

CMOS Image Sensors are entering a new age

Pierre Fereyre, Gareth Powell

e2v, Avenue de Rochepleine, BP123, Saint-Egrève, F-38521, France.

I. INTRODUCTION

In the early 90's, it was suggested that Charge Coupled Devices (CCDs) were in the process of becoming extinct and considered to be 'technological dinosaurs' [1]. The prediction is not entirely wrong, if the announcement made by Sony in 2015 is taken as an example. Sony Corporation has officially announced the end of the mass production of CCD and has entered into the last buy order procedure. While this announcement was expected for many years, it has caused a stir within the professional imaging community [2]. It is interesting to note that many industrial or professional applications (where CMOS Image Sensor (CIS) technology was expected to dominate) are still based on CCD sensors. Which characteristics still make CCDs more attractive that are not available with a CIS? At the beginning the two technologies cohabited and CCDs quickly established themselves as a superior technology to meet stringent image quality requirements. CMOS technology was then in its infancy and limited by its inherent noise and the pixel complexity. Architectures were then essentially analog and the idea of integrating the image processing features (System On-Chip) was not yet considered. The shrinkage of the technological node according to Moore's Law has made this technology more and more competitive due to its rapid expansion in the 2000s. CIS are now in the race of continuously improving electro-optical performances, and in many aspects are often shown to be better than CCDs.

Considering the proposed metaphor of "the evolution of life" in the title of this white paper, the CIS can be compared to the emergence of mammals having withstood successive natural disasters; an evolutionary history that tells an epic story that covers 65 million years!

II. CCD AND CMOS: TWO DIFFERENT BRANCHES WITH A COMMON ORIGIN

With CCDs, photonic signals are converted into electron packets and are sequentially transferred to a common output structure, where the electric charge is converted to voltage. The signal is then buffered and carried off-chip. On a CCD, most of the functions take place on the camera's printed circuit board. If the application demands change, a designer can change the electronics without redesigning the imager. In a CMOS imager, the charge-to-voltage conversion takes place in each pixel. A CMOS imager converts charge to voltage at the pixel level, and most functions are integrated into the chip. It can be operated with a single power supply and has the ability of flexible readout with region-of-interest or windowing. CCDs are generally made in NMOS technology which is dedicated in performance with specifics like overlapping double poly-silicon, anti-blooming, metal shields and a specific starting material. CMOS are often consumer oriented, based on standard CMOS process technology for digital ICs with some adaptation for imaging (e.g. pinned photodiode). It is generally considered that the manufacturing of CMOS sensors is cheaper than CCD and that the performance is lower. This assertion is based on market volume consideration, but regarding other specialist business sectors the two technologies are equivalent or a CCD may work out as more economical [3]. As an example, the large majority of space programs are still based on CCD components not only optimizing the performance at process level on limited quantities and cost, but also to ensure the long-term supply – not always compatible with consumer demand! Similarly the science imaging market is also still served by a majority of high-end CCD based solutions, with new product developments in progress. Some dinosaurs evolved into birds and, in the main, they have great imaging capabilities ...

The system complexity is improved with CMOS as it generally embeds SOC (System-On-Chip) architecture like analog to digital conversion, correlated double sampling, clock generation, voltage regulators or features like image post-processing, that were formerly limited to the application system level design. Modern CIS are commonly made in 1P4M (1 poly, 4 metal layers) 180 nm down to currently 65 nm technology, allowing the design of pixels with very high conversion factor that can be combined with column gain amplification. The resulting photo-response and sensitivity of CMOS to light is generally better than CCD. The CCD benefit from significant noise advantages over CMOS imagers because of better substrate biasing stability with less on-chip circuitry, and almost no fixed-pattern-noise.

