

A Smart Sensor for Image Processing: Towards a System on Chip

A. Elouardi, S. Bouaziz, A. Dupret, L. Lacassagne, J.O. Klein, R. Reynaud
Institute of Fundamental Electronics, Bat. 220. Paris XI University, 91405 Orsay, France
Email: abdelhafid.elouardi@ief.u-psud.fr

Abstract—One of the solutions to reduce the computational complexity of image processing is to perform some low-level computations on the sensor focal plane. This paper presents a vision system based on a smart sensor. PARIS (Programmable Analog Retina-like Image Sensor) is the first prototype used to evaluate the architecture of an on-chip vision system based on such a sensor and a digital processor. The sensor integrates analog and digital computing units. This architecture makes the vision system more compact and increases the performances reducing the data flow exchanges with the digital processor. A system has been implemented as a proof-of-concept. This has enabled us to evaluate the performance needed for a possible implementation of a digital processor on the same chip. The approach is compared to two architectures implementing CMOS sensors and interfaced to the same processor. The comparison is related to image processing computation time, processing reliability, programmability, precision, bandwidth and subsequent stages of computations.

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 low-level image processing tasks (early-vision), a silicon retina integrates analog and/or digital processing circuits in the image-sensing element (pixels) [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).

In this paper, our main goal is to reach to a conclusion on the aptitude of the retinas to become potential candidates for a system on chip. 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 on chip vision system (SoC).

We propose a system-level architecture and a design methodology for the integration of an image processing within a CMOS retina on a single chip. We highlight a compromise between versatility, parallelism, processing speed and resolution. Our solution takes also into account the algorithms response times, the significant resolution of the sensor, while reducing power consumption for a use with embedded systems (in intelligent vehicles applications [3]) so as to increase the overall system performances.

We have done a comparison relating two different architectures dedicated for a vision system on chip. The first one implements a logarithmic CMOS APS imager (Active Pixel Sensor) and a microcontroller. The second involves the same microcontroller with a smart CMOS retina that implements hardware operators and analog processors. We have modelled two vision systems. The comparison is related mainly to image processing time and power consumption.

II. VISION SYSTEM BASED ON A CMOS RETINA

A. Circuit Description

PARIS (Parallel Analog Retina-like Image Sensor) is an architecture for which the concept of retinas is modeled implementing in the same circuit an array of pixels, integrating memories, and column-level analog processors [4]. The proposed structure is shown in figure 1.

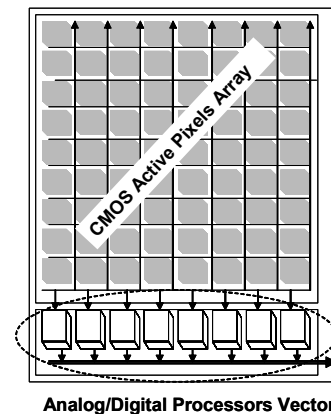


Fig. 1. PARIS architecture

This architecture allows a high degree of parallelism and a balanced compromise between communication and computations. Indeed, to reduce the area of the pixels and to allow the possibility to increase the fill factor, the image processing is achieved on a row of processors. Such approach presents the advantage to enable the design of complex processing units without decreasing the resolution. In return, because the parallelism is reduced to a row, the computations, which concern more than one pixel, have to be processed in a sequential way.

However, if a sequential execution increases the time of processing for a given operation, it allows a more flexible process. With this typical readout mechanism of the image, the column processing offers the advantages of parallel processing that permits low frequency and thus low power consumption. Furthermore, it becomes possible to chain basic functions in an arbitrary order, as in any digital SIMD (Single Instruction, Multiple Data) machine. The resulting low-level information extracted by the retina can be then processed by a digital microprocessor.

B. Pixels Description

The array of pixels constitutes the core of the architecture. Pixels can be randomly accessed. In some cases, the semi-parallel processing imposes to store intermediate and temporary results for every pixel in 4 MOS capacitors used as analog memories (figure 2).

