale image acquired. The image is inverted (c) and loaded to the flag register. Pixels along the edges of the processor array are set, becoming the starting points of an asynchronous propagation into the array. This has the effect of "flood-filling" the image, with images (d), (e) and (f) showing the propagation wave encroaching from the perimeter. Logic operations on images (c) and (f) leaves just the holes inside the closed shapes (g). The coordinates of the left most object are read out (h) and this object filled (i) and removed (j). The coordinate extraction and filling is repeated for the second object (j) and (k). Integration of light for the following frame occurs during the processing of the current frame. With frame capture every 10µs and the disc perimeter travelling at 55ms^{-1}, no motion blur is visible.

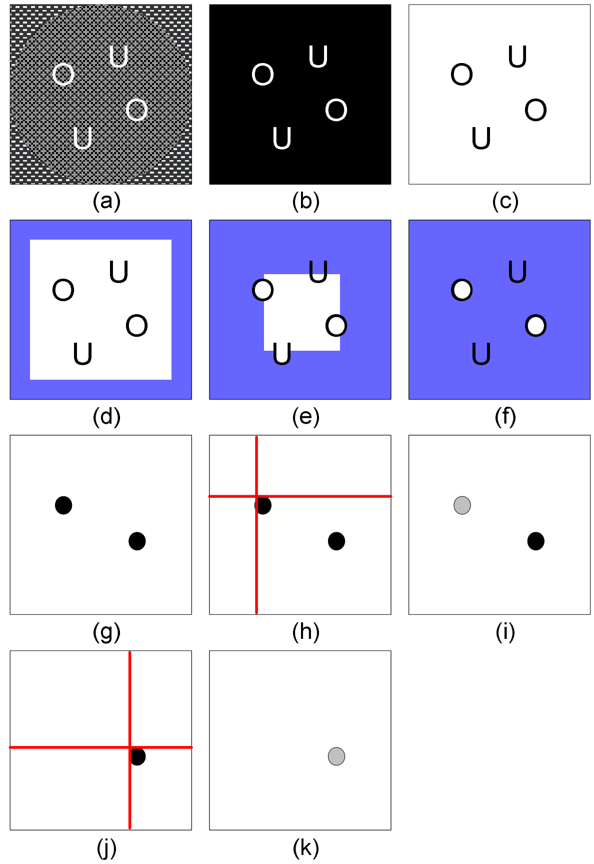

Figure 3. Algorithm progression; locating closed shaped objects.

IV. RESULTS AND DISCUSSION

Extraction of the coordinates of 5 objects can be performed at 25,000fps and that of a single object at 100,000fps. While the SCAMP-5 IC can be used for event visualisation (with very limited frame count), its applications are expected to lie in real-time control systems in which the chip would output an object's location from amongst a cluttered image.

REFERENCES

[1] Takeharu G. Etoh, Kohsei Takehara, Tomoo Okinaka, Yasuhide Takano, Arno Ruckelshausen and Dirk Poggemann, "Development of high-speed video cameras", Proc. SPIE 4183, 36 (2001)

[2] J. Conradt, M. Cook, R. Berner, P. Lichtsteiner, R. J. Douglas, and T. Delbruck, "A pencil balancing robot using a pair of AER dynamic vision sensors," IEEE Int. Symp. On Circuits and Systems (ISCAS), 2009, pp. 781-784

[3] 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.

[4] S. J. Carey, A. Lopich, P. Dudek , "A processor element for a mixed signal cellular processor array vision chip," IEEE Int. Symp. On Circuits and Systems (ISCAS), 2011, pp.1564-1567.

[5] P. Dudek, "Implementation of SIMD vision chip with 128x128 array of analogue processing elements", IEEE Int. Symp. On Circuits and Systems (ISCAS), pp.5806-5809, 2005.