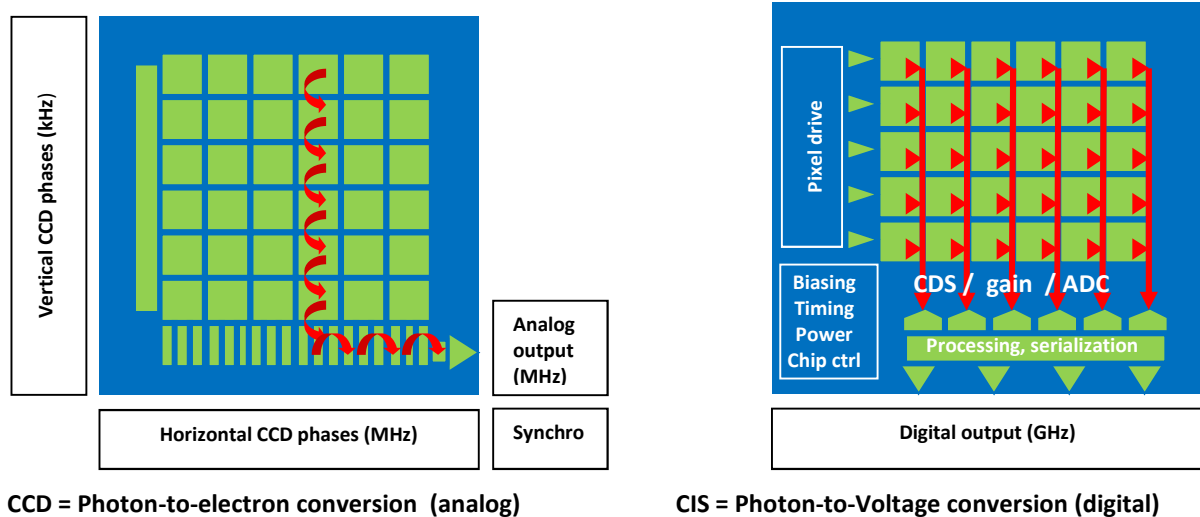


Figure 1 – CCD and CMOS architecture comparison

Characteristic	CCD	CMOS
Signal from pixel	Electron packet	Voltage
Signal from chip	Analog Voltage	Bits (digital)
Readout noise	low	Lower at equivalent frame rate
Fill factor	High	Moderate or low
Photo-Response	Moderate to high	Moderate to high
Sensitivity	High	Higher
Dynamic Range	High	Moderate to high
Uniformity	High	Slightly Lower
Power consumption	Moderate to high	Low to moderate
Shuttering	Fast, efficient	Fast, efficient
Speed	Moderate to High	Higher
Windowing	Limited	Multiple
Anti-blooming	High to none	High, always
Image Artefact	Smearing, charge transfer inefficiency	FPN, Motion (ERS), PLS
Biasing and Clocking	Multiple, higher voltage	Single, low-voltage
System Complexity	High	Low
Sensor Complexity	Low	High
Relative R&D cost	Lower	Lower or Higher depending on series

Table 1 – CCD-CMOS Characteristics comparison

On the other hand, CIS sampling frequencies being potentially lower, the bandwidth required for reading out the pixel can be reduced and the temporal noise is thereby lower. Shuttering exposes all pixels of the array at the same time. For CMOS, this approach consumes pixel area because it requires extra transistors in each pixel. Each pixel has an open-loop output amplifier, and the offset and gain of each amplifier fluctuates considerably because of wafer processing variations, making both dark and illuminated non-uniformities worse than those in CCD. CMOS imagers have lower power dissipation than equivalent CCD. The power dissipation of other circuits on the chip can be lower than that of a CCD using companion chips from an optimized analog system. Depending on the delivery volume, CMOS may be less expensive at the system level compared to CCD, when considering the cost of externally implementing related circuit functions. A summary of CCD and CMOS characteristics is presented in Table 1. Some features are to the benefit of one or the other technology without really segmenting the overall performance or cost. However, the point definitely discriminating CMOS is the flexibility of implementation with the system-on-chip (SOC) approach and the lower power consumption.

I. NOISE PERFORMANCE: A COMMON MISCONCEPTION

The bandwidth of the video imaging chain must be adjusted carefully to minimize the read noise level into the digitization stage. However, the bandwidth must be sufficient to avoid introducing artifacts in the image. This rule of thumb applies similarly to CCD and CMOS. The minimum threshold of bandwidth is determined by the time required for the sampled signal to settle on a level sufficiently close to the ideal signal. The induced error should be negligible compared to the Least Significant Bit (LSB) [4]. To determine the required bandwidth, the following criterion is applied:

$$f_c \approx N \cdot \ln(2) f_s \quad (1)$$

With f_c the amplification chain bandwidth, f_s the signal frequency and N the ADC resolution.