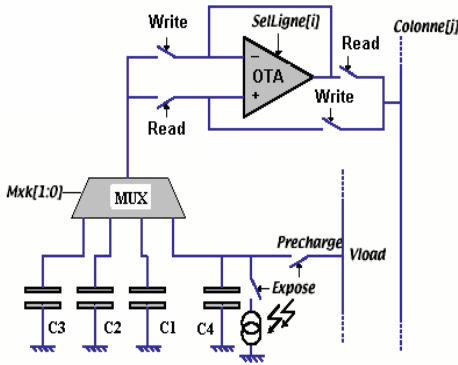


Fig. 2. Pixel diagram

The selected mode for the transduction of the light is the integration mode. The photosensor is then used as a current source that discharges a capacitor previously set to a voltage V_{ref} . One of the four analog memories is used to store the analog voltage deriving from the sensor. Two vertical bipolar transistors, associated in parallel, constitute the photosensor. For a given surface, compared to classic photodiodes, this disposal increases the sensitivity while preserving a large bandwidth [5] and a short response time can be obtained in a snapshot acquisition. The pixel area is $50 \times 50 \mu\text{m}^2$ when the Fill Factor is equal to 11%.

This approach eliminates the input/output bottleneck between different circuits even if there is a restriction on the implementation area, particularly for column width. Still, there is suppleness when designing the processing operators area: the implementation of the processing is more flexible relatively to the length of the columns. Pixels of the same column exchange their data with the corresponding processing element through a Digital Analog Bus (DAB). So as to access any of its four memories, each pixel includes a bidirectional (4 to 1) multiplexer. A set of switches makes possible to select the voltage stored in one of the four capacitors. This voltage is

copied out on the DAB thanks to a bi-directional amplifier. The same amplifier is used to write the same voltage on a chosen capacitor.

C. Programmable Analog Processors Vector

The pixels array is associated to a vector of processors operating in an analog/digital mixed mode. In this paper, we shall detail only the analog processing unit: APU (figure 3). Each APU implements three capacitors, one OTA (Operational Transconductance Amplifier) and a set of switches that can be controlled by a sequencer. The capacitance C_{out} plays the same role as the accumulator in a digital processor. The charge, loaded in C_{in1} , is transferred to C_{out} . According to the switches "Add" and "Sub", the charge of C_{in1} can be added or subtracted to the charge of C_{out} . The multiplication by a constant consists in applying a voltage V_{in} to the capacitor C_{in1} while C_{in2} is reset. Next, C_{in1} and C_{in2} are connected together. Since C_{in1} and C_{in2} are at equal value, the charge in C_{in1} is divided by two. Iterating the operation N times, this step leads to a charge in C_{in1} given in the equation (1) and more detailed examples of operations can be found in the reference [4].

$$Q_{in1} = \langle C_{in1} \cdot V_{in1} \rangle / 2^N \quad (1)$$

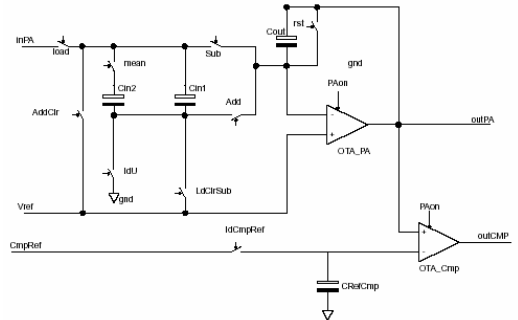


Fig. 3. Analog processor architecture

In order to validate this architecture, PARIS1 is the first prototype circuit that has been designed including 16×16 pixels and 16 analog processing units. This first circuit allows validating the integrated operators through some image processing algorithms like edge and movement detection. The vision chip has been design in a $0.6 \mu\text{m}$ CMOS technology. A microphotography and a view of the first prototype, PARIS1 circuit, are given in figure 4. The main characteristics of this vision chip are summarized in the following table:

TABLE I

| | |
|-----------------------------------|------------------------|
| Circuit area | 10 mm ² |
| Resolution (pixels) | 16x16 |
| Number of APUs | 16 |
| Pixel area | 50x50 μm ² |
| Area per processing unit | 50x200 μm ² |
| Clock frequency | 10 MHz |
| Processing Unit power consumption | 300 μW |
| Pixel power consumption | 100 μW |

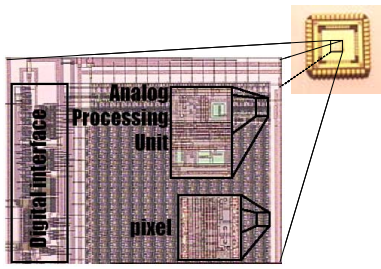


Fig. 4. Microphotography of PARIS1 sensor

D. PARIS Retina Based Vision System

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. On another side, a major advantage of retinas versus a CCD camera is the ability to integrate additional circuitry on the same chip along the array of pixels. As microcontrollers have become more prevalent and their abilities have increased, and since they have asset of high integration, high computing power and low consumption, these characteristics make them suited for the CMOS/APS sensors or smart retinas (known as intelligent sensors) as a finite state machine (FSM) giving instruction to a SIMD device. Such microcontrollers support various Operating Systems and communication drivers. This suggests that it should be possible to associate a CMOS Retina with a low cost microcontroller to implement an on chip vision system.

The retina, used as a standard peripheral of the microcontroller, is dedicated to image acquisition and low-level image processing. Thanks to the analog processing units, this retina extracts the low-level information (e.g. edges detection). Hence, the system, supported by the processor, that gives the high-level information, becomes more compact and can achieve processing suitable for real time applications.

To evaluate this architecture, we have implemented a prototype based on this architecture. It is a three design parts. The first two chips are the smart retina and the 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, that was developed for custom integrated circuits. The aim of the evaluation is the integration of the microprocessor with the retina (PARIS1 and ARM7TDMI) on a single chip.

The advantage of this architecture remains in the parallel execution of a large number of low level operations in the array by integrating operators shared by groups of pixels (rows or columns). 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]. Consequently, we obtain an architecture for which the PARIS1 circuit is dedicated for the regular and parallel image

processing. This circuit requires a programmable sequencer, from where the advantage of integrating a microprocessor with significant capacity of computing and low fuel consumption. Figure 5 shows the global architecture of the system and figure 6 gives an overview of the experimental module implemented for test and measurements.

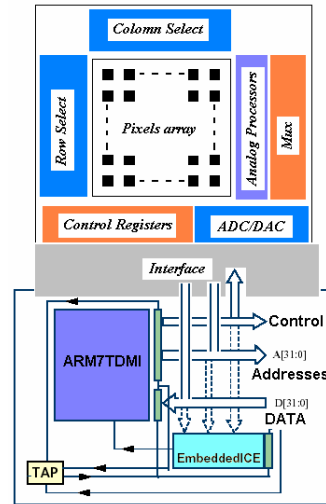


Fig. 5. Global architecture



Fig. 6. PARIS1 based vision system

E. Exposure time calibration and on chip image processing

Exposure time is an important parameter to control image contrast. This is the motivation for our development of a continuous auto-calibration algorithm that can manage this state for our vision system. This avoids pixels saturation and gives an adaptive amplification of the image, which is necessary to the post-processing.

The calibration concept is based on the fact that since the photo-sensors are used in an integration mode, a constant luminosity leads to a voltage drop that varies according to the exposure time. If the luminosity is high, the exposure time must decrease, on the other hand if the luminosity is low the exposure time should increase. Hence lower is the exposure time simpler is the image processing algorithms. This will globally decrease response time and simplify algorithms.

The algorithm consists in keeping the exposure time in the interval where all variations are linear and the exposure time is minimal. Figure 7 gives an example of images showing the adaptation of the exposure time to the luminosity.

