Smart cameras on a chip: using complementary metal-oxide-semiconductor (CMOS) image sensors to create smart vision chips

Dominique Ginhac

Chapter 7: Smart cameras on a chip: using complementary metal oxide semiconductor (CMOS) image sensors to create smart vision chips

D. Ginhac, University of Burgundy, France

[Dominique Ginhac (dginhac@u-bourgogne.fr)]

Abstract: In this chapter, we introduce the fundamental concept of smart cameras on a chip or smart vision chips that simultaneously integrate on the same die image capture capability and highly complex image processing. Successive technology scaling has made possible the integration of specific processing elements designed at chip-level, at column-level or at pixel-level. To illustrate this continuous evolution, we survey three different categories of vision chips, exploring first the pioneering works on artificial retinas, then describing the most significant computational chips, and finally presenting the most recent image processing chips able to perform complex algorithms at a high frame rate.

Keywords: smart camera on a chip, vision chips, focal-plane image processing

1 Introduction

Today, digital smart cameras are rapidly becoming ubiquitous, due to reduced costs and increasing demands of multimedia applications. Improvements in the growing digital imaging world continue to be made with two main image sensor technologies: charge

coupled devices (CCD) and CMOS sensors. Historically, CCDs have been the dominant image-sensor technology. However, the continuous advances in CMOS technology for processors and memories have made CMOS sensor arrays a viable alternative to the popular CCD sensors. This led to the adoption of CMOS image sensors in several high-volume products, such as webcams, mobile phones or tablets for example. New technologies provide the potential for integrating a significant amount of VLSI electronics into a single chip, greatly reducing the cost, power consumption, and size of the camera (Fossum, 1993; Seitz, 2000; Litwiller, 2001). By exploiting these advantages, innovative CMOS sensors have been developed (see Fossum, 1997 and Bigas et al., 2006 for two detailed surveys on CMOS image sensors). Numerous works have focused on major parameters such as sensitivity (Krymski and Tu 2003; Murari et al. 2011), noise (Sumi 2006), power consumption (Hanson et al. 2010), voltage operation (Xu et al. 2002; Gao and Yadid-Pecht 2012), high-speed imaging (Dubois et al. 2008; El-Desouki et al. 2009) or dynamic range (Schrey et al. 2002; Fontaine 2011).

Moreover, the main advantage of CMOS image sensors is the flexibility to integrate signal processing at focal plane down to the pixel level. As CMOS image sensors technologies scale to 0.13 μm processes and under, processing units can be realized at chip level (system-on-chip approach), at column level by dedicating processing elements to one or more columns, or at pixel-level by integrating a specific processing unit in each pixel (El Gamal et al. 1999, El Gamal and Elthoukhy, 2005). By exploiting the ability to integrate sensing with analog or digital processing, new types of CMOS imaging systems can be designed for machine vision, surveillance, medical imaging, motion capture, pattern recognition among other applications. This extends the basic concept of electronic

camera on a chip proposed by Fossum (1997) to the more sophisticated concept of smart camera on a chip (also called vision chip) including both analog signal processing functions, analog-to-digital conversion, digital signal and image processing as described in Fig. 1. In this book chapter, we first define the concept of smart vision chips or smart cameras on a chip that integrate both sensing and complex image processing on the same chip. Image processing tasks are wide, spanning from basic image quality enhancements to complex applications managed by analog or digital microprocessors integrated in the sensor. These tasks can be performed globally at chip level (system-on-chip approach), regionally at column level by dedicating processing elements to one or more columns (typically analog to digital converters), or locally at pixel-level by integrating a specific unit in each pixel.

The remainder of this chapter is chronologically organized. It first describes the pioneering works on spatial and spatio-temporal image processing vision chips. These vision chips can be viewed as the first smart sensors and came to light during the 90's under the well-known term of silicon artificial retinas. Secondly, it talks about computational chips that have turned the first generation of vision chips into fully programmable smart vision chips reusable in many fields of application. In this section, we successively address the Cellular Neural Networks paradigm and the software-programmable SIMD processor arrays.

Thirdly, this chapter deals with high-speed image processing chips. Such chips integrate a processing element within each pixel based on a Single Instruction Multiple Data (SIMD) architecture, enabling massively parallel computations and leading to high framerates up

to thousands of images per second. In this section, we survey the state-of-the-art vision chips, both in the analog and digital domains.

Finally, we set out recent trends on smart vision chips. From a technological point of view, three-dimensional integrated imagers, based on 3D stacking technology, become an emerging solution to design powerful imaging systems because the sensor, the analog-to-digital converters and the image processors can be designed and optimized in different technologies, improving the global system performance. Combined with a BSI technology, 3D vision chips equipped allow high-speed signal processing and have an optical fill factor of 100%. From a conceptual point of view, electronic imaging aims at detecting individual photons. Single-photon imaging can be seen as the next step to reach in the design of smart imaging systems. Such vision chips are able to detect single photons by combine a high sensitivity with excellent photon timing properties in the range of a few tens of picoseconds. So, joint optimizations of the sensor, of the analog to digital converters and the processors offer opportunity to improve the sensor performance and allow the emergence of new applications such as real time 3D imaging with Time-of-Flight cameras (delivering simultaneously intensity images (2D) and ranges of the observed scene), medical imaging, molecular biology, astronomy and aerospace applications.

2 The concept of smart camera on a chip

CMOS image sensors have become increasingly mature and now dominate image sensor market shipments. Despite a large variety of applications, imaging systems always embed the same basic functions allowing the formation of a 2-D image from a real illuminated scene. These basic functions consist of 1) optical collection of photons (e.g. a lens), 2)

conversion of photons to electrons (e.g. a photodiode), 3) readout of the collected signal, and 4) logic control for driving the sensors. Note that readout may include some basic analog processing in order to enhance the image quality by removing temporal noise and fixed pattern noise (FPN). However, embedding such processing functions into a single chip does not turn a standard camera into a smart camera. The fundamental differences between a smart camera and a standard camera is that a smart camera must include a special intelligent image processing unit to run specific algorithms, in which the primary objective is not to improve images quality but to extract information and knowledge from images (Shi and Lichman, 2006). The close colocation of sensing and processing in a smart camera transforms the traditional camera into a smart sensor (Rinner and Wolf 2008).

In CCD technology, integrating electronics dedicated to specific image processing onto the silicon is by essence impractical (Litwiller 2005) because analog-to-digital conversion and signal processing functions are performed outside CCD sensors. On the contrary, CMOS image sensors and smart camera on a chip are intimately closed because CMOS technologies provide the ability to integrate complete imaging systems within the pixel sensor (Aw and Wooley 1996; Loinaz et al. 1998; Smith et al. 1998). Basically, a smart camera on a chip or a vision chip includes on the same die image capturing, analog-to-digital conversion, and analog / digital image processing as seen in Fig. 1.

