NEXT GENERATION CMOS IMAGER

FOR BROADCAST CAMERAS

Peter G.M. Centen

Grass Valley, the Netherlands

ABSTRACT (150WORDS)

More than 20 years ago Tubes where replaced by CCD imagers as the image capture element in a broadcast camera. After maturing for many years, present CCDs far exceed the performance of Tubes. With the arrival of 1080p50/60 the transition from CCDs towards CMOS imagers is inevitable.

Laws of physics dictate that each packet of photons has shot noise. As such it will pose a physical limit to the performance of a camera, especially in the arena of interchanging frames-per-second, pixel dimensions and lens f-number. This aspect raises some interesting challenges with regard to 1080p50 and beyond.

Recently a full HDTV CMOS imager was developed for use in broadcast cameras, competing in performance with CCDs.

This paper discusses underlying issues, like resolution, pixel-size, noise, sensitivity and frame-rate in relation to CCDs and CMOS imagers for broadcast use.

INTRODUCTION

Without any doubt the performance of a camera starts with the imagers applied.
In the high-end imager-application everybody rides the edge of what is technically possible because all manufacturers are in a prisoner's dilemma. If we don't do it others will.
Looking back in history we discover that the MOS-imager was invented in 1967 by Wecker and Noble, 3 years before the invention of the CCD imager by Boyle and Smith, in 1970. MOS imagers had their early advent in the late 60's taken over in a fast ramp-up by CCDs in the early 70's, [1]. After 1995, when the CCDs reached maturity, visible CMOS started its ramp-up.
The intriguing question is: "*why did it take so long for CMOS imagers to enter the market even when they where conceived before the CCDs*".
The buzz word is Lithographic Feature Size.
In general the design of a CCD-pixel is simpler than a CMOS-pixel. A CMOS pixel contains several active elements and a CCD several poly gates.
E.G. in a CMOS imager with a 3T-pixel, 3-transistors and 1-photodiode per pixel are needed and the silicon real estate needs to be divided between those.

A first condition for a CMOS-imager to be viable is to achieve the small pixel dimensions.

THE CMOS IMAGER

Many people believe that CMOS imaging is cheap and of low quality. But that is a self fulfilling prophecy. Applying the solutions, known from CCDs, to the CMOS imager demands additional process steps and masks. It results in an imager that is more expensive, and as a consequence not easily executed in a mass-market.
The statement: the CMOS imager is expensive and of high quality is equally valid.

When designing a CCD or a CMOS imager the parameters that needs to be at the technological-edge are always the same. These parameters are:
-Temporal noise or read noise,
-Sensitivity, or quantum efficiency and fill-factor,
-Qmax or overexposure margin; together with the read noise it defines dynamic range,
-Dark current and Fixed Pattern Noise
-Random non-uniformity or gain differences between pixels.
And to a lesser degree al the yield related issues like scratches, beauty, leaking pixels, deviating pixels.

CCDs have a long history of improvement in all of these parameters. Often improvements in one of the parameters are reached while maintaining equal or better performance in the other ones. As such CCD performance represent a benchmark and CMOS-imagers needs to meet those.
The performance of a camera is determined by the imager it employs, likewise the performance of the imager is based in the pixel performance.

Pixel size 20 times feature size

In an already old, but still very valid paper [2], it is argued that the ratio between lithographic feature size and pixel size is about 20. It makes it clear that for a

HDTV CMOS imager, with 5µm pixels, to be viable one needs a CMOS (imaging) process with a feature size of 0.25µm or less. After a CMOS technology node becomes available it often takes many years of technological improvements before the imaging options are available too. And that point is reached only recently.

The first 2/3-inch HDTV CMOS imager was published [3] in 2001. Underscoring the conclusions drawn, but also answering the question of why CMOS imagers to become feasible in broadcast recently.

