GDAŃSK UNIVERSITY OF TECHNOLOGY FACULTY OF ETI ANNALS
No 5 Information technologies 2007

Piotr Dudek

School of Electrical and Electronic Engineering
The University of Manchester
PO Box 88, Manchester M60 1QD
United Kingdom
p.dudek@manchester.ac.uk

VISION CHIPS WITH PIXEL-PARALLEL CELLULAR PROCESSOR ARRAYS

Abstract

Vision chips are microelectronic devices which combine image sensing and processing on a single silicon die. In a way somewhat resembling the vertebrate retina these VLSI chips perform preliminary image processing directly on the sensory plane and are capable of very high processing speed at very low power consumption. This makes them particularly suitable for applications such as autonomous robots and other embedded machine vision systems. This paper discusses the principles of using massively parallel fine-grain SIMD processor arrays for low-level image processing and reviews the design and implementation of vision chips developed at the University of Manchester, including the SCAMP-3 chip, which contains 16,384 processors. Application examples and experimental results are presented.

1. INTRODUCTION

Computer vision is a complicated task and a complex set of algorithms has to be executed on image data to achieve the image understanding goal. Computer vision algorithms are usually hierarchically organised, as illustrated in Figure 1. On a high-level, machine perception aims at understanding of visual scenes, objects and their relations, employing symbol-manipulation based algorithms, concepts of artificial intelligence etc. While these algorithms may be quite complex, their computational requirements are often smaller than those of lower-level algorithms. This is because the high-level algorithms operate on a reduced data set - they take as input not a raw pictorial data, but rather an abstract description of its contents. The intermediate-level image processing algorithms are the ones that reduce the pictorial information to a symbolic description of the image contents. These algorithms include, for example: feature extraction and shape detection, object boundary detection, pattern classification, region segmentation and labeling, clustering, etc. Low-level image processing algorithms, dubbed ‘early vision’, are the ones at the front-end of the image processing chain, and are usually local methods, inasmuch as

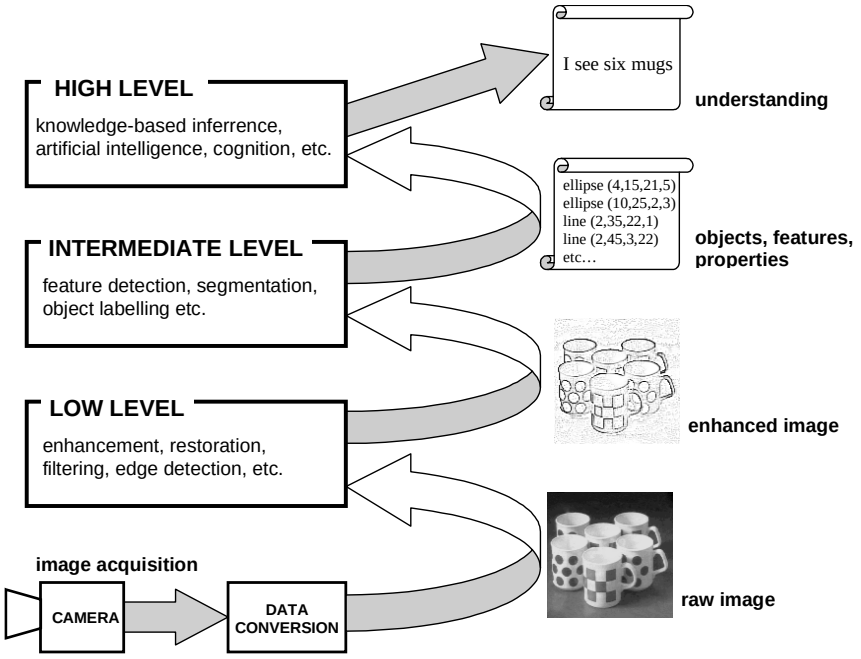

Figure 1. The hierarchical structure of computer vision.

they process raw image data piece-wise, with little or no concern for its content. These algorithms provide initial transformations of the image, such as contrast adjustment, noise removal, restoration, enhancement, edge detection or partial segmentation. Examples of low-level image processing tasks are presented in Figure 2. The image acquisition and data conversion are at the front-end of the computer vision system. They transform the optical image into an electronic data format - usually a quantisation in 2-D space is performed, and an image is represented by an array of elementary image elements called pixels. For grey-scale images, pixel values represent light intensities at the corresponding focal-plane locations.