The key advantages are 1) to release the host computer of complex pixel processing tasks by integrating the image sensor and the processors into a single chip; 2) to accelerate processing speed by using parallel processing elements; and 3) to minimize the data transfer between cameras and the outside world by only outputting extracted feature

information (Zhang et al., 2011). To summarize, the smart camera on a chip has the advantages of small size, high processing speed, low power consumption, and can be tailor-made for broad applications.

< Insert Figure 1 here >

As an illustrative example, the VISoc single chip smart camera designed by Albani et al. (2002) integrates a 320x256-pixel CMOS sensor, a 32-bit RISC processor, a neural co-processor, a 10-bit analog-to-digital converter and I/O onto a 6x6 mm2 single chip in a 0.35 μm standard CMOS process.

The greatest promise of CMOS smart cameras arises from the ability to flexibly integrate both sensing and complex image processing on the same chip (El Gamal and Eltoukhy, 2005). As CMOS image sensors technologies scale further down, smart vision chips are able to integrate focal-plane image processing tasks early in the signal chain. The range of pixel processing is wide, spanning from simple amplifiers dedicated to SNR enhancements to complete programmable digital or analog microprocessors in each pixel. Processing units can be realized at chip level (system-on-chip approach), at column level by dedicating processing elements to one or more columns (typically analog to digital converters), or at pixel-level by integrating a specific unit in each pixel. Historically, most of the researches have dealt with chip-level and column-level (Dickinson et al, 1995, Kemeny et al., 1997, Hong and Hornsey, 2002 ; Yadid-Pecht and Belenky, 2003; Acosta-Serafini et al. 2004; Kozlowski et al. 2005; Sakakibara et al. 2005). Indeed, pixel-

level processing has been generally dismissed for years because pixel sizes are often too large to be of practical use. However, as CMOS scales down, integrating a processing element at each pixel or group of neighboring pixels becomes more feasible since the area occupied by the pixel transistors decreases, leading to an acceptable small pixel size. A fundamental tradeoff must be made between three dependent and correlated variables: pixel size, processing element area, and fill-factor. This implies various points of view (Ginhac et al, 2008):

1. for a fixed fill-factor and a given processing element area, the pixel size is reduced with technology improvements. As a consequence, reducing pixel size increases spatial resolution for a fixed sensor die size.
2. for a fixed pixel size and a given processing element area, the photodiode area and the fill-factor increase as technology scales since the area occupied by the pixel transistors in each processing element decreases. It results in better sensibility, higher dynamic range and signal-to-noise ratio.
3. for a fixed pixel size and a given fill-factor, the processing element can integrate more functionalities since the transistors require less area as technology scales. Consequently, the image processing capabilities of the sensor increase.

In summary, each new technology process offers 1) to integrate more processing functions in a given silicon area, or 2) to integrate the same functionalities in a smaller silicon area. This can benefit the quality of imaging in terms of resolution or noise for example by integrating specific processing functions such as correlated double sampling (Nixon et al, 1995), anti blooming (Wu et al., 2001), high dynamic range (Decker et al., 1998), and even all basic camera functions (color processing functions, color correction,

white balance adjustment, gamma correction) onto the same camera-on-chip (Yoon et al., 2002). However, shrinking pixels size inevitably reach foreseen physical limits leading to poor performance of small pixels because of the reduced incident light on each pixel. Maintaining reasonable pixel performance – quantum efficiency, crosstalk, pixel capacity, angular signal response – and image quality – color reproduction, high dynamic range, limited chromatic aberration – when shrinking pixel size to small values such as 1.4 μm and smaller (typically used in mobile devices) is a big challenge for image sensor designers (Xiao et al., 2009).

3 The development of vision chip technology

From an historical point of view, the pioneering works have concentrated efforts on spatial and spatio-temporal image processing vision chips in the field of machine vision applications. These vision chips can be viewed as the first smart sensors and came to light during the 90's under the well-known term of silicon artificial retinas. Based on models of the vertebrate retina, they are able to implement some of its characteristics such as adaptation to local and global light intensity, and edge enhancement.

Generally, silicon artificial retina are arrays of identical pixels including significantly more transistors per pixel than the three or four found in typical APS-based sensors. This additional electronic circuitry performs pixel parallel processing over images immediately after they are captured without time-consuming and power- consuming image transfer. Spatial vision chips vision chips mainly implement basic neighborhood functions such as edge detection, smoothing, stereo processing, and contrast enhancement. On the other hand, spatio-temporal image processing vision chips are

mainly devoted to motion detection functions requiring the implementation of simultaneous time-space processing. Moini (2000) proposes an exhaustive overview of significant development up to 1997 about these two kinds of smart sensors while more recent developments are covered by Ohta (2008).

Artificial retinas were pioneered by Carver Mead in the late eighties, when he developed the first silicon retina that implements the first stages of retinal processing on a single silicon chip (Mead and Mahowald, 1988). This retina is based on models of computation of the vertebrate retina including specific structures such as the cones, the horizontal cells, and the bipolar cells. First of all, the cones, i.e. the light detectors, have been implemented using phototransistors and MOS-diode logarithmic current to voltage converters. Secondly, the outputs of the cones are then averaged, both spatially and temporally, by the horizontal cells. This averaging step is performed electronically using a hexagonal network of active resistors as seen on Fig. 2. Finally, bipolar cells detect the difference between the averaged output of the horizontal cells and the input.

< Insert Figure 2 here >

These first works leads to several other research projects dedicated to the design of analog artificial retina based on CMOS photodetectors combined with CMOS signal processing circuitry. Delbruck (1993) describes a two-dimensional silicon retina that computes a complete set of local direction-selective outputs. The chip motion computation uses unidirectional delay lines as tuned filters for moving edges. As a result,

the detectors are sensitive to motion over a wide range of spatial frequencies. Brajovic and Kanade (1996) describe a VLSI computational sensor using both local and global interpixel processing that can perform histogram equalization, scene change detection, and image segmentation in addition to normal image capture. Deutschmann and Koch (1998) present the first working analog VLSI implementation of a one-dimensional velocity sensor that uses the gradient method for spatially resolved velocity computation. Etienne-Cummings et al. (1999) implement a retina for measuring two-dimensional visual motion with two one-dimensional detectors. The pixels are built around a general-purpose analog neural computer and a silicon retina. Motion is extracted in two dimensions by using two 1-D detectors with spatial smoothing orthogonal to the direction of motion. In 2000, the same team presents a silicon retina chip with a central foveated region for smooth-pursuit tracking and a peripheral region for saccadic target acquisition (Etienne-Cummings et al., 2000). This chip has been used as a person tracker in a smart surveillance system and a road follower in an autonomous navigation system.

