Vision Chips with In-pixel Processors for High-performance Low-power Embedded Vision Systems

Julien N.P. Martel
Institute of Neuroinformatics
University of Zurich & ETH Zurich
jmartel@ini.ethz.ch

Piotr Dudek
School of Electrical and Electronic Engineering
The University of Manchester
p.dudek@manchester.ac.uk

Abstract

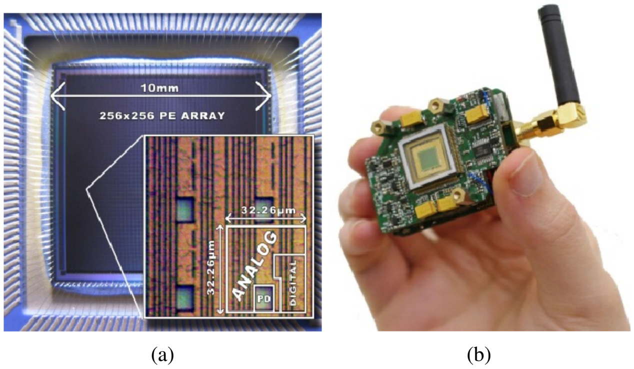
We present the design of vision systems with in-pixel processors, specifically the cellular processor array architecture and implementation of the SCAMP-5 vision-chip. These sensor-processor devices provide a high-speed power-efficient solution to low-level vision processing in embedded systems. We discuss how these devices can be programmed and emulated, including a high-level domain specific language with a compiler responsible of taking care of the peculiarities of theses devices (e.g. analog errors). We discuss application examples, where such vision chips have been used to create systems that provide very low-latencies and running on a low-power budget.

Categories and Subject Descriptors C.1.3 [*Processor Architectures*]: Other Architecture Styles—Analogue computers; D.3.4 [*Programming languages*]: Processors—Compilers, Code generation, Optimization

Keywords vision chip, cellular processor array, embedded vision, SIMD

1. Introduction

The processing requirements of real-time vision on mobile platforms necessitate the development of highly parallel processor architectures and careful balancing of performance and power consumption. It is well known that it is not the processing circuitry but data-transfers (i.e. processor-memory transfers, sensor-processor interface, etc.) that are the bottleneck in terms of achievable performance, and a major contributor to the power dissipation of the system. Our approach to this problem is to eliminate these communications bottlenecks through moving the processing right next to the sensors, and the local memory, in a tightly coupled sensing/processing fine-grain massively parallel system. We have been developing integrated circuits based on these ideas using analogue, digital, and 3D-integration technologies. (Dudek and Carey 2006) (Carey et al. 2013a) (Lopich and Dudek 2011) (Walsh and Dudek 2015) (Dudek et al. 2010)



(a)

(b)

Figure 1: Vision chips: (a) microphotograph of a SCAMP-5 256 × 256 mixed-signal processor array chip, fabricated in 180 nm CMOS technology (b) smart camera system based on SCAMP-3 chip

2. SCAMP-5 vision-chip

The latest SCAMP-5 vision chip developed at The University of Manchester (Carey et al. 2013a) shown in Figure 1 integrates 65,536 processing elements (ALUs + local registers) embedded in a 256 × 256 imager array. The silicon area constraints imposed by such level of integration calls for unconventional circuit solutions, and the device implements a mixed-signal datapath, with arithmetic operations carried out in the analogue domain. Nevertheless, this is achieved without compromising the programmability. The processor array executes code, typically operating on image-wide register arrays with a single (one clock-cycle) instruction. The processor architecture is shown in Figure 2.

The fully software-programmable SIMD architecture delivers 655 GOPS at 1.2 W power consumption (achieved using a 15-year old 180 nm CMOS technology!), making it a powerful front-end for low-power embedded vision systems. Pixel-parallel algorithms are executed on the vision chip, producing a low-bandwidth stream of data (e.g. extracted keypoints, locations of objects of interest, spatio-temporal events etc.) to be further processed by a low-power microprocessor.

The fully-parallel interface allows the transfer of a complete image frame from the image sensor array to the processor array in one clock cycle (100 ns) for an equivalent sensor-processor bandwidth of 655 GB/s. This allows implementation of algorithms with ultra-fast frame rates, that could not be even contemplated on conventional architectures. For example, we demonstrated a high-dynamic-range tone-mapping image acquisition algorithm that processes data from 1,000 images acquired with various exposure set-