Given the complexity of the computer vision tasks it is of little surprise that the implementation of real-time vision requires significant computing power, often beyond the capabilities of conventional processors. It is also known, that due to the large amount of image data that needs to be processed on the lower levels of the image processing chain, an efficient implementation of the low-level algorithms is crucial for real-time image processing. This is mostly due to the fact that these simple operations need to be performed on every pixel, over the entire input image, which will typically consist of tens or hundreds of thousands of pixels.

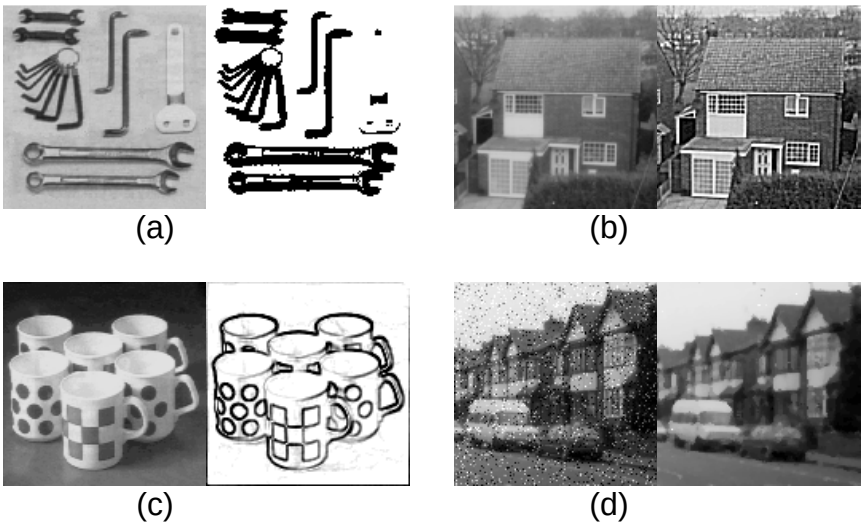

Figure 2. Examples of low-level image processing; left - input images, right - processed images: (a) thresholding, (b) sharpening filter, (c) edge detection, (d) noise removal (median filter).

There are two distinctive features of low-level image processing algorithms, which make them highly parallel in nature. Firstly, identical operations are performed on every pixel in the image. Secondly, the computations are localised, i.e. the output value of the pixel depends usually on the original value of the pixel, and sometimes on the values of its nearest neighbours. For example, brightness and contrast adjustments are achieved by linear modification of a pixel's value. Many algorithms, like linear filtering (e.g. smoothing and sharpening) or edge detection, rely on convolution of the original image with a 3×3 kernel, and therefore the output value of the pixel is computed as the weighted sum of the pixels in a 3×3 neighbourhood centred on this pixel. Similarly non-linear filters, e.g. median filter, are computed by considering the pixel's immediate neighbours. In each case, the same formula is applied to every pixel in the image.

The inherent fine-grain parallelism present in the low-level image processing algorithms can be exploited using parallel computation techniques. Modern superscalar microprocessors include parallel units. However, the degree of parallelism exhibited by these machines is not very large. Moreover, these microprocessors are usually optimised for floating-point operations, and are not particularly suited to perform most low-level vision tasks. Digital Signal Pprocessors (DSP) that include multiple execution units, can be used in some less demanding cases. Supercomputers, especially massively parallel machines (i.e. the ones that contain a very large number of processing nodes) can deliver the required performance. However, in most computer vision applications, other factors, most notably the cost of a system, are of great importance. Similarly, especially in portable and mobile applications, size and power dissipation must be taken into account.