4 From special purpose chips to smart computational chips

One of the main drawbacks of the above-mentioned works is that these vision chips are not general-purpose. In other words, many vision chips are not programmable to perform different vision tasks. They are often built as special-purpose devices, performing specific and dedicated tasks, and not really reusable in another context (Dudek and Hicks, 2001). This inflexibility is particularly restrictive and even unacceptable for such vision systems that aim to flood several consumer markets. So, the main challenge when designing a smart vision system is to design a compact but versatile and fully programmable processing element, known as computational chip.

For this purpose, the processing function can be based on the paradigm of Cellular Neural Networks (CNN), introduced by Chua and Yang in 1988 (Chua and Yang, 1988a; 1988b). CNN can be viewed as a very suitable framework for systematic design of image processing chips (Roska and Rodriguez-Vazquez, 2000). The complete programmability of the interconnection strengths, its internal image-memories, and other additional features make this paradigm a powerful front-end for the realization of simple and medium-complexity artificial vision tasks (Espejo et al. 1996; Morfu et al. 2008). Some proof-of-concept chips operating on preloaded images have been designed (Rekeczky et al. 1999; Czuni and Sziranyi 2000). Only a small amount of researches have integrated CNN on real vision chips. As example, Espejo et al. (1998) report a 64x64 pixel programmable computational sensor based on a CNN. This chip is the first fully operational CNN vision-chip reported in literature which combines the capabilities of image-transduction, programmable image-processing and algorithmic control on a common silicon substrate. It has successfully demonstrated operations such as low-pass image filtering, corner and border extraction, and motion detection. More recently, Galan et al. (2003) have focused on the development of CNN-based sensors with a chip including 1024 processing units arranged into a 32x32 grid corresponding to approximately 500000 transistors in a standard 0.5 µm CMOS technology. An enhanced 128x128 version was also described in Rodriguez-Vazquez et al. (2004). The chip designed in a 0.35 µm standard CMOS technology, contains about 3.75 million transistors and exhibits peak computing figure of 330 GOPS. Each processing element in the array contains a reconfigurable computing kernel capable of calculating linear convolutions on 3x3 neighborhoods in less than 1.5 µs, Boolean combinations in less

than 200 ns, arithmetic operations in about 5 µs, and CNN-like temporal evolutions with a time constant of about 0.5 µs. Successive evolutions of these chips are presented on the historical roadmap depicted in Fig. 3.

<Insert Figure 3 here>

However, hardware realization of such chips has turned out to be difficult because they suffer from large area and high power consumption. In the above-mentioned vision chips, the pixel size is often over than 100 µm x 100 µm. Obviously, these dimensions cannot be considered as realistic dimensions for a real vision chip and numerous pioneering works have been abandoned nowadays. However, a major part of this crucial problem should be resolved in future years by using the new emergent CMOS technologies. Indeed, CMOS image sensors directly benefit from technology scaling by reducing pixel size, increasing resolution and integrating more analog and digital functionalities on the same chip with the sensor.

Other architectures in this category are the SCAMP family - Simd Current-mode Analog Matrix Processor (Dudek, 2005; Dudek and Carey, 2006) of software-programmable SIMD (Single instruction, multiple data) processor arrays implementing a variety of low-level image processing tasks. A SIMD processors array is built around multiple processing elements that simultaneously perform the same operation on different data. In the field of image sensors, the key idea is the introduction of an Analog Processing

Element (APE) per pixel, operating on the pixel value. The APE executes software instructions in a similar manner to a digital processor, but it operates on analog samples of data. The SCAMP-3 chip, described on the Fig. 4, fabricated in a 0.35 µm CMOS technology contains a 128x128 processor array and achieves cell density of 410 processors/mm2 (a single cell measures under 50 µm x 50 µm).

<Insert Figure 4 here>

The same team worked also on other complementary vision chips called ACLA (Asynchronous Cellular Logic Array) and ASPA (Asynchronous / Synchronous Processor Array). ACLA (Dudek, 2006; Lopich and Dudek, 2011) is an asynchronous cellular processor array that facilitates binary trigger- wave propagations, extensively used in various image-processing algorithms. A proof- of-concept array of 2460 cells has been fabricated in a 0.35 µm CMOS process.

The ASPA family (Lopich and Dudek 2008) includes vision chips embedding fine- grain processor arrays based on novel control schemes, where individual processors are triggered, as data are available at their neighbors, optimizing speed and power consumption of the devices. The aim is to provide image processing engines suitable for both low-level, pixel-based operations (filtering, feature detection etc.) as well as more global, object-based algorithms, such as object reconstruction, skeletonization, watershed

transform, distance transform etc. The latest chip in this family (ASPA-3) has a 160x80 processor array fabricated in a 180 nm CMOS technology with a chip area of 50 mm2. Another approach, which is potentially more programmable, is the PVLSAR (Programmable Versatile Large Scale Artificial Retina) retina chip (Paillet et al., 1998; 1999). The PVLSAR is a highly integrated CMOS smart sensor device comprising an SIMD (Single Instruction Multiple Data) array of 128x128 pixel processors. Each pixel processor contains a photodiode as the optical sensor and a logical unit. The retina chip is a fine grain massively parallel SIMD processing unit with optical input. It is fully programmable and very powerful especially on low-level image processing. The PVLSAR can perform a plethora of retinotopic operations including early vision functions, image segmentation, and pattern recognition.

To summarize, Table 1 and Table 2 respectively show an overview of some representative analog and digital computational chips based on different alternatives for implementing vision processing at focal plane.

[Insert Table 1 near here]

[Insert Table 2 near here]

5 From video rate applications to high-speed image processing chips

The random access readout of CMOS image sensors provides the potential for high-speed readout and window-of-interest operations at low power consumption (Gamal and

Elthoukhy, 2005), especially when dealing with low-level image processing algorithms. Indeed, such low-level image processing tasks are inherently pixel-parallel in nature. So, integrating a processing element within each pixel based on a Single Instruction Multiple Data (SIMD) architecture is a natural candidate to cope with the temporal processing constraints (Cembrano et al. 2004). This approach is quite interesting for several aspects. First, SIMD image-processing capabilities at focal plane have not been fully exploited because the silicon area available for the processing elements is very limited. Nevertheless, this enables massively parallel computations allowing high framerates up to thousands of images per second. The parallel evaluation of the pixels by the SIMD operators leads to processing times, independent of the resolution of the sensor. In a standard system, in which low-level image processing is externally implemented after digitization, processing times are proportional to the resolution leading to lower framerates as resolution increases. Several papers have demonstrated the potentially outstanding performance of CMOS image sensors (Krymski et al., 1999; Stevanovic et al. 2000, Kleinfelder et al., 2001). Krymski et al. (1999) describe a high speed (500 frames/s) large format 1024 × 1024 Active Pixel Sensor (APS) with 1024 ADCs. Stevanovic et al. (2000) describe a 256 × 256 APS which achieves more than 1000 frames/s with variable integration times. Kleinfelder et al. (2001) describe a 352 × 288 Digital Pixel Sensor (DPS) in which analog-to-digital (A/D) conversion is performed locally at each pixel, and digital data is read out from the pixel array in a manner similar to a random access digital memory, achieving 10,000 digital frames/s capturing and 1 Giga-pixels/s for readout.