For example, if N=12, then the proper value is:

$$f_c \approx 8.3 f_s$$

The noise is the result of two contributions: 1/f flicker noise and thermal noise as shown Figure 2. Flicker noise is frequently present in nature and its spectral density is relevant for the fluctuations in the earth's rate of rotation, undersea currents, weather, climate phenomena, etc. A study of the flickering of a common candle has demonstrated that it fluctuates as 1/f. Regarding the MOS device and elements of the amplification chain, the flicker noise is the consequence of charge being trapped in the gate oxide. The traps are present because of defects generated by the technological process. The filling and emptying of these traps lead to fluctuations in current flowing in the transistor channel [6] also

referred to as "random telegraph noise" (RTS). Each individual trap can be modeled using a Lorentzian mathematical model that is suitable to describe a resonant behavior. The sum of Lorentzian (i.e the number of traps present at the surface of the MOSFET channel) results in a 1/f spectrum that fits very well with the actual noise spectral density. As a result, the 1/f magnitude inversely depends on the surface of the MOSFET channel area – not entirely intuitive.

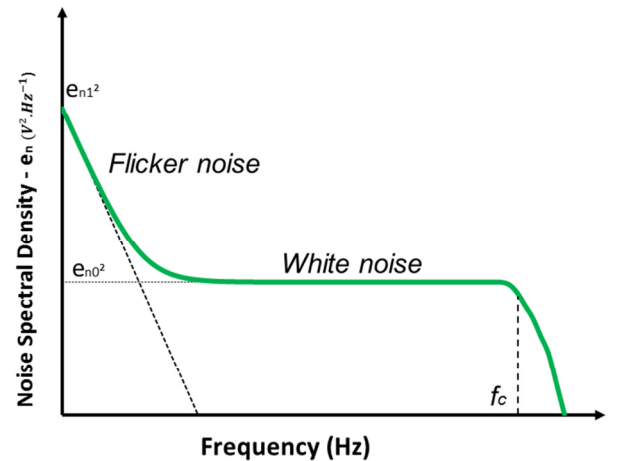


Figure 2 – Spectral noise density

To eliminate or at least reduce the variation of the common mode of the amplifier, the reset noise of the floating node, and the technological dispersions of the transistors in the case of CIS, the video channel usually integrates a Correlated Double Sampling (CDS) stage. This element transforms the video signal transfer function according to the following formula:

$$H_{CDS}(f) = 2 \cdot \sin\left(\pi \frac{f}{n f_s}\right) \quad (2)$$

With f_s the sampling frequency, n the CDS factor (typically $n=2$). As shown Figure 3 depending on the sampling rate, this filtering will eliminate more or less

the $1/f$ noise frequency component particularly as the sampling frequency f_s is high (in other words, the trap and release mechanism becomes slower than the CDS frequency). The combination of the H_{CDS} filtering and low pass filter of the amplification chain can be simplified as an equivalent band pass filter as depicted in Figure 3. The eqBP1 corresponds to a first order band pass filter. In order to make equivalent the integrated noise power with the H_{CDS} function, the noise spectral function of eqBP1 is divided by a factor of two. The eqBP2 is a frequency notch approximation of eqBP1. To make equivalent the integrated noise power, the lower and upper limits of the eqBP2 filter are respectively multiplied by $(\frac{\pi}{2})^{-1}$ and $\frac{\pi}{2}$.

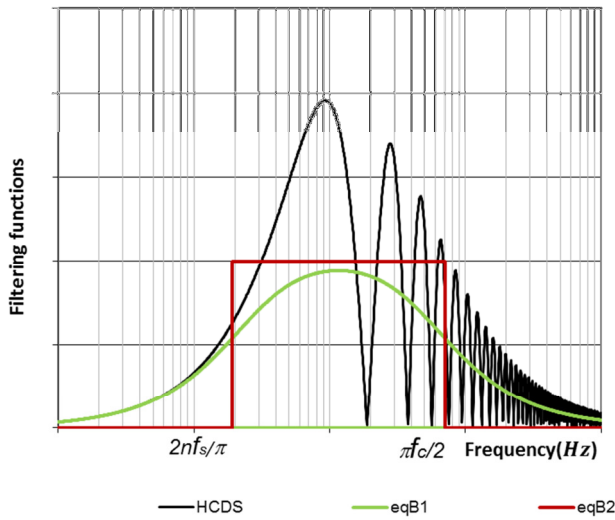


Figure 3 – Noise filtering functions

Considering the general case starting from the Figure 2 and Figure 3, the noise expression can be simply expressed according to the following formula:

$$e_T^2 = \int_0^\infty |e_n|^2 |H_{CDS}|^2 df \quad (3)$$

$$e_T^2 \approx 2 \int_{2nf_s/\pi}^{\pi f_c/2} \left(e_{n0}^2 + \frac{e_{n1}^2}{f} \right) df \quad (4)$$