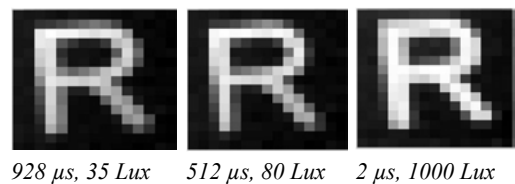


Fig. 7. Exposure time adaptation to the luminosity

The aim of this study is to investigate what image processing algorithms can be integrated on smart sensors as a part of early vision sequences and to examine their merits and the issues that designers should consider in advance.

In this paper, we do not wish to limit implementations to application-specific tasks, but to allow for general-purpose applications such as DSP-like image processors with programmability. The idea is based on the fact that some of early level image processing in the general-purpose chips are commonly shared with many image processors, which do not require programmability on their operation.

We have successfully implemented and tested different 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 the exposure time self calibration.

Since the input signal is always smaller than the input range, no saturation occurs. When successive operations are performed, the coefficient applied to the input signals must be chosen so that their sum remains lower than maximum range to prevent saturations. The real limitation comes from the dynamic (the lower bound is due to noise, component mismatch, non-linearity.) of the analog processor. Finally, calibration may be locally achieved thanks to the random access to pixels. Figure 8 gives examples of processed images.

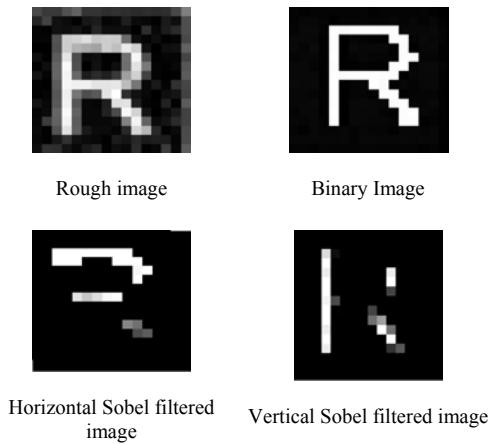


Fig. 8. Examples of on chip processing images

III. VISION SYSTEM BASED ON A CMOS SENSOR

A. Implementation

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 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 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 sensor chip has an on-chip 10 bit flash ADC and a digital gain/offset control. It behaves like a one Mbyte memory. The entire architecture is shown in figure 9. The figure 10 gives an overview of the CMOS sensor and the experimental module.

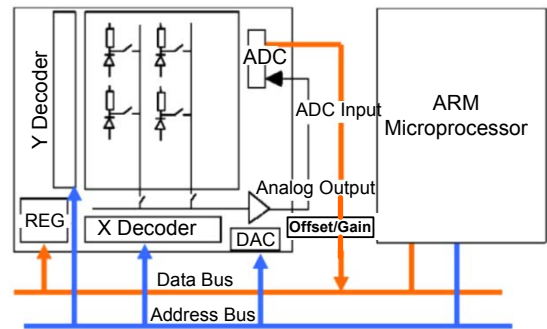


Fig. 9. System architecture based on the CMOS sensor

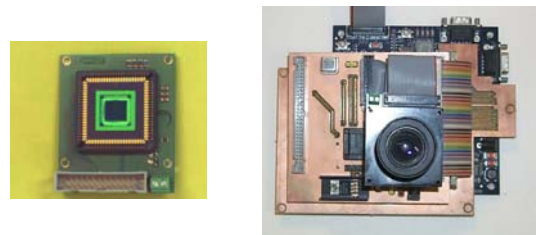


Fig. 10. Overview of the CMOS sensor (1024x1024 pixels) and the first experimental module

B. Calibration and Image processing

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 ($v \rightarrow v + \text{offset}$, 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 11).

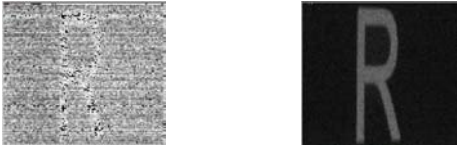


Fig. 11. Result of an FPN correction

For the FUGA1000 sensor based vision system, images are processed on the ARM microcontroller. We have implemented several algorithms of image processing similar to those established for PARIS based vision system. Other more complicated algorithms which require diversified computing with exponential power have been also implemented. We recall that to carry out comparisons relating to the processing times, we chose to use the same processor (ARM7TDMI) for the different implemented systems.