Secondly, the high speed imaging capability of CMOS image sensors can benefit the implementation of new complex applications at standard rates and improve the performance of existing video applications such as motion vector estimation (Handoko et al., 2000; Lim and El Gamal, 2001; Liu and El Gamal, 2001a), multiple capture with dynamic range (Yang et al., 1999; Yadid-Pecht and Belenky, 2001; Stoppa et al., 2002), motion capture (Liu and El Gamal, 2001b), and pattern recognition (Wu and Chiang, 2004). Indeed, standard digital systems are unable to operate at high framerates, because of the high output data rate requirements for the sensor, the memory, and the processing elements. Integrating the memory and processing with the sensor on the same chip removes the classical input output bottleneck between the sensor and the external processors in charge of processing the pixel values. Indeed, the bandwidth of the communication between the sensor and the external processors is known as a crucial aspect, especially with high-resolution sensors. In such cases, the sensor output data flow can be very high, and needs a lot of hardware resources to convert, process and transmit a lot of data. So, integrating image processing at the pixel-level can alleviate the high data rate problem because the pixel values are pre-processed on-chip by the SIMD operators before sending them to the external world via the communication channels. This will result in data reduction, which allows sending the data at lower data-rates, and reduces the effect of the computational-load bottleneck.

Thirdly, one of the main drawbacks to design specific circuits integrating sensing and processing on the same chip is that these vision chips are often built as special-purpose devices, performing specific and dedicated tasks, and not reusable in another context (Dudek and Hicks, 2001). So, it can be widely beneficial to integrate a versatile device,

whose functionality can be easily modified. Moreover, except the basic operations such as convolutions with small masks, the majority of computer vision algorithms require the sequential execution of different successive low-level image processing on the same data. So, each processing element must be built around a programmable execution unit, communication channels, and local memories dedicated to intermediate results. Because of the very limited silicon area, the processing units are necessarily very simple, providing the best compromise between various factors such as versatility, complexity, parallelism, processing speeds and resolution.

To sum up, the flexibility to integrate processing down to the pixel level allows us to rearchitect the entire imaging system to achieve much higher performance (El Gamal and Eltoukhy, 2005). The key idea is 1) to capture images at a very high framerate, 2) to process the data on each pixel with a SIMD programmable architecture exploiting the high on-chip bandwidth between the sensor, the memory and the elementary processors and 3) to provide results at the best framerate depending on the complexity of the image processing.

To illustrate this concept, we designed a massively parallel, SIMD vision chip implementing low-level image processing based on local masks (Dubois et al. 2008; Ginhac et al. 2010). The core includes a two-dimensional array of 64 × 64 identical processing element (PE). Each of the PE is able to convolve the pixel value issued from the photodiode by applying a set of mask coefficients to the image pixel values located in a small neighborhood. The key idea is that a global control unit can dynamically reconfigure the convolution kernel masks and then implements the most part of low-level image processing algorithms. This confers the functionality of programmable processing

devices to the PEs embedded in the circuit. As seen in Fig. 5, each individual PE includes the following elements: 1) a photodiode dedicated to the optical acquisition of the visual information and the light-to-voltage transduction 2) a set of two Analog Memory, Amplifier and Multiplexer structures called $[AM]^2$, which serve as intelligent pixel memories and are able to dissociate the acquisition of the current frame in the first memory and the processing of the previous frames in the second memory, and (3) an Analog Arithmetic Unit named A^2U based on four analog multipliers, which performs the linear combination of the four adjacent pixels using a 2 × 2 convolution kernel. In brief, each PE includes 38 transistors integrating all the analog circuitry dedicated to the image processing algorithms. The global size of the PE is 35 µm x 35 µm (1225 µm2). The active area of the photodiode is 300 µm2, giving a fill-factor of 25 %. In terms of pixel size and fill-factor, this chip share similar characteristics with the vision chips previously described in Table 1 and Table 2. The chip can capture raw images up to 10 000 frames per second and runs low-level programmable image processing at a framerate of 2 000 to 5 000 frames per second.

<Insert Figure 5 here>

Based on the same principle, a large number of equivalent pixel-level image sensors have been designed during the past ten years, taking advantage of pixel-level processing elements to achieve massively parallel computations and thus, to exploit the high-speed imaging capability of CMOS image sensors. In an increasingly digital world, we can imagine that the most part of state-of-the-art imaging systems has become almost entirely

digital, including analog to digital conversion and digital processing in the chip. But, for low-level image processing, an analog or a mixed- approach can offer superior performance leading to a smaller, faster, and lower power solution than digital processors (Martin et al. 1998). Indeed, low-level image processing usually involves basic operations using local masks. These local operations are spatially dependent on other pixels around the processed pixel. Since the same type of operations is applied to a very large data set, these low-level tasks are computationally intensive and require a high bandwidth between the image memory and the digital processor. Following this idea, analog vision chips such as those described in Brea et al (2004), Rodriguez-Vazquez et al. (2004), Dudek and Hicks (2005), Chi et al. (2007), Massari and Gottardi (2007), or Kim et al (2008) are characterized by a very compact area, optimized dedicated processing, high processing speed, and impressive performance, but suffers from low flexibility.

On the opposite, as integrated circuits keep scaling down following Moore's Law, recent trends show a significant number of papers discussing the design of digital imaging systems that take advantage of the increasing number of available transistors integrated in each pixel in order to perform analog to digital conversion, data storage and sophisticated digital imaging processing. Following the first digital pixel sensors designed by Kleinfelder et al (2001), numerous works have been conducted to optimize this new design concept. While on-pixel conversion provides a number of advantages, there are still many challenges and issues to be solved and more particularly the dynamic range limitation due to the number of bits used for the conversion (Kitchen et al., 2005), the optimization of the silicon area of the memory (Zhang et al, 2011), the compression of

the data (Zhang et al., 2007). Then, digital data can be processed by specific digital computational elements such as those described in Komuro et al. (2004), Leon-Salas et al. (2007), Miao et al. (2008), Komuro et al. (2009), or Lin et al (2009). Such vision chips offer more versatility and flexibility, more programmability and perform more complex algorithms from low- to mid level image processing.