Combining the equation (1) and (4) the total integrated readout noise is approximated as follows:

$$e_T \approx \sqrt{2} \left(\underbrace{N f_s e_{n0}^2}_{\text{White}} + \underbrace{Ln(N/n) e_{n1}^2}_{\text{flicker}} \right)^{\frac{1}{2}} \quad (5)$$

It is verified that the formulation matches pretty well with the numerical simulation.

The CCD read noise can be extremely low for applications such as astronomy, or science for which the image is read out at a very low frequency. The

system design includes electronics which frequency bandwidth is minimized to avoid integration of the time fluctuations of the signal. For these applications the $1/f$ component of noise is dominant. For high-speed video applications, noise is much higher and leads to a significant degradation of the signal to noise ratio. The theory is confirmed with measurements under actual conditions of noise performance of different CCD video cameras [5]. The CMOS image sensor has the advantage of a column parallel readout scheme (see Figure 1). The threshold readout frequency is therefore divided by the number of columns compared with CCD. Consequently, the readout noise of CIS is generally dominated by the $1/f$ contribution. It is leading to continuous effort of improvement of CMOS technology for imaging. It has been demonstrated recently that very good noise performance in the range of 1 e- and even below is achievable [7] [8].

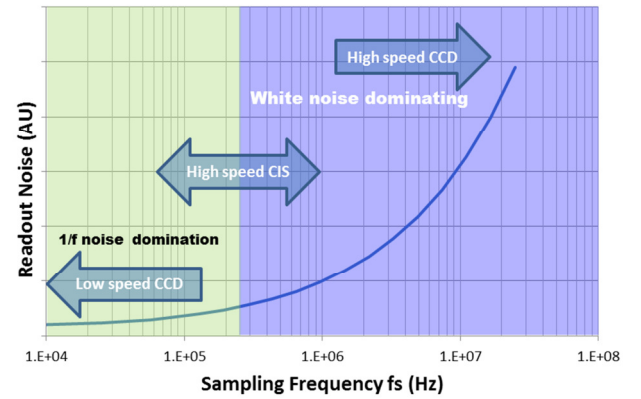


Figure 4 – Read out noise as a function of f_s

II. MTF AND QE: THE PILLARS OF IMAGE QUALITY

Quantum efficiency (QE) takes a direct part in the image sensor electro optical performance because any loss of photon-to-electron conversion efficiency leads directly to a decrease in the Signal-To-Noise-Ratio (SNR). The impact is twofold since the QE comes to the numerator of SNR (Signal) and also the denominator (noise) which is true if the shot noise (square root of the signal) is the dominant contributor. On this point, CCD and CMOS can be considered at the same level, but CCD has historically benefited from many years of technological process iterations for QE enhancement. It is a relatively recent advance in the CIS domain. Based on the physical properties of silicon where longer wavelengths penetrate deeper into the photo-

sensitive conversion zone, thick epitaxial material is used to increase the QE in the upper red and NIR wavelengths. Following the Beer-Lambert law, the absorbed energy has an exponential dependence on the material thickness. CCD dedicated to high-end applications benefit from these technologies including the use of thick silicon materials and back side illumination (BSI) to recover a high broadband QE and sensitivity in the near infrared (NIR).

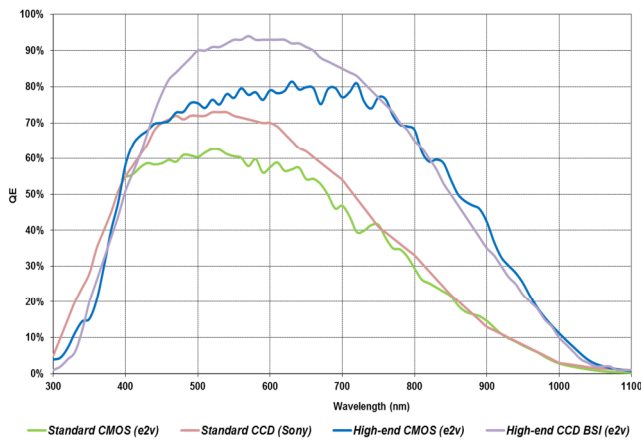


Figure 5 – QE benchmark

The interline transfer CCD (ITCCD) is based on a specific manufacturing process for implementing the said “vertical overflow drain” (VOD) or “vertical anti-blooming” (VAB) that has been developed in the early 1980’s [12]. The performance of the VAB is very good but has the drawback of cutting the response in the red, rejecting the NIR part of the spectrum.