In order to achieve a high-performance image-processing system, several solutions have been proposed that employ a simple model of parallel computation known as Single Instruction Multiple Data (SIMD). Multiple processors (or processing elements, PEs) are organized in an array and all perform the same operation in a given time, thus greatly simplifying the control path of a parallel system. Simple PEs, which use integer or even one-bit arithmetic units, allow a large number of processing nodes to be integrated onto a single chip, and therefore enable shrinking of the size of a massively parallel computer to a single-chip level. An alternative way to achieve a very high-performance low-cost image-processing system is to use an application-specific solution. These systems usually offer direct hardware implementation of specific vision algorithms, using advanced parallel architectures. Another performance boost, coupled with cost-saving, can be achieved by employing analogue techniques. Analogue signal processing circuits can offer significant advantages in terms of speed, power dissipation and silicon area over their digital counterparts. The drawback of most of these architectures is, however, that they usually offer little flexibility as compared with software-programmable general-purpose digital processors. Yet another approach is to place processing elements next to image sensors, combining the image acquisition and parallel processing in a single solid-state integrated circuit. These chips are sometimes called 'smart sensors', 'artificial retinas', 'focal plane image processors' or 'vision chips'. Images are processed in the focal-plane, which eliminates the I/O bottleneck between the image sensor and the processor. Special-purpose analogue vision chips offer efficient and fast implementation of early vision tasks. The most challenging task when considering a versatile vision chip is how to ensure programmability given small area available for processing elements in a pixel-per-processor array. The designs presented in this paper fit in the gap between fine-grain massively parallel SIMD digital computers, and smart sensor chips, offering the flexibility of the general-purpose SIMD processor arrays and speed, low-cost and power efficiency of the pixel-parallel 'vision chip' approach.

2. PIXEL-PARALLEL VISION CHIPS

The basic concept of the pixel-parallel vision chip is illustrated in Figure 3. The chip combines the functionality of a camera, and image processor, in a single integrated circuit. The incoming light is focussed onto the chip surface, and the images are captured using photosensors. The processors, which are placed within the sensor array, execute then a software program, operating concurrently, each processor operating on data belonging to respective pixel in the image. The entire architecture can be represented as consisting of multiple layers of data arrays (Figure 4), and operations are performed on these arrays, on a pixel-wise basis. Shift operations are also available, that correspond to nearest neighbour communication, while local autonomy of processing elements enables executing data-dependent conditional operations.

The integration of a large number of processors on a single chip is a challenging task, requiring optimisation of the processor area. We have proposed an approach based on a sampled-data "analogue microprocessor" [1] to achieve this goal, and demonstrated several generations of the 'SCAMP' integrated circuit implementations [2-4]. Unlike conventional digital computers which use Boolean logic (operating with values of '0' and '1' only), the

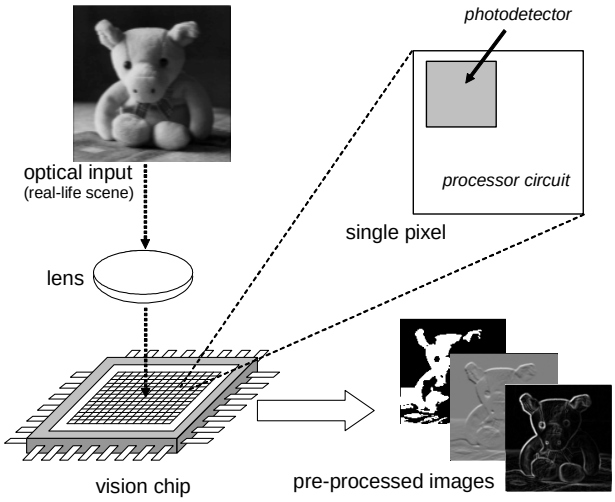

Figure 3. A vision chip with a processor-per-pixel array

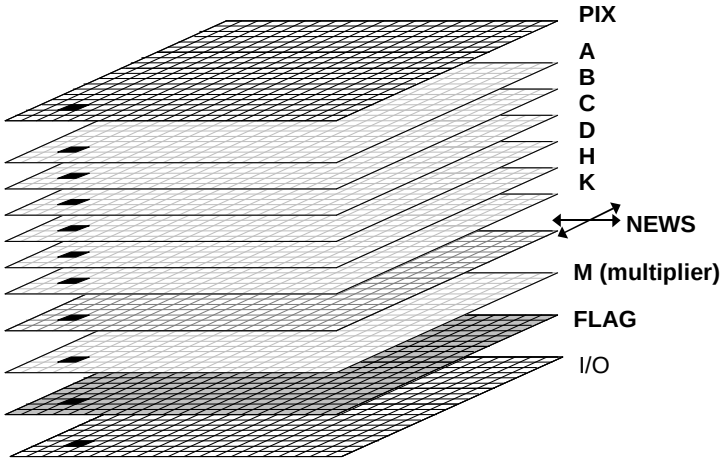

Figure 4. Register-array layers in a pixel-parallel architecture (single APE is marked)

operation of the PE on the SCAMP chip is analogue, based on storage of charge, current summation, etc. Nevertheless, from the programmer's point of view, the system behaves like a parallel SIMD computer. The analogue operation, however, results in a very compact and power efficient implementation of the processor.