6 Future trends

CMOS active-pixel image sensors have become increasingly mature because of the continuous technological advances. They now dominate image sensor market shipments, both in volume and in revenue. Successive improvements result in the shrink of pixel size. State-of-the-art CMOS image sensors used typically in mobile devices are built with 1.4 μm generation pixels. However, with such pixel pitches, the number of incident photons is limited and sensor optical response can be blocked or interfered by metal layers in traditional front-side illumination (FSI) sensor structure. So, there is a growing interest in the mass production of backside-illuminated (BSI) devices. BSI CMOS sensors have the metal wiring layer positioned below the photodiode layer which means light is not reflected by pixel wiring and then lost. Due to this design, the photodiodes receive more light and the sensor is able to produce higher quality images in dark or low light scenes.

As photo-detector size shrinks, approaches that decouple sensing from readout and processing by employing separate stacked structures for photo-detection and processing is also of growing interest. The main idea is to exploit the features of the 3D technologies for the fabrication of a stack of very thin and precisely aligned CMOS APS layers, each of these stacks being optimized for a given function. 3D chips are obtained by

segmenting 2D chips into functional blocks, stacking these blocks, and interconnecting them with short signal paths As an example, a CMOS digital smart vision system consisting of a CMOS sensor, analog to digital converters and digital programmable processing elements is particularly suitable for a 3D integration. Combined with a BSI technology, image sensor devices equipped with 3D technologies allow high-speed signal processing and have an optical fill factor of 100%. The optimization of the sensor, the ADC and the processors fabricated in different technologies offers opportunity to improve the sensor performance and decrease the cost (Suntharalingam et al. ,2009; Motoyoshi and Koyanagi 2009; Yeh et al. 2011].

The ability to integrate dedicated signal processing functions within the pixel site is finally the most significant difference between CMOS and CCD sensors. In CMOS sensors, data processing can take place concurrently with image acquisition and can contribute to acquire complementary data, such as 3D information. Real time 3D imaging is a rapidly emerging field, due to the fact that 3D image acquisition systems can be used in a variety of applications such as automobile, robot vision systems, security and so on. Time-of-Flight (TOF) image sensors using high frequency modulation of near-infrared light have emerged as a viable alternative to stereo and structured light imaging for capturing range information (Lee et al., 2011). However, effort will have to be done to improve performance and precision that remain critical to mass adoption in consumer electronics applications.

Finally, the ultimate sensitivity in electronic imaging is the detection of individual photons. Single-photon imaging can be seen as the next step to reach in the design of smart imaging systems (Seitz and Theuwissen, 2011). Single-Photon Avalanche Diodes

(SPADs) are, as their name suggests, highly sensitive optical detectors capable of distinguishing single photons. They combine this high sensitivity with excellent photon timing properties in the range of a few tens of picoseconds. In the past, SPADs were designed using custom processes. Recent works have dealt about integration of SPADs on standard CMOS processes. One of the major challenges still remaining is the creation of large arrays of SPADs with reduced pitch and capable of 1) single photon sensitivity and 2) precise photon timing with sub-nanosecond resolution. Applications of such innovative sensors are in many fields: 3D scanning, medical imaging, molecular biology, astronomy and aerospace applications, etc and more generally in any situation requiring brilliant pictures captured under extreme low-light conditions.

7 Conclusion

In this chapter, we have introduced the fundamental concept of smart cameras on a chip or smart vision chips that simultaneously integrate on the same die image capture capability and highly complex image processing. Successive technology scaling has made possible the integration of analog or digital processing elements into single pixels.
To illustrate this continuously evolution, we chronologically surveyed three different categories of vision chips, exploring first the pioneering works on artificial retinas, then describing the most significant contributions to computational chips, and finally presenting the most recent state-of-the-art high speed image processing chips able to perform complex algorithms at a high framerate. During this detailed survey, we have outlined the challenges of implementing complex image processing applications at focal-plane and the underlying complexity of resolving the fundamental tradeoffs between

high-performance tailor-made chips and less powerful but more reusable programmable chips.

This chapter ends with a particular focus on trending topics, mentioning recent technological innovations such as backside illumination (BSI) and 3-dimensional stacking. The democratization of these recent technological developments may lead to the design of very innovative smart image processing chips. We also mention the new challenges of designing efficient single-photon imaging chips able to sense every individual photon, both spatially and temporally with high precision.

8 References

Acosta-Serafini, P., Ichiro, M., & Sodini, C. (2004), A 1/3 VGA linear wide dynamic range CMOS image sensor implementing a predictive multiple sampling algorithm with overlapping integration intervals, IEEE Journal of Solid-State Circuits, 39(9), 1487–1496.

Albani, L., Chiesa, P., Covi, D., Pedegani, G., Sartori, A., and Vatteroni, M. 2002. Visoc: A smart camera soc. In Solid-State Circuits Conference, 2002. ESSCIRC 2002. Proceedings of the 28th European. 367 –370.

Aw, C. H. and Wooley, B. (1996), A 128x128-pixel standard-CMOS image sensor with electronic shutter, IEEE Journal of Solid State Circuits 31, 12, 1922–1930.

Bigas, M., Cabruja, E., Forest, J., and Salvi, J. (2006), Review of cmos image sensors, Microelectronics Journal 37, 5, 433 – 451.

Brajovic, V. and Kanade, T. (1996), A sorting image sensor: an example of massively parallel intensity-to- time processing for low-latency computational sensors. In

Proceedings of IEEE International Conference on Robotics and Automation, Vol. 2. 1638 –1643.

Brea, V. M., Vilariño, D. L., Paasio, A. and Cabello, D. (2004) Design of the processing core of a mixed-signal CMOS DTCNN chip for pixel-level snakes, IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 51, no. 5, pp. 997–1013, May.

Cembrano, G., Rodriguez-Vazquez, A., Galan, R., Jimenez-Garrido, F., Espejo, S., and Domniguez-Castro, R. (2004), A 1000 fps 128x128 vision processor with 8-bit digitized I/O, IEEE Journal of Solid State Circuits, 39, 7, 1044 – 1055.

Chi, Y., M. Mallik, U., Clapp, M. A., Choi, E., Cauwenberghs, G. and Etienne-Cummings, R. (2007), CMOS camera with in-pixel temporal change detection and ADC, IEEE J. Solid-State Circuits, vol. 42, no. 10, pp. 2187–2196, Oct..

Chua, L. and Yang L. (1988a). Cellular neural networks: theory, IEEE Transactions on Circuits and Systems, 35, 10, 1257 –1272.

Chua, L. and Yang L. (1988b). Cellular neural networks: applications, IEEE Transactions on Circuits and Systems, 35, 10, 1273 –1290.