In discussing and comparing CMOS and CCD imagers an opening question could be:
Is it possible to develop a CMOS imager constraint by the condition that:
it should be a 1920x1080i50 imager with 11 mm image diagonal with which,
a 3-imager camera can be developed, has a signal-to-noise of 54-60 dB in Y in 30MHz bandwidth at 2000 lux, f/8-f/11, 90% and 3200 K. Has an overexposure margin of 500% and can be used in 1920x1080i50 or even 1280x720p50 all with competitive image quality?
CCDs did just that, you only have to look at all the cameras available in the market today.
In the following paragraphs this issue will be addressed.

The Image Diagonal

In the consumer market there is a tendency to regard, more pixels as better. But more pixel in even less silicon area results in small and damn small pixels. The present trend is at about 2.5µm squared and even exotic pixels of 0.9µm are designed!! Getting light in these pixels is a real challenge.

In broadcast the image diagonals was for a long time 2/3-inch and to a lesser extend ½-inch. The diagonals being downgraded in recent years to 1/3-inch for Pro/AV like applications.

The basic problem is that of sensitivity [4]. From 2/3-inch down to ½'-inch the pixel area shrinks with a factor two. The 1/3-inch imager has an even four times smaller pixel area. So for a given resolution (read pixel count) the ½-inch has a 1-fstop, lesser sensitivity and the 1/3-inch lags with 2-fstops, behind in sensitivity. This gets even worse when one compares a single-chip camera with the 3-chip camera. Opting for a single chip camera with the same resolution then due to the Bayer pattern, the same area needs to be divided between 2-green and 1-blue and 1-red pixel, and one loses an additional 2-fstop in sensitivity too. So a 3-imager 2/3-inch camera has 4-fstops more sensitivity then a 1-chip 1/3-inch camera with the same resolution.
NOTE: that is why those small format digital cameras are so (very) noisy at low light conditions. Not to mention the problems one runs into with the optics where the demand to resolve higher number of lp/mm is 2 times higher for the 1/3-inch camera compared with the 2/3-inch, [4].

The Pixel

A CCD pixel is either of the Frame-Transfer [5] type where the pixel acts as storage, light sensitive and transport element at the same time. The other type is the Inter-Line/Frame Inter-Line [5] where storage, transport and light capture are separated. Both types have their merits. The Frame Transfer with the inherent flexibility in scanning that became known as Dynamic Pixel Management [6, 7] and the Inter-Line for its shutter-less operation.

A CMOS pixel consists of a photodiode and several transistors. The pixel-type is often expressed in the number of transistors per pixel. E.G the 3-transistor pixel is denoted as 3T-pixel, 4-transistor as 4T-pixel and the 1.5 transistor-per-pixel as 1.5T-shared pixel. In contrast to the pixel-number race there is no race in the number-of-transistors-per-pixel. The transistors are used for additional functionality, not found in CCDs. A striking example is the enabling of high dynamic range through a knee-function in each pixel [8]. This is not possible with a CCD.

Sensitivity

Due to the area taken up by the transistors the fill-factor of the pixel, defined as the area of the photodiode divided by the pixel area is much less than 100%, values of 50% are not uncommon. Without any other counter-measures a small fill-factor results in a low sensitivity.

Solving the problem of getting light in the pixel/photodiode is an issue that has been addressed in CCDs already a long time ago. The solution is in the application of µ-lenses on each pixel [9]. It focuses light on the photodiode and refocuses light that normally would have fallen on the non-sensitive part. It increases the, effective, fill-factor towards 80%-90%. Other approaches to enhance sensitivity are backside illumination [10] and light-piping. [11]

Both CMOS and CCD arrive at comparable photon efficiency numbers.