The SCAMP-3 chip [4] (shown in Figure 5) is fabricated in a 0.35um CMOS technology. It contains 128×128 pixel-processors and comprises over 1.8 million transistors (most of which are working in analogue mode). The main parameters of the chip have been gathered in Table I. An interesting aspect of working with the analogue processors is the limited

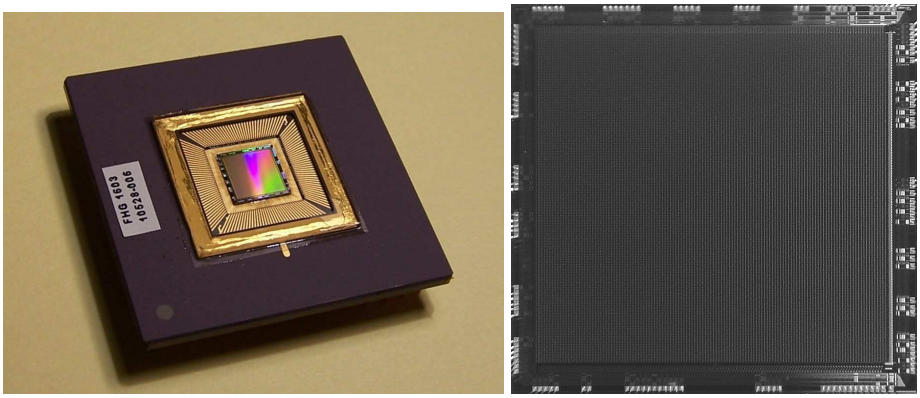

Figure 5. The SCAMP-3 chip in a package (left) and the silicon die microphotograph (right).

Table I. SCAMP-3 Chip Specifications

Number of parallel processors:	16,384
Performance:	20 GOPS
Power consumption:	240 mW max
Supply voltage:	3.3V and 2.5V (analogue)
Image resolution:	128x128
Sensor technology:	photodiode, active pixel, 0.35µm CMOS
Pixel size:	49.35µm×49.35µm (includes processor)
Pixel fill-factor:	5.6%.
Fixed pattern noise:	1%
Accuracy of analogue processors (instruction error):	
Storage linearity:	0.52%
Storage error fixed pattern noise:	0.05% rms
Division-by-2 fixed pattern noise:	0.12% rms
Random noise:	0.52% rms
Image processing performance benchmarks (execution time):	
Sharpening filer 3x3 convolution:	17 µs
Sobel edge detection:	30 µs
3x3 median filter:	157 µs

accuracy and noise of the storage, transfer and arithmetic operations. Effective solutions, both regarding the circuit design and the algorithmic approaches, have been proposed to keep the errors at a level acceptable for the intended applications [5]

To enable evaluation of the vision chip in various applications, entire “smart camera” system, including control, interfacing, and relevant software has been developed [6]. The system has been applied to execute some basic image processing algorithms as well as

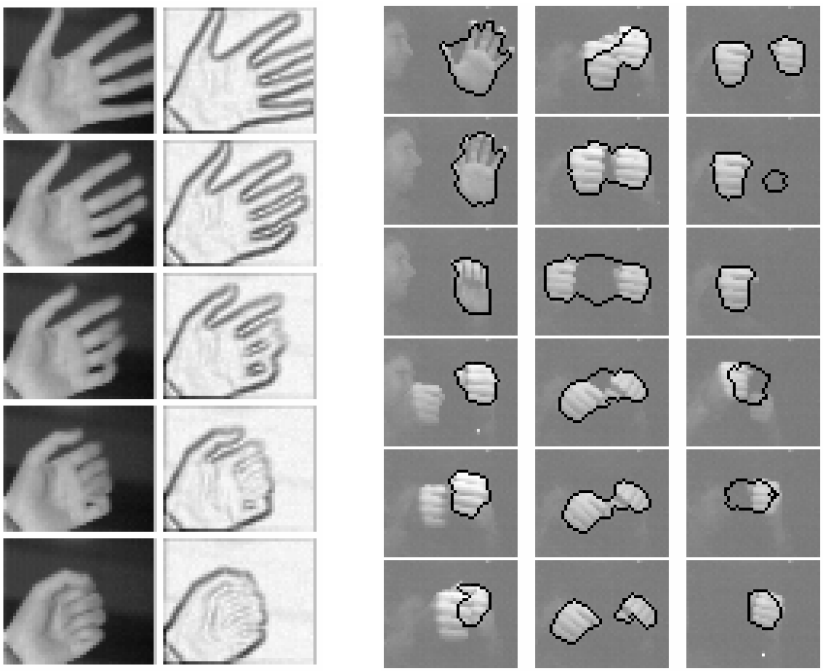

Figure 6. Real-time image processing on the SCAMP vision chip; sequences illustrating edge detection (left) and active contours (right)