Czuni, L. and Szianyi, T. (2000), Motion segmentation and tracking optimization with edge relaxation in the cellular nonlinear network architecture. In Proceedings of the 6th IEEE International Workshop on Cellular Neural Networks and Their Applications, CNNA 2000, 51 –56.

Decker, S., McGrath, D., Brehmer, K., & Sodini, C. (1998), A 256 x 256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output, IEEE Journal of Solid-State Circuits, 33(12), 2081–2091.

Delbruck, T. (1993), Silicon retina with correlation-based, velocity-tuned pixels, IEEE Transactions on Neural Networks, 4, 3, 529 –541.

Deutschmann, R.and Koch, C. (1998). An analog VSLS velocity sensor using the gradient method, In Proceedings of the 1998 IEEE International Symposium on Circuits and Systems, ISCAS '98, Vol. 6. 649 –652.

Dickinson, A., Mendis, S., Inglis, D., Azadet, K. and Fossum, E.R. (1995) CMOS Digital Camera with Parallel Analog to Digital Conversion Architecture, in Program of 1995 IEEE Workshop on CCDs and Advanced Image Sensors, Dana Point, CA, April 20-22.

Dubois, J., Ginhac, D., Paindavoine, M., and Heyrman, B. (2008), A 10 000 fps CMOS sensor with massively parallel image processing, IEEE Journal of Solid-State Circuits, 43, 3, 706 –717.

Dudek, P. 2005. Implementation of SIMD vision chip with 128x128 array of analogue processing elements. In IEEE International Symposium on Circuits and Systems, ISCAS, vol 6.,5806 – 5809

Dudek P. and Hicks, P. J. (2001), An analogue SIMD focal-plane processor array, in Proceedings of the International Symposium on Circuits and Systems (ISCAS), vol. 4, 6-9, pp. 490–493.

Dudek P. and Hicks, P. J. (2005), A general-purpose processor-per-pixel analog SIMD vision chip, IEEE Trans. Circuits Syst. I, vol. 52, no. 1, pp. 13–20, Jan. 2005.

Dudek P. (2006), An asynchronous cellular logic network for trigger-wave image processing on fine-grain massively parallel arrays. IEEE Transactions on Circuits and Systems II: Express Briefs, 53, 5, 354–358.

Dudek P. and Carey, S. 2006. General-purpose 128x128 SIMD processor array with integrated image sensor, Electronics Letters, 42, 12, 678 – 679.

Dudek, P., Barr, D., Lopich, A., and Carey, S., (2006) Demonstration of real-time image processing on the SCAMP-3 vision system, 10th International Workshop on Cellular Neural Networks and Their Applications, CNNA '06.

El-Desouki, M., Jamal Deen, M., Fang, Q., Liu, L., Tse, F., and Armstrong, D. (2009), CMOS image sensors for high speed applications, Sensors 9, 1, 430–444.

El Gamal, A. and Eltoukhy, H. (2005), CMOS image sensors, IEEE Circuits and Devices Magazine, 21, 3, 6 – 20.

El Gamal, A., Yang, D., and Fowler, B. (1999) Pixel level processing – Why, What and How? In Proceedings of the SPIE Electronic Imaging ’99 conference. Vol. 3650. 2–13.

Etienne-Cummings, R., Van Der Spiegel, J., and Mueller, P. (1999), Hardware implementation of a visual-motion pixel using oriented spatiotemporal neural filters, IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, 46, 9, 1121 –1136.

Etienne-Cummings, R., Van Der Spiegel, J., Mueller, P. and Zhang, M. (2000), A foveated silicon retina for two-dimensional tracking, IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, 47, 6, 504 –517.

Etienne-Cummings, R., Kalayjian, Z., And Cai, D. (2001), A programmable focal-plane MIMD image processor chip. IEEE Journal of Solid-State Circuits, 36, 1, 64–73.

Espejo, S., Carmona, R., Dominguez-Castro, R., and Roriguez-Vazquez, A. (1996), A 0.8µm CMOS programmable analog-array-processing vision-chip with local logic

and image-memory, In Proceedings of the 22nd European Solid- State Circuits Conference. ESSCIRC ’96, 280 –283.

Espejo, S., Dominguez-Castro, R., Linan, G., and Roriguez-Vazquez, A. (1998). A 64x64 CNN universal chip with analog and digital I/O. In IEEE International Conference on Electronics, Circuits and Systems, Vol. 1, 203 –206.

Fontaine, R. (2011), A review of the 1.4μm pixel generation. In International. Image Sensor Workshop. Hokkaido, Japan, 1–4.

Fossum, E. (1993), Active pixel sensors: Are CCDs dinosaurs? International Society for Optical Engineering (SPIE), 1900, 2–14.

Fossum, E. (1997), CMOS image sensor: Electronic camera on a chip, IEEE Transactions on Electron Devices, 44(10), 1689– 1698.

Galan, R., Jimenez-Garrido, F., Dominguez-Castro, R., Espejo, S., Roska, T., Rekeczky, C., Petras, I., and Rodriguez-Vazquez, A. (2003), A bio-inspired two-layer mixed-signal flexible programmable chip for early vision, IEEE Transactions on Neural Networks, 14, 5, 1313 – 1336.

Gao, Q. and Yadid-Pecht, O. (2012), A low-power block-based CMOS image sensor with dual vdd. IEEE Sensors Journal, 12, 4, 747 –755.

Ginhac, D., Dubois, J., Paindavoine, M. and Heyrman, B. (2008), A SIMD Programmable Vision Chip with High Speed Focal Plane Image Processing, EURASIP Journal of Embedded Systems, vol. 2008, Article ID 961315, 13 pages.

Ginhac, D., Dubois, J., Heyrman, B., and Paindavoine, M. (2010), A high-speed programmable focal- plane SIMD vision chip. Analog Integrated Circuits and Signal Processing 65, 389–398.

Handoko, D., Takokoro, K., Kumahara, M., and Matsuzawa, A. (2000), A CMOS image sensor for local-plane motion vector estimation," in Symposium of VLSI Circuits, vol. 3650, pp. 28–29.

Hanson, S., Foo, Z., Blaauw, D., and Sylsverter, D. (2010), A 0.5 v sub-microwatt CMOS image sensor with pulse-width modulation read-out. IEEE Journal of Solid-State Circuits, 45, 4, 759 –767.

Hong, C. and Hornsey, R. (2002) On-chip binary image processing with CMOS image sensors, Proc. SPIE 4669, 125

Kemeny, S.E., Panicacci, R., Pain, B., Matthies, L. and Fossum, E.R. (1997) Multiresolution image sensor, IEEE Trans. Circuits and Systems for Video Technology, vol. 7(4), pp. 575-583 .