ASR-MOV Workshop, CGO'16 March 12th, 2016, Barcelona, Spain

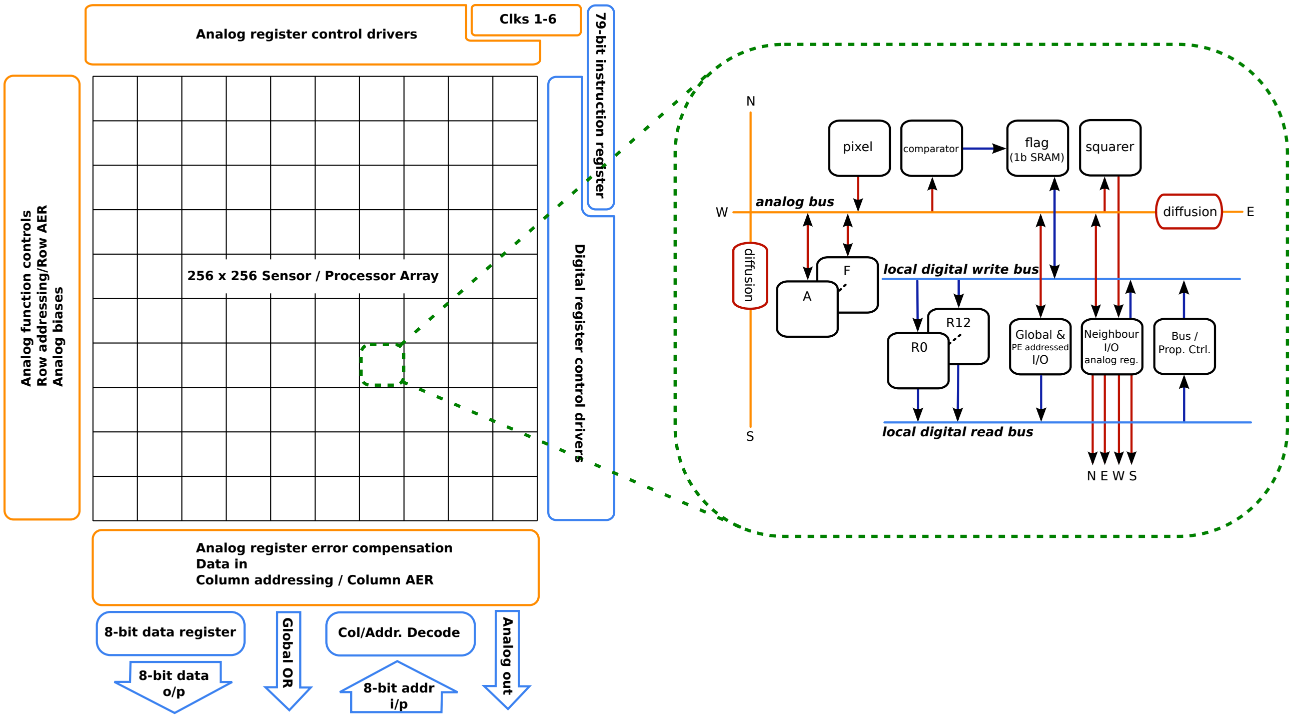


Figure 2: The SCAMP-5 architecture (Carey et al. 2013a). An individual processing element (PE) cell includes a photodetector (pixel), and a processor (ALU, registers, control and I/O circuits). Nearest-neighbour connected PE nodes operate as a pixel-parallel SIMD array processor. An Instruction Processing Unit (IPU) (not shown) dispatches a sequence of 79-bit instruction words specifying the algorithm to be executed. All the PEs execute the same instruction on their own local data. Local activity flags can be used to provide conditional code execution. Arithmetic operations are carried out using analog current-mode circuits, allowing summation, subtraction, division and squaring to operate directly on analog data samples (e.g. gray-level pixel values) without the need for analog to digital conversion. Logic circuits implement binary registers and logic operations. Operations such as flood-fill and low-pass spatial filtering are accelerated in hardware using an asynchronous propagation network formed from interconnecting adjacent PE cells. Readout can be in the form of a binary or analog frames, or global data (e.g. regional summation). In particular, a direct readout of active (foreground) pixel coordinates can be obtained from the asynchronous address extraction circuitry, providing rapid event-based identification of pixels of interest

tings per frame at 20 fps (Martel et al. 2016) as well as an object-tracking algorithm running at 100,000 fps (Carey et al. 2013a).

Conversely, when running at lower frame rates, ultra low-power operation is possible. For instance, we have demonstrated a complete vision system, capable of carrying out image analysis in a loiterer detection application running continuously at 8 fps over 10 days powered by three standard AAA batteries (Carey et al. 2013b).

Further application examples include an inference procedure that can be efficiently carried out jointly on visual quantities such as the optic-flow, the spatial-temporal image gradients, and the ego-motion of the observer(Martel et al. 2015a), an on-chip reconstruction of images from their spatial gradients (Martel et al. 2015b).

3. Toolchain: emulator and compiler tools

We developed a programming toolchain: emulator, and compiler tools that account for the peculiarities of the processors instruction set and operation (e.g. error compensation techniques to improve analogue processing accuracy).These have been designed to work with devices such as SCAMP-5 and similar cellular processor array vision chips .

Emulating the devices in software

Devices can be emulated in software using tools such as APRON (Barr et al. 2009). Specifically for SCAMP-5: The analog and digital operations, the PE communication to shift data on the array as well as the errors in the analog domain are modelled in the emulator. This provides the user a convenient way to prototype and later debug programs by getting a similar behaviour than if the real device was actually used. It also includes several useful features such as stepping through the code, a local variable evaluator, the ability to display all the registers as images etc.