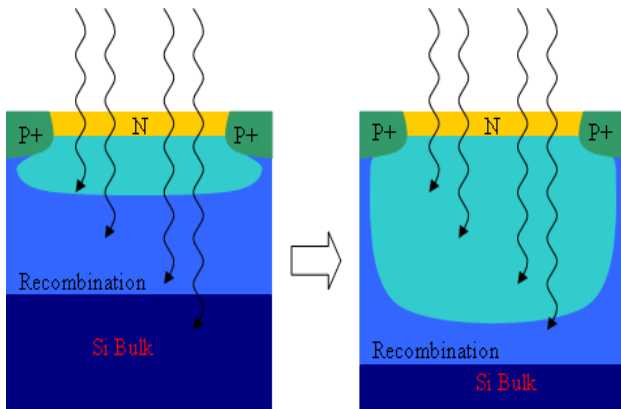


Figure 6 – Deep depletion approach

For this reason they cannot benefit from BSI. As they are using horizontal anti-blooming (HAB) the high-

end CCD doesn’t have this limitation. This is also the case for CMOS image sensors. Thin detection layers do not have the disadvantage of crosstalk because the charges are prevented from diffusing from pixel to pixel. As a result, the spatial resolution or modulated transfer function (MTF) of ITCCD and standard CIS is good. To gain sensitivity in the NIR domain necessitates a significant increase of material thickness. However, thick material usually leads to MTF degradation through increased electro-optical crosstalk. Image quality is a combination of MTF and QE (so-called Detective Quantum Efficiency) where consideration must be given to both the spatial and temporal domains. Deep depletion photodiodes with adapted silicon doping methods are used to recover MTF as shown in Figure 6. The CIS are generally made on technologies inspired from those used for integrated circuits (particularly DRAM/memory processes) and generally do not use the specific recipe described above. However, recent publications have shown implementation of specific processes, performing excellent QE improvement and relatively close to those obtained with high-end CCD as depicted in Figure 5 [9][10]. Latest CMOS technology trends represent a quantum leap forward. Techniques such as light guides, deep trench isolation (DTI), buried μ lens, and even stacked die containing pixel transistor underneath the photosensitive area now being employed.

III. SOME INHERENT IMPERFECTIONS

The “pinned photodiode” (PPD) or the “hole accumulation diode” (HAD) was initially developed to remove the lag and perform a full charge transfer from the photodiode to the ITCCD register [12]. One of the major developments of CMOS imagers is the adoption of the ITCCD photodiode structure in the early 2000’s [11], represented in Figure 7. In CMOS, pixel architectures are frequently referred to by the number of transistors per pixel. Most CMOS imagers tend to use an electronic rolling shutter, which is the most beneficial in integration and can be realized with as little as three transistors (3T). While commendable for its simplicity, the 3T pixel architecture suffers from a higher pixel generated temporal noise from kT/C (or thermal) noise in the circuit, which cannot be simply removed.