Kim, D., Cho, J., Lim, S., Lee, D. and Han, G. (2008), A 5000 S/s single- chip smart eye-tracking sensor, in IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers, , pp. 46–594.

Kitchen, A.; Bermak, A.; Bouzerdoum, A.; (2005), A digital pixel sensor array with programmable dynamic range, IEEE Transactions on Electron Devices, , vol.52, no.12, pp. 2591- 2601.

Kleinfelder, S., Lim, S., Liu, X. and El Gamal, A. (2001), A 10 000 Frames/s CMOS Digital Pixel Sensor, IEEE Journal of Solid-State Circuits, vol. 36, no. 12, pp. 2049–205.

Komuro, T., Kagami, S., and Ishikawa, M. (2004), A dynamically reconfigurable SIMD processor for a vision chip, IEEE J. Solid-State Circuits, vol. 39, no. 1, pp. 265–268, Jan.

Komuro, T., Iwashita, A., and Ishikawa, M. (2009), A QVGA-size pixel-parallel image processor for 1,000-fps vision, IEEE Micro, vol. 29, no. 6, pp. 58–67, Nov.-Dec.

Kozlowski, L., Rossi, G., Blanquart, L., Marchesini, R., Huang, Y., Chow, G., Richardson, J., & Standley, D. (2005), Pixel noise suppression via SoC management of target reset in a 1920 x 1080 CMOS image sensor, IEEE Journal of Solid-State Circuits, 40(12), 2766–2776.

Krymski, A. Van Blerkom, D. Andersson, A. Bock, N. Mansoorian, B. and Fossum, E. (1999), A high speed, 500 frames/s, 1024x1024 CMOS active pixel sensor," Digest of Technical Papers of Symposium on VLSI Circuits, pp. 137–138.

Krymski, A. and Tu, N. (2003), A 9-v/lux-s 5000-frames/s 512x512 CMOS sensor, IEEE Transactions on Electron Devices, 50, 1, 136 – 143.

Lee, T.Y.; Lee, Y.J.; Min, D.K.; Lee, S.H.; Kim, W.H.; Kim, S.H.; Jung, J.K.; Ovsiannikov, I.; Jin, Y.G.; Park, Y.D.; Fossum, E.R.; Chung, C.H., (2011), A 192×108 pixel ToF-3D image sensor with single-tap concentric-gate demodulation pixels in 0.13 μm technology, IEEE International Electron Devices Meeting (IEDM), pp.8.7.1-8.7.4.

Leon-Salas, W. D., Balkir, S., Sayood, K., Schemm, N., and Hoffman, M. W., (2007), A CMOS imager with focal plane compression using predictive coding, IEEE J. Solid-State Circuits, vol. 42, no. 11, pp. 2555–2572, Nov. 2007.

Lim, S., and El Gamal, A. (2001), Integrating Image Capture and Processing – Beyond Single Chip Digital Camera, in Proceedings of the SPIE Electronic Imaging '2001 conference, vol. 4306, San Jose, CA.

Lin, Q.-Y., Miao, W., Zhang, W.-C., Fu, Q.-Y., and Wu, N.-J. (2009) A 1000 frames/s programmable vision chip with variable resolution and row-pixel-mixed parallel image processors, Sensors, vol. 9, pp. 5933–5951.

Litwiller, D. (2001), CCD vs. CMOS: Facts and fiction, Photonics Spectra, 35, 154–158.

Liu, X., and El Gamal, A. (2001a) Photocurrent estimation from multiple non- destructive samples in a CMOS image sensor, in Proceedings of the SPIE Electronic Imaging '2001 conference, vol. 4306, San Jose, CA, january 2001.

Liu, X. and El Gamal, A. (2001b). Simultaneous image formation and motion blur restoration via multiple capture. In IEEE International Conference on Acoustics, Speech and Signal Processing. Vol. 3. 1841–1844.

Loinaz, M., Singh, K., Blanksby, A., Inglis, D., Azadet, K., and Ackland, B. (1998), A 200mv 3.3v CMOS Color Camera IC Producing 352 × 288 24-b Video at 30 Frames/s, IEEE Journal of Solid-State Circuits 33, 12, 2092–2103.

Lopich, A. and Dudek, P. (2008). ASPA: Focal plane digital processor array with asynchronous processing capabilities. In IEEE International Symposium on Circuits and Systems, ISCAS 2008, 1592 – 1595.

Lopich, A. and Dudek, P. (2011), Asynchronous cellular logic network as a co-processor for a general purpose massively parallel array. International Journal of Circuit Theory Applications, 39, 9, 963–972.

Martin, D., Lee, H.-S., and Masaki, I. (1998), A mixed-signal array processor with early vision applications, IEEE Journal of Solid-State Circuits, 33, 3, 497 –502.

Massari, N. and Gottardi, M. (2007), A 100 dB dynamic-range CMOS vision sensor with programmable image processing and global feature extraction, IEEE J. Solid-State Circuits, vol. 42, no. 3, pp. 647–657, Mar.

Mead, C. and Mahowald, M. (1988), A silicon model of early visual processing, Neural Networks 1, 1, 91 – 97.

Miao, W., Lin, Q., Zhang, W., and Wu, N.-J.(2008) A programmable SIMD vision chip for real-time vision applications, IEEE J. Solid-State Cir- cuits, vol. 43, no. 6, pp. 1470–1479, Jun. 2008.

Moini, A. (2000). Vision chips. Kluwer Academic, Boston ; London

Morfu, S., Marquié, P., Nofiélé, B., and Ginhac, D. (2008), Nonlinear systems for image processing. In Advances in Imaging and Electron Physics, P. W. Hawkes, Ed., vol. 152. Elsevier, 79 – 151.

Motoyoshi, M. and Koyanagi, M. (2009), 3D-LSI technology for image sensor. Journal of Instrumentation 4, 03, P03009.

Murari, K. Etienne-Cummings, R. Thakor, N. and Cauwenberghs, G (2009), Which photodiode to use: A comparison of CMOS-compatible structures, IEEE Sensors Journal, 9, 7, 752 –760.

Nixon, R. H., Kemeny, S. E., Staller, C. O., & Fossum, E. R. (1995). 128 x 128 CMOS photodiode-type active pixel sensor with on-chip timing, control, and signal chain electronics. In M. M. Blouke (ed.), Proceedings of the SPIE charge-coupled devices and solid state optical sensors V, April (Vol. 2415, pp. 117–123).

Paillet, F., Mercier, D., and Bernard, T. (1998), Making the most of 15k-lambda-2 silicon area for a digital retina. In Proceedings of SPIE, vol. 3410, 158–167.

Paillet, F., Mercier, D., and Bernard, T. (1999), Second generation programmable artificial retina. In Proceedings of the twelfth annual IEEE International ASIC/SOC Conference, 304 –309.