The Federico Garcia Lorca [7] filter is an example of the implemented image processing. This filter is a simplification of the Deriche filter [8], the recursive implementation of the optimal Canny filter. The smoother is applied horizontally and vertically on the image, in a serial way. Then a derivator is applied. Garcia Lorca derivator is, after simplification of Deriche derivator, a 3x3 convolution kernel instead of a recursive derivator. The response of such a filter is:

$$y(n) = (1 - \gamma)^2 x(n) + 2\gamma y(n-1) - \gamma^2 y(n-2) \quad (2)$$

with

$$\gamma = e^{-\alpha} \quad (3)$$

$x(n)$ is the pixel source value. $y(n)$ is the pixel destination value and n is the pixel index in a one dimensional table representing the image. γ is an exponential parameter allowing much more filtering flexibility, depending on the noise within the image. If the image is very noisy we use a very smoothing filter: $\alpha = [0.5, 0.7]$ otherwise we use bigger values of α : $\alpha = [0.8, 1.0]$. Figure 12 gives examples of smoothing filter and

derivator filter implemented with the FUGA-ARM vision system and applied to 120x120 pixels images.

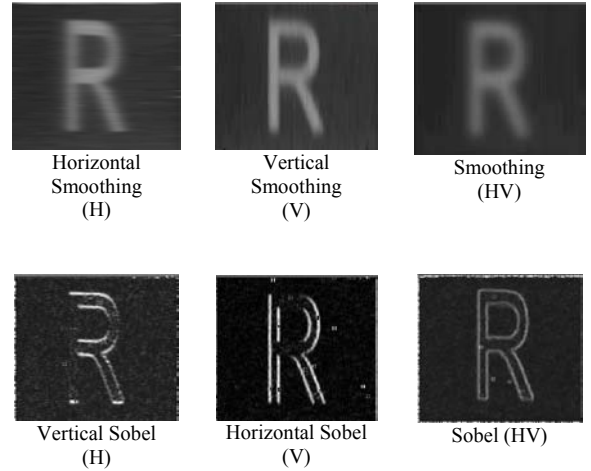


Fig. 12. Examples of image processing implemented with the FUGA1000 sensor based vision system

IV. COMPARISON

To achieve scene recognition, a vision system consists in the following principal tasks: image acquisition, low-level and high-level data processing. Our goal is to compare the vision system implementing the logarithmic CMOS sensor and the ARM microcontroller with the one based on PARIS retina. The comparison is related to the power consumption and image processing speed.

We have used the edge detection algorithms (Sobel filter operator, Garcia Lorca filter) to take several measurements of the computing time relating to the two architectures described below. For the retina-based system, these computations are carried out by the analog processors integrated on-chip. For the CMOS sensor based system, these computations are carried out by the ARM microcontroller.

A. Power Consumption

A characterization of the power consumption for PARIS based vision system has been achieved [9]. The total power of an NxN resolution and N analog processing units is:

$$P = \alpha.N^2 + \beta.N \quad (\mu W) \quad (4)$$

In this equation, α (100 μW) is the power consumption per pixel and β (300 μW) is the power consumption per analog processing unit. The 16x16 pixels circuit has a consumption of 30.4 mW.

The consumption of the FUGA1000 sensor is 0.25 mW per pixel and that of the ARM microcontroller is 14 mW (The

comparison does not take into account the power consumption of the microcontroller peripherals. RAM, ROM and logic glue consumption are excluded). This gives a consumption of 78 mW for a 16x16 pixels resolution.

Hence, when comparing the power consumption between the "CMOS sensor/ARM processor" like-system and the PARIS retina based vision system, at the same frequency; we conclude that the on chip solution allows better performances and lower power consumption.

B. Computing Time

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 13 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 [9].

