CMOS Image Sensor versus Retina Experience

A. Elouardi, S. Bouaziz, A. Dupret, L. Lacassagne, J.O. Klein, R. Reynaud
Institut d'Électronique Fondamentale, Orsay, France, elouardi@ief.u-psud.fr

Abstract
This paper presents a comparison relating two different architectures dedicated for a vision system on chip. The first one implements a logarithmic CMOS imager and a microprocessor. The second involves the same processor with a CMOS retina which implements hardware operators and analog microprocessors. We have modeled two vision systems. The comparison is related to image processing speed, processing reliability, programmability, precision, bandwidth and subsequent stages of computations.

Keywords
Vision System Architecture, System on Chip, Retinas, CMOS/APS Sensors.

I- INTRODUCTION

To face the computational complexity induced by the computer vision algorithms, an alternative approach consists to perform some image processing on the sensor focal plane. The integration of pixels array and image processing circuits on a single monolithic chip makes the system more compact and allows enhancing the behavior and the response of the sensor. To achieve some simple low-level image processing tasks (early-vision), a silicon retina integrates analogue and/or digital processing circuits in the image-sensing element [1] or at the edge of the image sensor array [2]. The energy dissipation is also lower than with classical approach using multi chip (microprocessor, sensor, logic glue …etc).
This paper is built to get a general conclusion on the aptitude of the retinas to become potential candidates for systems on chip, consequently to reach an algorithm-architecture and system adequacy. Hence this paper focuses on the VLSI compatibility of retinas, more particularly, of integrating image processing algorithms and their processors on the same sensor focal plane to provide a smart vision system on chip.

II- VISION SYSTEM BASED ON A CMOS SENSOR

In recent years CMOS image sensors have started to attract the attention in the field of electronic imaging that was previously dominated by charge-coupled devices (CCD). The reason is not only related to economic considerations but also to the potential of realizing devices with imaging capabilities not achievable with CCDs. For applications where the scene light intensity varies over a wide range, dynamic range is a characteristic that makes CMOS image sensors attractive in comparison with CCDs. An example is a typical scene encountered in an outdoor environment where the light intensity varies over a wide range, as, for example, six decades. Image sensors with logarithmic response offer a solution in such situations. Since the sensor is a non-integrating sensor there is no control of the integration time. Because of the large logarithmic response the sensor can deal with images with large contrast without the need for iris control, simplifying the system vision. This makes the sensors very well suited for outdoor applications. Due to the random access, regions of interest can to be read-out and processed. This reduces the image processing, resulting in faster and/or cheaper image processing systems.
We have modeled a vision system based on a logarithmic CMOS sensor (FUGA1000) and an ARM microprocessor. The CMOS sensor is a random addressable 1024x1024 pixels. It has a logarithmic light power to signal conversion. This monolithic digital camera chip has an on-chip 10 bit flash ADC and digital gain/offset control. It behaves like a one Mbyte memory. The entire architecture is shown in figure 1. The figure 2 gives an overview of the CMOS sensor and the experimental module.

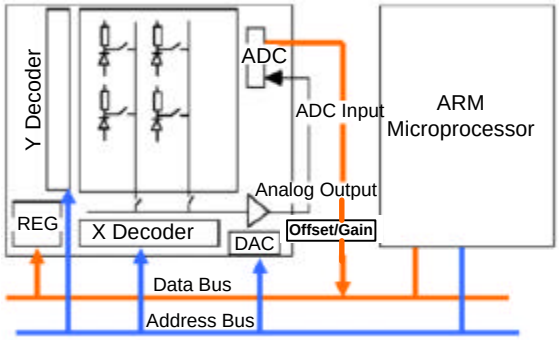


Figure 1. System architecture based on the CMOS sensor

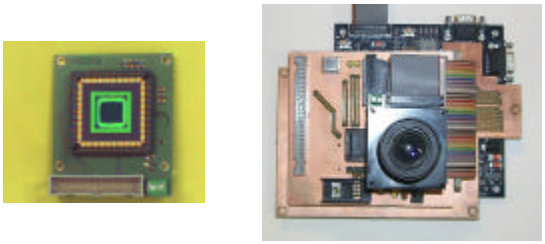


Figure 2. Overview of the CMOS sensor (1024x1024 pixels) and the first experimental module

The major drawback of the logarithmic sensor is the presence of a time-invariant noise in the images. The Fixed Pattern Noise is caused by the non-uniformity of the sensor characteristics. In particular, threshold voltage variations introduce a voltage-offset of each pixel. The FPN noise is

removed from the images by adding to each pixel value the corresponding offset (v → v + offset, where v is the raw pixel value and offset is the FPN correction corresponding to the pixel). For the CMOS sensor, the FPN suppression is performed by the ARM microprocessor (this operation can be achieved by an FPGA circuit for example) in real time and it is transparent. The sensor is shipped with one default correction frame (figure 3).