Rekczky, C., Tahy, A., Végh, Z., and Roska, T. (1999), CNN-based spatio-temporal nonlinear filtering and endocardial boundary detection in echocardiography. International Journal of Circuit Theory and Applications 27, 1, 171–207.

Rinner, B. and Wolf, W.(2008), An introduction to distributed smart cameras, Proceedings of the IEEE 96, 10, 1565 –1575.

Rodriguez-Vazquez, A., Linan-Cembranon, G., Carranza, L., Roca-Moreno, E., Carmona-Galan, R., Jimenez-Garrido, F., Dominguez-Castro, R., and Meana, S. (2004), Ace16k: the third generation of mixed-signal SIMD-CNN ACE chips toward VSOCs. Circuits and Systems I: Regular Papers, IEEE Transactions on 51, 5, 851 – 863.

Roska, T. and Rodriguez-Vazquez, A. (2000), Review of CMOS implementations of the CNN universal machine-type visual microprocessors, In Proceedings of the IEEE International Symposium on Circuits and Systems, ISCAS 2000 Geneva, Vol. 2. 120 –123.

Sakakibara, M., Kawahito, S., Handoko, D., Nakamura, N.,Higashi, M., Mabuchi, K., and Sumi, H. (2005). A high-sensitivity CMOS image sensor with gain-adaptative column amplifiers, IEEE Journal of Solid-State Circuits, 40(5), 1147–1156.

Schrey, O., Huppertz, J., Filimonovic, G., Bussmann, A., Brockherde, W., and Hosticka, B. (2002), A 1 k x1 k high dynamic range CMOS image sensor with on-chip

programmable region-of-interest readout. IEEE Journal of Solid-State Circuits, 37, 7, 911 –915.

Seitz, P. (2000), Solid-state image sensing, Handbook of computer Vision and Applications, 1, 165–222.

Seitz, P. and Theuwissen, A. (2011), Single Photon Imaging, Springer Series in Optical Sciences. Springer-Verlag.

Shi, Y. and Lichman, S. (2006), Smart cameras: A review, CCTV Focus 37, 38–45.

Smith, S., Hurwitz, J., Torrie, M., Baxter, D., Holmes, A., Panaghiston, M., Henderson, R., Murrayn, A., Anderson, S., and Denyer, P. (1998), A single-chip 306x244-pixel CMOS NTSC video camera. In ISSCC Digest of technical papers. San Fransisco, CA, 170–171.

Stevanovic, N., Hillebrand, M., Hosticka, B. and Teuner, A. (2000), A CMOS image sensor for high-speed imaging," IEEE International Solid-State Circuits Conference, Digest of Technical Papers, pp. 104–105.

Stoppa, D., Somoni, A., Gonzo, L., Gottardi, M. and. Dalla Betta, G.F. (2002), Novel CMOS Image Sensor With a 132-dB Dynamic Range," IEEE Journal of Solid-State Circuits, vol. 37, no. 12, pp. 1846–1852.

Sumi, H. (2006), Low-noise imaging system with CMOS sensor for high-quality imaging. In IEDM '06 International Electron Devices Meeting, 2006, 1 –4.

Suntharalingam, V.; Berger, R.; Clark, S.; Knecht, J.; Messier, A.; Newcomb, K.; Rathman, D.; Slattery, R.; Soares, A.; Stevenson, C.; Warner, K.; Young, D.; Lin Ping Ang; Mansoorian, B.; Shaver, D., (2009), A 4-side tileable back illuminated

3D-integrated Mpixel CMOS image sensor, IEEE International Solid-State Circuits Conference - Digest of Technical Papers, ISSCC 2009, pp.38-39,39a.

Ohta, J. (2008), Smart CMOS images sensors and applications. CRC Press.

Wu, C.-Y. and Chiang, C.-T. (2004), A low-photocurrent CMOS retinal focal-plane sensor with a pseudo-bjt smoothing network and an adaptive current schmitt trigger for scanner applications, IEEE Sensors Journal, 4, 4, 510 – 518.

Wuu, S., Chien, H., Yaung, D., Tseng, C., Wang, C., Chang, C., & Hsaio Y. (2001). A high performance active pixel sensor with 0.18 lm CMOS color imager technology. In Electron devices meeting, 2001. IEDM technical digest. International (pp. 555–558).

F. Xiao, J. Farrell, P. Catrysse, B. Wandell (2009). Mobile Imaging: The Big Challenge of the Small Pixel. In Digital Photography V, Proceedings of the SPIE Electronic Imaging ’2009 Conference, Vol. 7250, January 2009

Xu, C., Zhang, W., Ki, W.-H., and Chan, M. (2002) A 1.0-v vdd CMOS active-pixel sensor with complementary pixel architecture and pulsewidth modulation fabricated with a 0.25-µm cmos process, IEEE Journal of Solid- State Circuits, 37, 12, 1853 – 1859.

Yadid-Pecht, O., and Belenky, A. (2003), In-pixel autoexposure CMOS APS, IEEE Journal of Solid-State Circuits, 38(8), 1425– 1428.

Yadid-Pecht, O., and Belenky, A. (2001), Autoscaling CMOS aps with customized increase of dynamic range. In IEEE International Solid-State Circuits Conference, Digest of Technical Papers, ISSCC (2001), 100 –101.

Yang, D., El Gamal, A., Fowler, B., and Tian, H.(1999), A 640 x 512 CMOS Image Sensor with Ultra Wide Dynamix Range Floating-Point Pixel- Level ADC, IEEE Journal of Solid-State Circuits, vol. 34, pp. 1821– 1834.

Yeh, S.. Lin, J., Hsieh, C., Yeh, K., and Li, C. (2011), A new CMOS image sensor readout structure for 3D integrated imagers, In IEEE Custom Integrated Circuits Conference (CICC), 1 –4.

Yoon, K., Kim, C., Lee, B., & Lee, D. (2002), Single-chip CMOS image sensor for mobile applications, IEEE Journal of Solid- State Circuits, December.

Zhang, M. and Bermak, A. (2007), Compressive Acquisition CMOS IMage Sensor: From the Algorithm to Hardware Implementation, IEEE Trans. Very Large Scale Integr. (VLSI) Syst, 18, 490–500.

Zhang, W., Fu, Q., Wu, N.J. (2011), A Programmable Vision Chip Based on Multiple Levels of Parallel Processors, IEEE Journal of Solid-State Circuits, vol.46, no.9, pp.2132-2147, Sept. 2011

Zhang X., Leomant S., Lau K.L., Bermak A. (2011), A Compact Digital Pixel Sensor (DPS) Using 2T-DRAM, Journal of Low Power Electronics and Applications, 1(1):77-96.