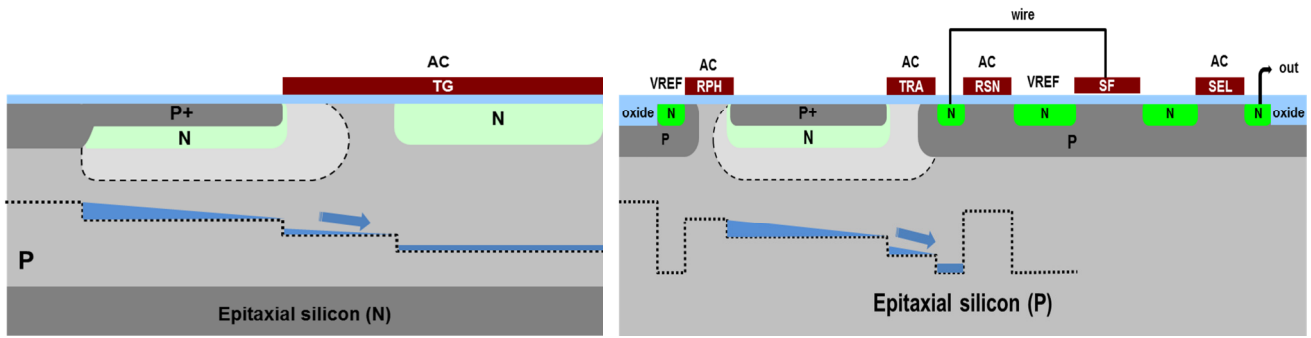


Figure 7 – ITCCD and 5T CMOS pixels side by side

The pinned photodiode was initially introduced in CIS to remove the noise from the reset of the floating diffusion and has led to the introduction of the four transistors pixel (4T). The 4T architecture performs a Correlated Double Sampling (CDS) to remove the reset temporal noise. This architecture also enables transistor sharing schemes between pixels to reduce the number of effective transistors per pixel to less than two. Evidently, fewer transistors in the pixel free up more area for the photosensitive part or fill factor to more directly couple light into the pixel. However ERS introduces image distortion when capturing video or images containing fast motion as shown in Figure 8. The PPD was exploited in a second stage to perform the global shutter (GS) capture. It removes the ERS artifacts and furthermore the temporal noise, as well as dark current and fixed pattern noise. The fifth transistor adjacent to the PPD (5T) is used to drain the excess charges and also to adjust the integration time in overlap mode (readout during integration).

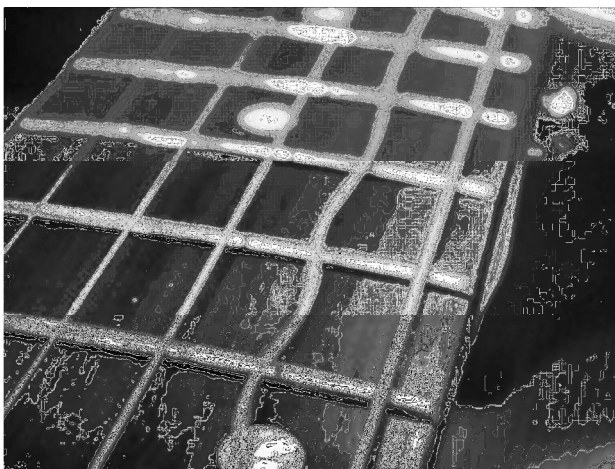


Figure 8 –Image artefacts: CMOS ERS distortion

The GS mode is used with ITCCD but is in some case sensitive to smearing effects.

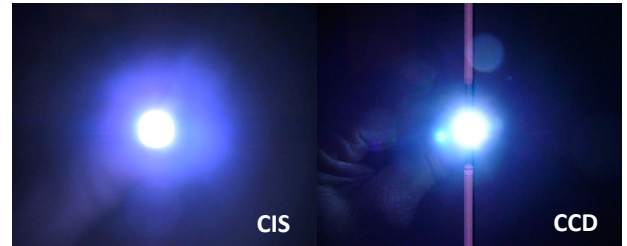


Figure 9 –Image artefacts: CCD smearing

The smearing appears during charges transfer and produces vertical stripes in the image as shown in Figure 9. This defect is particularly visible in high contrast images and should not be confused with the blooming that potentially produces a similar artefact. In order to reduce this problem, the frame-interline-transfer (FIT) CCD architecture is generally implemented. The FIT has also the advantage of higher video rate. The CMOS equivalent parameter of smearing is the Global Shutter Efficiency (GSE), sometimes referred to as parasitic light sensitivity (PLS) that corresponds to the ratio of sensitivity of the sensing node to the photodiode. GSE is generally between -88dB to -100dB on ITCCD [13], and in the range of -74dB to -120dB on CMOS and even -160dB using 3D stacked architectures [14]. The use of advanced and customized pixel micro lens (e.g. zero-gap) makes a significant difference, in the sensitivity over wavelength response and limits fill factor loss due to the presence of transistors in the CMOS pixel. It is also a contributor to improvement of GSE performance.