The principle of the emulator is also to use the same language than the one used to generate the instructions for SCAMP-5. Therefore, once the program is written and is emulated in software, the same code can be translated by the compiler for the target architecture and then run in hardware.

A low-level language mapping to the instruction code words

The low-level assembly language used in the emulator is translated in instruction code words interpreted by SCAMP-5. The language has been designed such that each statement in the program can be directly translated in a machine-level instruction code word. The compilation is thus syntax-directed: a single rule in the grammar directly translates in a sequence of actions on-chip.

A peculiarity of this low-level language is that the statements act on all the processors simultaneously following the SIMD paradigm. For instance when writing:

```
R1 = R2 R3
where(R1)
  NEWS = A
  A = EAST
all
```

The first statement performs a digital OR for all the 1-bit registers R1 in all the PEs. The “where” selects all the PEs whose R1 bit is set to 1. Then the analog register A of all the processing elements is copied in the register “NEWS” used to communicate with the neighbours. Finally the neighbour register from the eastern processing element is copied back to A, thus having performed a shift operation in the west direction on the array.

A High-level Domain Specific Language and compiler

With the low-level language, the user has to take care of performing all the operations with the error compensation schemes as well as the register allocation of the variables in use.

To ease the programmability of the device, we developed a high-level domain specific language and its compiler that can either generate code in the low-level language used in the emulator or directly instruction code words for the SCAMP-5 device. The compiler can optimize the number of instructions generated, minimize the analog errors by reordering statements according to algebraic identities specified by the user, reorder instructions to minimize the time a variable spends in an analogue register to prevent the decay of the variable.

To perform these optimizations, we model the problem of allocating a register to a variable, the reordering of statements and the decay of analogue registers in time in a Mixed Integer/Real valued Linear Program (MIRLP) which is solved until optimality. This flexible approach allows us to include more optimizations in our compiler by encoding them appropriately in the MIRLP.

4. Conclusion

Vision chips with pixel-parallel processor arrays provide high-performance at low power consumption, which are desirable features for real-time embedded vision systems. On-sensor processing can reduce the data flow in a system, providing useful information, rather than raw images, to the higher-level processors. The vision chips are fully programmable, however, to take a full advantage of the possibilities offered by these novel devices, suitable software development tools, as well as new algorithms, exploiting the high sensor-processor bandwidth and massively parallel processing capabilities, need to be developed.

References

D.R.W. Barr and P. Dudek, “Apron: A cellular processor array simulation and hardware design tool”, in *EURASIP Journal on Advances in Signal Processing*, pp. 1–9, 2009.

S.J. Carey, A. Lopich, D.R.W. Barr, B. Wang and P. Dudek, “A 100,000 fps vision sensor with embedded 535 GOPS/W 256×256 SIMD processor array”, in *Proc. of the VLSI Circuits Symposium 2013*, pp. C182–C183, June 2013

S.J. Carey, D.R.W. Barr and P. Dudek, “Low Power High-Performance Smart Camera System based on SCAMP Vision Sensor”, in *Journal of Systems Architecture*, Vol 59, Issue 10, Part A, pp. 889–899, November 2013.

P. Dudek and S.J. Carey, “A General-Purpose 128x128 SIMD Processor Array with Integrated Image Sensor”, in *Electronics Letters*, vol.42, no.12, pp.678-679, June 2006

P. Dudek, A. Lopich and V. Gruev, “Vision Sensor with a SIMD Processor Array in a Vertically Stacked 3D Integrated Circuit Technology”, in *Workshop on 3D integration, Design, Automation and Test in Europe, DATE 2010*, Dresden, March 2010.

A. Lopich and P. Dudek, “Asynchronous Cellular Logic Network as a Co-Processor for a General-Purpose Massively Parallel Array”, in *International Journal of Circuit Theory and Applications*, Volume 39, Issue 9, pp. 963-972, September 2011

J.N.P. Martel, M. Chau, P. Dudek and M. Cook, “Toward Joint Approximate Inference of Visual Quantities on Cellular Processor Arrays”, in *Proc. of the IEEE International Symposium on Circuits and Systems, ISCAS 2015*, 2015.

J.N.P. Martel, M. Chau, P. Dudek and M. Cook, “Pixel Interlacing to Trade-off the Resolution of a Cellular Processor Array against more Registers”, in *Proc. of the IEEE International Symposium on Circuits and Systems, ECCTD 2015*, 2015.

J.N.P. Martel, L.K. Muller, S.J. Carey and P. Dudek, “Parallel HDR Tone mapping and Auto-focus on a Cellular Processor Array Vision Chip”, in *Proc. of the IEEE International Symposium on Circuits and Systems, ISCAS 2016*, 2016 *(To Appear)*.

D. Walsh and P. Dudek, “An Event-Driven Massively Parallel Fine-Grained Processor Array”, in *IEEE International Symposium on Circuits and Systems, ISCAS 2015*, Lisbon, pp.1346-1349, June 2015