more complex tasks, for example image segmentation using so-called Pixel Level Snakes [7], which are an active contour technique applicable to pixel-parallel processor arrays (Figure 6). The system has been also used in demonstrations involving mobile robots, in applications such as target tracking, and autonomous navigation. In addition to image processing, the massively parallel array is suitable for other applications, and we demonstrated an implementation of a complex neural network model on the SCAMP-3 chip [8].

While the approach based on the analogue microprocessor have proven successful, recently we have also investigated architectures based on digital asynchronous/synchronous processor cells [9]. On one hand, this is due to the fact that certain operations can be very efficiently executed in a continuous-time system, using simple, binary logic structure of a cellular processor network [10], giving enormous benefit in terms of performance and efficiency as compared with the SIMD array solution. On the other hand, the exploration of digital architectures is motivated by the continuous scaling of the CMOS technology, and the fact that the performance of analogue cells, in particular repeatability of the circuit element parameters degrades as the transistor dimensions reduce (which leads to increased errors of analogue processing), while the purely digital solution can take full advantage of increased speed and reduced power consumption offered by technology scaling.

3. CONCLUSIONS

Pixel-parallel vision chips, with thousands of processors, can be feasibly produced in today's low-cost silicon technologies. The technology scaling will enable even greater performance and integration of the system functionality. The vision chips are particularly suited to embedded computer vision and image processing applications, such as autonomous robots, surveillance, biomedical applications (e.g. augmented vision, artificial retina), etc. where size and power consumption are important considerations in addition to demands for high computational performance and low cost. Our research continues, exploiting architectures, circuitry and algorithms suitable for future generations of these devices.

REFERENCES

[1] P.Dudek and P.J.Hicks, "A CMOS General-Purpose Sampled-Data Analogue Processing Element", IEEE Transactions on Circuits and Systems - II: Analog and Digital Signal Processing, vol. 47, no. 5, pp. 467-473, May 2000

[2] P.Dudek and P.J.Hicks, "A General-Purpose Processor-per-Pixel Analog SIMD Vision Chip", IEEE Transactions on Circuits and Systems - I, vol. 52, no. 1, pp. 13-20, January 2005

[3] P.Dudek, "A 39x48 General-Purpose Focal-Plane Processor Array Integrated Circuit", IEEE International Symposium on Circuits and Systems, ISCAS 2004, Vancouver, vol.V, pp.449-452, May 2004

[4] 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

[5] P.Dudek, "Accuracy and Efficiency of Grey-level Image Filtering on VLSI Cellular Processor Arrays", IEEE Workshop on Cellular Neural Networks and their Applications, CNNA 2004, Budapest, pp.123-128, Budapest, July 2004

[6] D.R.W.Barr, S.J.Carey, A.Lopich and P.Dudek, "A Control System for a Cellular Processor Array", IEEE International Workshop on Cellular Neural Networks and their Applications, CNNA 2006, pp.176-181, Istanbul, August 2006

[7] P.Dudek and D.L.Vilarino, "A Cellular Active Contours Algorithm Based on Region Evolution", IEEE International Workshop on Cellular Neural Networks and their Applications, CNNA 2006, pp.269-274, Istanbul, August 2006

[8] D.R.W.Barr, P.Dudek,J.Chambers and K.Gurney (accepted) IEEE International Joint Conference on Neural Networks, IJCNN 2007.

[9] A.Lopich and P.Dudek, "Architecture of a VLSI cellular processor array for synchronous/asynchronous image processing", IEEE International Symposium on Circuits and Systems, ISCAS 2006, pp.3618-3621, May 2006

[10] P.Dudek, "An Asynchronous Cellular Logic Network for Trigger-Wave Image Processing on Fine-Grain Massively Parallel Arrays", IEEE Transactions on Circuits and Systems - II, vol. 53, no.5, pp. 354-358, May 2006