IV. CMOS IMAGING TECHNOLOGY TO COME

There is a field where CCD technology is particularly suitable, which is for **time delay integration (TDI)**. The implementation of TDI (the integration and the summation of electrons synchronously with the scanning of the scene) is relatively straightforward with a charge transfer device. This technique contributes firstly to maximize the SNR and secondly

to preserve a good image definition (MTF). In recent years several attempts to reproduce the signal summation either in the analog domain (voltage) [18] or in the digital domain have opened the way forward for CMOS TDI. For space earth observation or for machine vision, CCD time delay integration architecture is still in demand for its low noise and high sensitivity performances. However, the most promising results were obtained, by considering the best of both technologies and the combination of charge transfer registers and column wise ADC converters on the basis of a CMOS process [17].

Despite the progress in this area, the sensitivity of CMOS image sensors are still limited by the read noise for extremely low light applications (i.e. few tens of μlux). The EMCCD [15] that operates with electronic multiplication has shown great potential for noise reduction, and made it of great interest to the scientific market. In general, CCD have been gradually replaced by CMOS imagers and in the same way, EMCCD can potentially move to **electron multiplying CMOS** (EMCMOS) [20]. Like EMCCD, this technology is planned to be used to improve image quality at extremely low light levels for science or surveillance applications. CMOS technology enables lighter and smarter systems, lower power consumption, and are less expensive for large scale volumes (the so-called SWAP-C approach). The principle of electron multiplication is to apply a gain to the signal before any noise addition by the readout chain. The noise is virtually divided by this gain and SNR is improved. As a result of the CCD principle, the signal is transferred in the form of electron packets and the multiplication is applied commonly to each pixel before reading out. For CMOS, the signal is in the voltage domain, which means that the multiplication must be applied prior to transfer to the floating node and before adding the noise from the source follower transistor [16].

Growing in popularity, three-dimensional (3D) imaging where depth information is measured, the **Time-Of-Flight** (TOF) is a convenient technique. In principle a source of pulsed artificial light, located on the sensor plane is emitted, and the return of the reflected wave is used in a correlated function to extract the distance. This technique was initially described in 1995 as a "lock-in imager" CCD [21]. The first attempt with CMOS was inspired by the CCD pixel [22]. A second method uses Current Assisted Photonic Demodulators (CAPD). Both approaches have led to mass production of industrial 3D sensors for a variety of applications including people

counting, safety control, metrology, industrial robotics, gesture recognition and automotive Advanced Driver Assistance Systems (ADAS). This is a typical example of a successful transition from a concept developed on CCD technology and perfected on CMOS for deployment on a high-volume industrial scale.

CMOS technology deployment has branched out into new fields of application. As an example, the **Single-Photon Avalanche Diode** (SPAD) is a solid state solution originally developed to replace photomultiplier tubes (PMT). This is similar to the replacement in the 80's of vidicon tubes in professional cameras by CCD. SPAD is basically a p-n junction biased in reverse voltage above the breakdown mode in the so-called Geiger mode. The structure is very unstable and any energetic disturbance leads to the avalanche effect. This property is exploited to perform single-photon detection. The avalanche is turned off using the principle of passive quenching by implementing a simple resistive component between the SPAD and the supply voltage, or active quenching when using an embedded MOSFET channel. This produces a digital signal representation of a quantum event. In principle the SPAD is a simple structure-based CMOS technology and doesn't need a very specific process as required for image sensors. Because it requires the use of complex circuitry, the operation of a SPAD array is much more complicated. The SPAD triggering and event counting is by definition asynchronous, like the arrival of photons. The choice of CMOS technology is judicious. For example it is possible to proceed in a very quick scan of the pixel array to determine those that have transitioned. The assembly of these frames produces a video sequence [23].

V. CONCLUSION

Early publications announcing the end of CCD were something of a prophesy [1], even if the transition is taking considerably longer than predicted. Furthermore the variety and inventiveness of pixel structures developed for CMOS imagers have surpassed the imagination. They have been made achievable with the transistors etchings downsize and evolution of the CMOS fabrication technology, now completely adjusted for CIS production. The major industrial imaging makers still compete on price but continuously on electro-optical performance. Users no longer limit themselves to take a picture but to

capture the best times of their lives, they expect a perfect picture in all light conditions. The industrial applications also benefit from these advances accomplished for a more general interest. Vision systems are increasingly based on imagers that follow the trend of consumer demand including the shrink of pixels. Speed, for example, is also an important economic factor since it maximizes the through-put of expensive production machines and automated

processes/inspection. New applications are pushing the sensors towards their extreme capabilities and do not tolerate any additional noise in the image, enabling single photon imaging. Beyond simple image capture and display, 3D augmented reality uses all the potential of CMOS technology to bring a different perception of space. CMOS sensors have definitely evolved and adapted to their environment just like a dominant species.