Figure 3. Result of an FPN correction

III- VISION SYSTEM BASED ON A CMOS RETINA

1- PARIS architecture

PARIS (Programmable Analogue Retina-Like Image Sensor) is an architecture for which the concept of retinas is respected by integrating, in the same circuit, the acquisition photo-elements and the column-level processing operators [3]. The architecture is showed in figure 4. It guarantees a high degree of parallelism and a balanced compromise between the communications and the computations. However in this architecture a pixels array is associated to a mixed analogue-digital processors vector. Since the approach consists in setting operators at the array edge [4], each processor is associated to a column and is able to carry out a wide range of low-level image processing algorithms [5]. Consequently, these functions are shared by a group of pixels, and the image processing is then carried out sequentially. Since images of the array are read row by row, the whole row is transferred serially to the output pixel by pixel. With this typical readout mechanism of CMOS image sensor array, the column processing offers the advantages of parallel processing that permits low frequency processing and thus low power consumption. Finally, the resulting low-level information provided by the retina can be then further processed by some external processor. The pixel features a high fill factor because the PE is taken out to the column, which increases the photosensitivity of the sensor. This approach also eliminate the input output bottleneck between different circuits although there is a restriction on the implementation area, particularly column width, the implementation is relatively flexible because of the freedom in vertical direction of the columns. Still, due to the narrow column width, linked to the pixel size, designers cannot have full flexibility of processing operators' area. Detailing this architecture, we distinguish two blocks: an array of pixels and an analogue/digital processors vector. Pixels can be randomly accessed. Each pixel consists of a photo-sensor and four analogue capacitors acting as memories. The vector of processors operates in an analogue/digital mixed mode. Figure 5 gives details of the analogue processor.

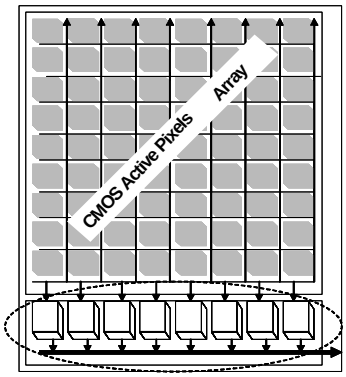
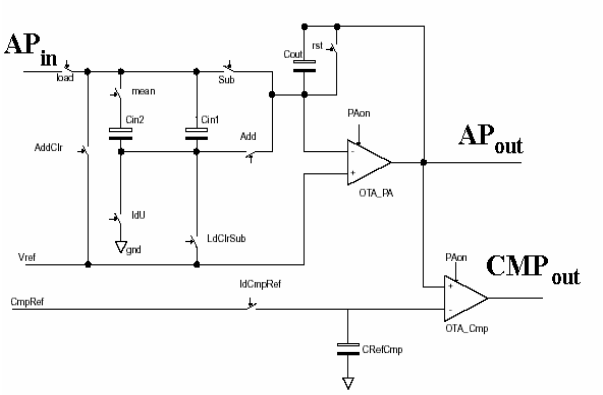


Figure 4. PARIS architecture



APout: Analogue Processor output, **CMPout**: Comparator output, **APin**: Analogue Processor input

Figure 5. Analog processor unit architecture

Combination of divisions, additions and subtractions executes the MAC (Multiply-and-Accumulate) required in an operation like spatial filtering. The analogue processor unit implements three switched integrating capacitors, one OTA (Operational Transconductance Amplifier) and a set of switches controlled by a digital microprocessor. The capacitor Cout is used as an accumulator. Thanks to the OTA, the charge stored in capacitor Cin1 is transferred towards Cout. Example of detailed operations can be found in [5]. In addition, with the comparator implemented in the APU, it is also possible to use this structure as an analogue to digital converter. PARIS1 is a 16x16 pixels VLSI prototype with 16 analogue processors though it is designed to support up to 256x256 pixels. However, this first circuit allows validating the integrated operators through some image processing algorithms like edge and movement detection.

2- System architecture

The computation capacities of microprocessors have increased; it is possible to perform pixel processing "on the fly" as the pixel values are scanned out of the retina and so

a full frame buffer is not necessary. Since microprocessors have asset of high integration, high computing power and low consumption, these characteristics make them suited for the CMOS/APS imager sensors or smart retinas (known as intelligent sensors) as a finite state machine (FSM) giving instruction to an SMD device. Such microprocessors support various operating systems and communication drivers. This suggests that it should be possible to associate a CMOS Retina with a low cost microprocessor or a microcontroller to implement a vision system on chip.