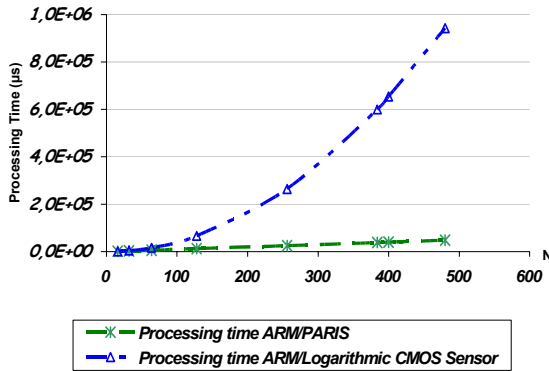


Fig. 13. Processing times for the two systems

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 analog processor vector.

The equation (5) gives the definition of the CPP (Cycle Per Pixel) of a processor. F_{CLK} is the processor frequency, T is the time computing, L is the rows number and C is the columns number:

$$CPP = (T \cdot F_{CLK}) / (L \cdot C) \quad (5)$$

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 like- system, the CPP factor is inversely proportional to the number of lines N . Figure 14 shows the evolution of the CPP for PARIS1 and CMOS sensor/ARM systems.

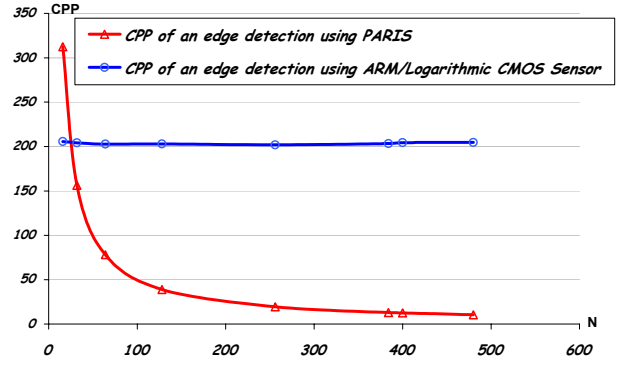


Fig. 14. Evolution of the CPP for the two systems

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-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 vision system [10].

REFERENCES

- [1] P. Dudek, J. Hicks, "A CMOS General-Purpose Sampled-Data Analogue Microprocessor". Proc. of the 2000 IEEE International Symposium on Circuits and Systems. Geneva, Suisse.
- [2] Y. Ni, J.H. Guan, "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] A. Elouardi, S. Bouaziz, R. Reynaud, "Evaluation of an artificial CMOS retina sensor for tracking systems". Proceeding of IEEE Intelligent Vehicles Symposium 2002, Versailles, France.
- [4] A. Dupret, J.O Klein, A. Nshare "A DSP-like Analog Processing Unit for Smart Image Sensors", International Journal of Circuit Theory and Applications 2002. 30: p. 595-609.
- [5] A. Dupret, E. Belhaire, J.C Rodier "A high current large bandwidth photosensor on standard CMOS Process" presented at EuroOpto'96, AFPAEC, Berlin, 1996.
- [6] D. Litwiller "CCD vs. CMOS: Facts and Fiction". The January 2001 issue of PHOTONICS SPECTRA, Laurin Publishing Co. Inc.
- [7] R. Deriche, "Fast algorithms for low level-vision". IEEE Transaction of Pattern Analysis and Machine Intelligence, vol 12-1, 1990.
- [8] F. Garcia Lorca, L. Kessal, D. Demigny "Efficient ASIC and FPGA implementation of IIR filters for real time edge detections", Proc. International Conference on Image Processing, IEEE ICIP 1997.
- [9] A. Dupret, J.O Klein, A. Nshare, "A programmable vision chip for CNN based algorithms". CNNA 2000, Catania, Italy: IEEE 00TH8509.
- [10] A. Elouardi, S. Bouaziz, A. Dupret, J.O Klein, R. Reynaud, "On Chip Vision System Architecture Using a CMOS Retina". Proceeding of IEEE Intelligent Vehicle Symposium, IV'04. Pages 206-211. ISBN 0-7803-8311-7. June 14-17, 2004. Parma, Italy.