REFERENCES

- [1] Active pixel sensors: Are CCD dinosaurs? ER Fossum IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology, 2-14A.
- [2] <http://www.vision-systems.com/articles/2015/03/sony-rumored-to-discontinue-production-of-ccd-sensors.html>
- [3] CCD vs. CMOS, Dave Litwiller, Photonics Spectra, 2001
- [4] Determination of the optimal electrical bandwidth in CCD- and CMOS-based image detector applications, Robert H. Philbrick, SPIE 5499, Optical and Infrared Detectors for Astronomy, 2004
- [5] CMOS vs. CCD: Changing Technology to Suit HDTV Broadcast, Lester J. Kozlowski, 2003
- [6] Fundamental performance differences between CMOS and CCD imagers: Part 1, James Janesick et al., SPIE 6276, High Energy, Optical, and Infrared Detectors for Astronomy II, 62760M, 2006
- [7] A 0.7 e-rms Temporal Readout Noise CMOS Image Sensor for Low Light Level Imaging, Y. Chen et al., IEEE International Solid-State Circuits Conference (ISSCC), 2012.
- [8] http://www.caeleste.com/caeleste_publications/2011CNES/20111207_CNES_caeleste_presentation.pdf
- [9] L²CMOS Image Sensor for Low Light Vision, Pierre Fereyre et al., International Image Sensor Workshop, 2011
- [10] Night Vision CMOS Image Sensors Pixel for SubmilliLux Light Conditions Amos Fenigstein, International Image Sensor Workshop, 2015
- [11] A Review of the Pinned Photodiode for CCD and CMOS Image Sensors Eric R. Fossum, et al., IEEE Journal Of The Electron Devices Society, Vol. 2, no. 3, may 2014
- [12] No image lag photodiode structure in the interline CCD image sensor, N Teranishi et al., Electron Devices Meeting, 1982 International (Volume:28), 1982
- [13] http://www.sony.net/Products/SC-HP/new_pro/may_2014/icx825_e.html
- [14] A 3D stacked CMOS image sensor with 16Mpixel global-shutter mode and 2Mpixel 10000fps mode using 4 million interconnections, Symposium on VLSI Circuits (VLSI Circuits), T. Kondo et al., pages C90 - C91, 2015
- [15] M. S. Robbins and B. J. Hadwen, "The noise performance of electron multiplying charge-coupled devices," IEEE Transactions on Electron Devices, vol. 50, no. 5, pp. 1227–1232, May 2003
- [16] Electron Multiplying Device Made on a 180 nm Standard CMOS Imaging Technology, Pierre Fereyre et al., International Image Sensor Workshop, June 2015
- [17] First Measurements of True Charge Transfer TDI (Time Delay Integration) Using a Standard CMOS Technology, F. Mayer et al., International Conference on Space Optics, 2012
- [18] CMOS long linear array for space application G. Lepage, Proc. SPIE 6068, Sensors, Cameras, and Systems for Scientific/Industrial Applications VII, 606807, 2006)
- [19] Time-Delay-Integration Architectures in CMOS Image Sensors G. Lepage et al., IEEE Transactions On Electron Devices, vol. 56, no. 11, November 2009
- [20] R. Shimizu and Al., "A Charge-Multiplication CMOS Image Sensor Suitable for Low-Light-Level Imaging" IEEE Journal of Solid-State Circuits, vol. 44, no. 12, pp. 3603-3608 December 2009
- [21] The lock-in CCD-two-dimensional synchronous detection of light, T. Spirig, P. Seitz et al., IEEE Journal of Quantum Electronics, Vol. 31, Iss. 9, p. 1705 – 1708, Sep 1995
- [22] Demodulation pixels in CCD and CMOS technologies for time-of-flight ranging Robert Lange et al., Proc. SPIE 3965, Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications, 177 (May 15, 2000)
- [23] 320x240 Oversampled Digital Single Photon Counting Image Sensor N. AW. Dutton, VLSI Circuits Digest of Technical Papers, 2014