Read Noise

Every CMOS pixel has a capacitive detection node, resembling the floating diffusion of the CCD. At the detection node charge is converted into a voltage (swing). The capacitance needs to be reset (discharged) every time a charge packet, containing video information, is sensed to prepare it for the next charge packet. When the capacitor is reset and is ready to accept the next charge packet reset noise, known as kTC noise, is generated. It remains at the capacitor or detection node until the next reset operation. It is a kind of random offset. This noise component is very large. Suppression of it is mandatory to obtain ample performance. The solution again found in the CCD [12] and known as correlated double sampling abbreviated as CDS. Correlated because the processing consists of sampling the black value after the capacitance has been reset and sampling again when charge has accumulated on the capacitor and subtracting both values. The values belong to one and the same reset interval and hence are correlated. This type of noise improvement can also be applied to CMOS imagers. An other solution to reduce the kTC-noise is known as soft reset [3]. Unfortunately it suffers from LAG related issues.

An important characteristic of CMOS imagers is related to the frame rate.

Higher frame rate means higher bandwidth. In a CCD the noise determining element is the on-chip amplifier where all the pixels go through. The external signal processing averages the signal and the read noise. The averaging time is proportional with clock frequency and for HDTV on a nano-second time scale. Comparing 1080p50 cameras with 1080i50, the averaging bandwidth increases with 2x and the sampled noise increases with sqrt(2).

In the CMOS imager case every column has its amplifier and signal processor. Signals are processed on a micro-second time scale. In 1080p50 the pixel-to-pixel scanning has to be done in doubled speed just as the A-to-D conversion. As such the noise will stay rather constant independent of whether the imager is used in 1x, 2x or 3x. or even 6x.
In theory the noise of the CMOS imager can be much smaller than that of a CCD.

VISIBILITY OF NOISE

THE main aspect of image quality is the visibility of noise in the images. Today HDTV cameras are struggling to obtain noise free images at f-numbers we know from SD cameras, (f/11). Neglecting Fixed Pattern Noise a distinction can be made between two aspects of the noise contribution. One aspect is the noise in the dark parts of the images; the other is the photon shot noise in the exposed parts.
Through the laws of physics, photon shot noise is proportional to the square root of the linear output signal.
More precise the photon shot noise is equal to the square root of the number of electrons (n) generated in the photodiode. The shot noise is therefore also proportional with the square root of the linear signal. Let the output signal be given as: $V_{out} = gain * n$ then the noise for that signal level is: $U_n = gain * \sqrt{N_{ro}{}^2 + n}$.
Doubling illumination doubles the linear output signal and increases the shot noise with 3dB.

The noise in each video frame is different. There is no correlation between noise from consecutive frames. The perception of noise and hence it perceived intensity is greatly reduced by the average action of the “eye” [13]. An image can look noise free on a monitor but the moment one focuses on one frame, e.g. during editing, the noise pop’s up. *A rule of thumb is: temporal camera noise has a 10dB lower stimulus value than frozen noise.*

Even today there is still room for improvement for the noise in dark areas. It is determined by the read noise of the pixel and the readout path (N_{ro}). The noise in the exposed parts, the shot noise is fully determined by the f-number chosen for the 0dB camera setting and the effective quantum/conversion efficiency. The present technology allows an overall conversion efficiency of some 60% for the green channel. This means that 60% of the photons that fall on the pixel are converted into usable signal electrons. *The physical limit for the photon-to-electron conversion, in silicon, is 100%.*

The visibility of noise in an image not only depends on the noise in the dark areas. Shot noise in the exposed parts is also important. Luckily the eye is less sensitive to noise when it’s superimposed on an illuminated background (Weber-Fechner law).

Some Broadcast Numbers

Given a full HDTV broadcast camera, with three 2/3-inch imagers and adjusted to a (0dB) sensitivity of 2000 lux; f/8; 3200K and 89.9% scene reflection.
At 50 frames/second one arrives at the number of photons per pixel as per Table 1. Also the numbers to reach 100% video in +12dB are given for the Red, Green and Blue channel.

Assuming an overall conversion efficiency for the photon-to-electron conversion of 60%, $QE * Aperture * (1 - reflection) * \mu lens\_gain$, then the charge accumulated in the green pixel under nominal illumination conditions is G=4400 electrons. The low number of signal electrons for blue of 1400 electrons is caused by the color temperature of 3200K. Hence the reason for the ever noisy Blue channel when the camera is used at 3200K.