To evaluate this architecture, we have implemented a prototype based on this idea. It is a three design parts. The first two chips are the smart retina and a microcontroller. The third part is a simple interface card implementing DAC/ADC converter (that can be integrated on the microcontroller) and decoders' circuits. The microcontroller is built around a CPU core: the 16/32-bit ARM7TDMI RISC processor. It is a low-power, general purpose microprocessor, operating at 50 MHz, that was developed for custom integrated circuits. The retina, used as a standard peripheral of the microcontroller, is dedicated to image acquisition and low-level image processing. The processor waits for the extracted low-level information and processes them to give high-level information. The system sends then sequences of entire raw images.

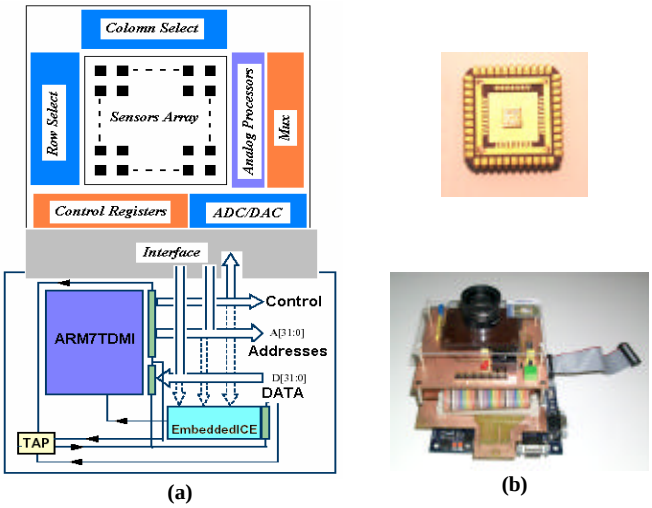


Figure 6. (a) System based on PARIS1 retina, (b) Overview of the retina and the second experimental module

With all components listed above, we obtain a system vision that uses a fully programmable smart retina. Thanks to the analogue processing units, this retina extracts the low-level information (e.g. edges detection). Hence, the system, supported by the microprocessor, becomes more compact and can achieve processing suitable for real time applications. The advantage of this architecture type remains in the parallel execution of a consequent number of low level operations in the array by integrating operators shared by groups of pixels. This allows saving expensive resources of computation, and decreasing the energy consumption. In term of computing power, this structure is more advantageous than that based on a CCD sensor associated to a microprocessor [6]. Figure 6 shows the global architecture of the system and an overview of the experimental module implemented for test and measurements. We have successfully implemented and tested a number of algorithms, including convolution, linear filtering, edge detection, segmentation, motion detection and estimation. Some examples are presented below. Images are processed at different values of luminosity using an exposure time self calibration algorithm. Figure 7 gives examples of processed images.

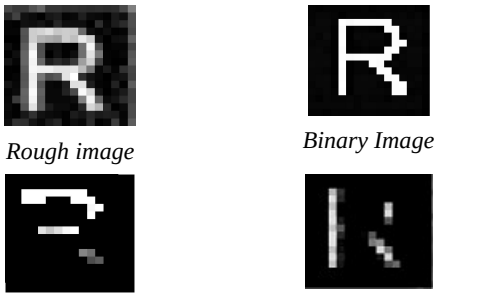


Figure 7. Examples of processed images

IV- COMPARISON

A vision system consists to proceed in three principal tasks: image acquisition, low-level data processing (signal processing) and high-level processing (scene recognition, artificial intelligence). Therefore, the information provided is used to enable the system to interact with the environment (active system) or to prevent a danger (passive system). Some systems use a hardware architecture allowing a complete analysis of the scene. This increases the iteration number and the complexity of computation. Consequently, it reduces the processing speed. Other systems are based on a hardware architecture which privileges the computing speed at the images quality. In those applications, the processing speed of such a vision system is a major factor to achieve applications in real time. The aim is to compare the vision system implementing the logarithmic CMOS imager (FUGA1000) and the ARM microprocessor with the one based on PARIS1 retina. This comparison is related to image processing speed, programmability and subsequent stages of computations. We have used the edge detection algorithm and a Sobel filter algorithm to take several measurements of the computation times relating to the two architectures described bellow. For the retina based system, these computations are carried out by the analogue processors integrated on chip. For the CMOS sensor based system, these computations are carried out by the ARM microprocessor. The two computation time graphics presented in the figure 8 translate the diverse computing times for different square image resolutions for both systems. It is significant to note that the acquisition time of the frames is not included in these measurements. The comparison is related

to the data processing computing time. Times relating to the PARIS retina were obtained by extension of the data processing timing obtained from those of the first prototype [5].