2000lux; 89.9%; 3200K; f/8 2/3"; 1920x1080p50			
	Red	**Green**	**Blue**
0 dB	8400 photons 3400 electrons	7400 photons 4400 electrons	2400 photons 1400 electrons
+ 12 dB	2100 photons 840 electrons	1850 photons 1100 electrons	600 photons 360 electrons

Table 1 -*The number of electrons and photons in a pixel for 100% video.*

Bear in mind though that they are for a 2/3-inch imager! Imagine how small these numbers become for a ½-inch or a 1/3-inch imager.

To get and impression of the photon shot noise involved, in Table 2 the shot noise at 100% exposure is given. Note that these are f/8 numbers and the SNR is already low at 36dB in green for 100% output level.

2000lux; 89.9%; 3200K; f/8 2/3"; 1920x1080p50			
0dB	**Red**	**Green**	**Blue**
Electrons	3400 electrons	4400 electrons	1400 electrons
Shotnoise	58 electrons	66 electrons	37 electrons
Shotnoise SNR	35 dB	36 dB	31 dB

Table 2 - *The number of electrons and shotnoise at 100% video.*

The Photon Shot Noise curve and 1920x1080p50

One can plot the noise as function of output signal [14]. For low exposure values the curve level flattens out asymptotically to the read noise. At high output levels photon shot noise is dominating and follows a square root law. These curves on a log-log scale are in the asymptotical case, two straight lines of which one is a constant and the other one has an angle of 0.5. From this curve several electrical properties can be determined [14, 15]

In the figures 1, 2, 3 two reference photon shot noise curves are drawn, Ref. SD and Ref. HD. They are a pragmatic way of defining acceptable and not acceptable and a gray zone in-between, where it is unclear which way to go.

In the SD era a status was reached for what was regarded as noise free images. The noise in black defined as -60dB and a 0dB sensitivity of f/11 for 2/3", (Green solid line with label Ref. SD).

The second photon shot noise curve is for the 2/3-inch HDTV camera (Red solid line with label Ref. HDTV) scanned in 1080i with 54dB in black and a sensitivity of f/11. They where regarded at the edge with respect to the noise impression, or slightly over the edge. The blue solid curve denoted Limit is reached when there is only photon shot noise with a conversion efficiency of 100%. One cannot pass this line without violating the laws of physics.

In the figures 1, 2, 3 the X-axis shows the imager output relative to 100% in 0dB. The Y axis shows the noise in Y (luminance with γ=1) as a function of the relative output. For low output levels the noise is dominated by the electronics read noise and the curve flattens out. *Note: The SNR in Y is mainly determined by the noise in green and: SNY=SNG + 1.5 a 2.0 dB*

The photon shot noise curves given in the figures 1, 2, 3 are based on the intrinsic noise performance of the imager itself. One can always improve these curves through the use of noise reducers. From these plots one can estimate the amount of noise reduction needed to arrive at acceptable noise levels. But noise reducers never go without side effects. Hence the focus on the intrinsic noise performance.

Figure 1: the camera output format is 1080i50 and the input is 1080p50: The noise is plotted as a function of output level. There are 2 solid curves denoted with ref SD and the other with ref HD. The dashed-dot curve is for 1080i50 derived from 1080p50 at 2000 lux, 89.9%, 3200K, f/10. The dashed curve depicts the 1080i50 derived from 1080p300. To stay in line with present day specification the curves are plotted for f/10 and for a conversion efficiency QE=60%.
To arrive at the same shot noise level as for SD the 0dB camera setting should be at f/5.7. And for the read noise to reach -60dB the f-number is f/7.2.
Note: For 1080p300 to be usable for down conversion to 1080p50 the read noise needs to decrease considerably. After down conversion there is no difference between the photon shot noise of a 1080p50 or a 1080p50 from 1080p300.

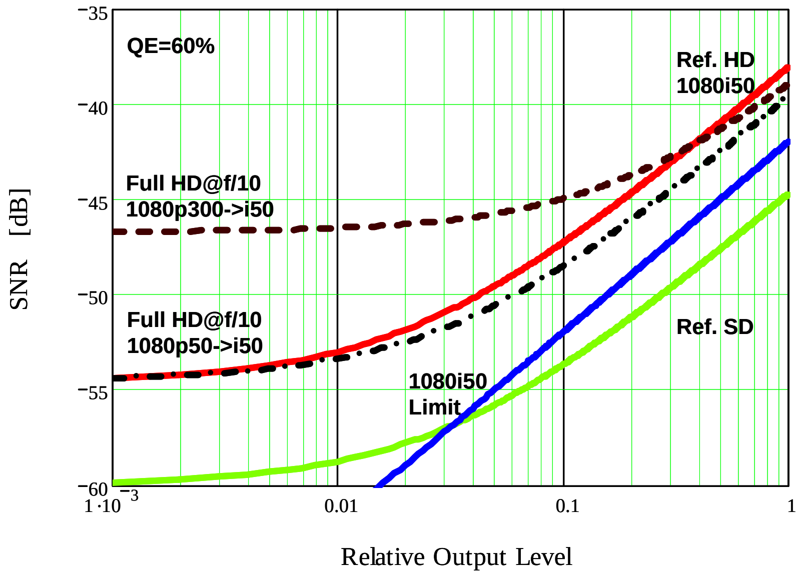

Figure 1 - Noise in Y as a function of output level for 1080i50 output format derived from different input formats at f/10. As a reference the noise is given for the interlace scanned SD and HD cameras. QE is assumed at 60%.

The transition into 1080p50 as the output format makes the noise performance even worse. In **Figure 2** the effect of conversion efficiency is depicted. The solid curves are again the SD and HD reference curves and the dashed curve plots the noise for 1080p50 at f/10. The curve lay's clearly outside the usable region. Even if one would achieve a conversion efficiency of 100% (dashed-dot curve) one would still be on the edge of noisiness.

Given that theoretical maximum of 100%: a 1080p50 camera with 0dB at f/5.3 would have the same shot noise as the SD reference curve and for f/7.8 the read noise would be at -60dB.

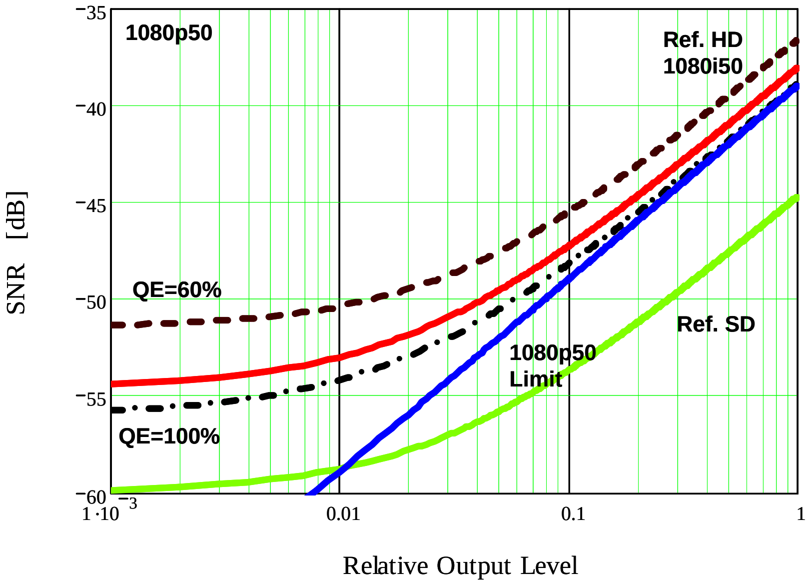

Figure 2 - Noise in Y as a function of output level for 1080p50 at f/10. QE=100% is the theoretical maximum.