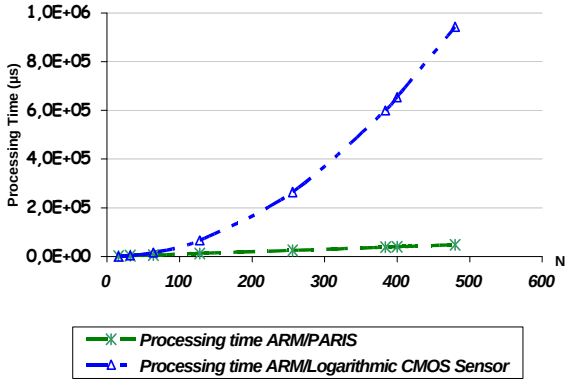


Figure 8. Processing times for the two systems

We deduce that the computation time for the CMOS sensor/ARM processor like-system varies according to the pixels square number N^2 (quadratic form). Hence, the computation time for Retina-like system varies according to the number of line N (linear form) thanks to the analogue processor vector. Consequently, the microprocessor of the CMOS sensor like-system carries out a uniform CPP (Cycle Per Pixel) relative to regular image processing independently of the number of proceeded pixels. For PARIS lik-system, the CPP factor is inversely proportional to the number of lines N. Figure 9 shows the evolution of the CPP for PARIS1 and CMOS sensor/ARM systems.

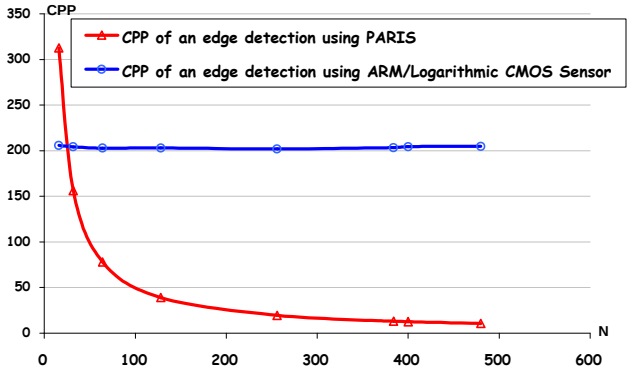


Figure 9. Evolution of the CPP for the two systems

A characterization of the power consumption for PARIS1 based system has been achieved [5]. The total power of an NxN resolution and N analogue processing units is:

$$P = a.N^2 + ß.N \ (\mu W)$$

When a (100 µW) is the power consumption per pixel and ß (300 µW) is the power consumption per analog processing unit. The circuit has a consumption of 30.4 mW. The consumption of the corresponding system (CMOS sensor/ARM microprocessor) is 253 mW (5V operation, 10MHz, RAM, ROM and logic glue consumption are excluded). Hence, When comparing the power consumption between the CMOS sensor/ARM like-system and the PARIS retina at 10MHz frequency, we conclude that the on chip solution allows better performances and low power consumption. This comparison does not take into account the power consumption of the ARM processor peripherals. In addition, 10MHz clock of the ARM processor ARM is not a realist comparison related to the necessary computing power, which needs a greater clock frequency. It's very difficult to calculate the power dissipation of systems including multiple peripherals, logic glues, user custom circuits and microcontroller including processor core. IC manufacturers hide technical information relating to the power consumption.

V- CONCLUSION

The based methodology leads to a general conclusion, that of the ability of retinas to become potential candidates to design high performance vision systems with high resolution, to provide a low-level information and consequently to reach an algorithm-architecture-system adequacy (A3 methodology). In this context, the chosen application makes it possible to build up a conclusion on an integration of an on chip retina based system.

VI- REFERENCES

[1] Dudek P., Hicks J., A CMOS General-Purpose Sampled-Data Analogueue Microprocessor, Pro. of the 2000 IEEE International Symposium on Circuits and Systems. Geneva, Suisse.

[2] Ni Y., Guan J.H., "A 256x256-pixel Smart CMOS Image Sensor for Line based Stereo Vision Applications", IEEE, J. of Solid State Circuits, Vol. 35 No. 7, Juillet 2000, pp. 1055-1061.

[3] Nshare A., Klein J.O., Dupret A., "Improved ARAM for PARIS, an original vision chip", CNNA2002, Frankfurt/Main, World Publishing 2002, pp 347-354.

[4] Dupret A., Klein J.O., Nshare A., A programmable vision chip for CNN based algorithms. CNNA 2000, Catania, Italy: IEEE 00TH8509.

[5] Dupret A., Klein J.O., Nshare A., "A DSP-like Analogue Processing Unit for Smart Image Sensors", International Journal of Circuit Theory and Applications 2002. 30: p. 595-609.

[6] Dave Litwiller, “CCD vs. CMOS: Facts and Fiction”. The January 2001 issue of PHOTONICS SPECTRA, Laurin Publishing Co. Inc.

[7] Elouardi A, Bouaziz S., Dupret A., Klein J.O, Reynaud R., “On Chip Vision System Architecture Using a CMOS Retina“. Proceeding of IEEE IV’04. Pages 206-211. ISBN 0-7803-8311-7. June 14-17, 2004. Parma, Italy.