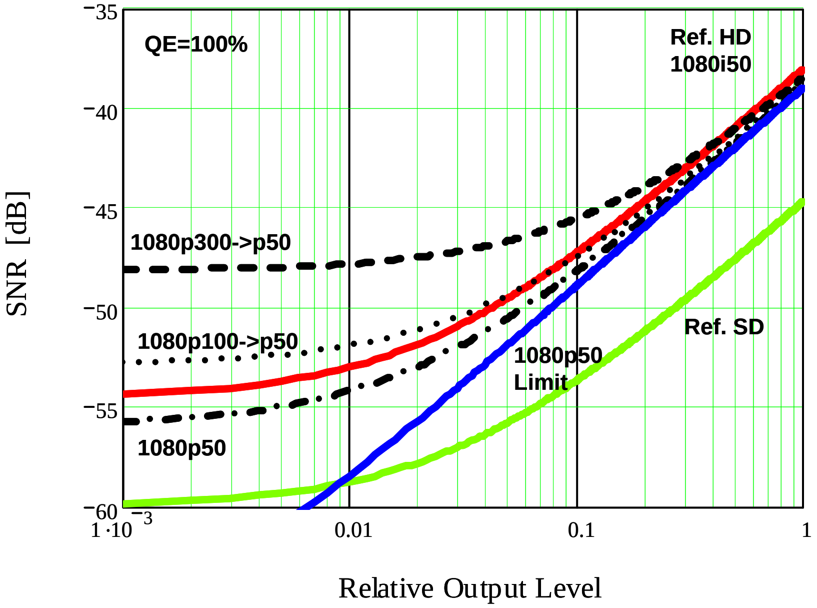

Figure 3 - Noise in Y as a function of output level for an output scanning format of 1080p50 and different input scanning formats at f/10. QE=100%.

In **figure 3** the camera output format is 1080p50 with 1080p50, 1080p100 and 1080p300 as input formats. The 0dB settings are at 2000lux, f/10 and the conversion efficiency is at the theoretical limit of 100%.
Again the tendency for the photon shot noise to converge at high output levels is clearly visible. The noise in the dark parts is problematic. Again the blue solid curve denoted Limit is reached when there is only photon shot noise with a conversion efficiency of 100%. One cannot pass this line without violating the laws of physics.

THE NEXT STEP IN IMAGERS: A CMOS IMAGER FOR BROADCAST

Based on the considerations outlined in the CMOS Imager paragraph, Grass Valley embarked on its own CMOS imager development. The Xensium imager was presented at the International Solid State Circuits Conference (ISSCC) in February 2007, [16]. Its design was a close collaboration between the Thomson Silicon Components design group in Villingen, Germany, and the Grass Valley Camera R&D division in Breda, the Netherlands.

In the architectural choices made for Xensium we leveraged on the massive experience in imagers and in signal processing. The design philosophy had three main elements, of which the first two where:
-the imager does not need to be perfect as long as the images from the camera are.
-the imager is in essence an analog device and therefore allow only for a minimum of digital circuits. So keep the design of the imager as close to its basic functions as possible, Figure 4.

In line with this approach was to place the majority of complexity, and hence the flexibility, in an external FPGA. In the end it was essential in getting Xensium realized in time.

The advantage of this type of architecture is in the flexibility one has in readout type and frame rates and performance.

The imager was decided to be a full HDTV 2/3” imager with 16:9 aspect ratio. An image diagonal less then the 11mm would result in several of the degrading issues discussed in [4]. Given the full HDTV resolution of the imager and the 11mm image diagonal one finds that the pixels are 5 μm x5 μm.

Broadcast applications are very demanding. To underpin the effort made by the team: to reach the goals, needed to fulfill the demands, over 15 patents where applied for.

Figure 4: Through the ‘SERIAL CONTROL INTERFACE’ a diversity of settings can be programmed. Like Region of interest, analog gain, type of horizontal multiplexing.
A state machine is on-chip for generating the basic timing. The imager has several independent line scanning registers to enable flexible vertical scanning, among which 1080p, 1080i and 720p. The chip has 2 A-to-D converters which input receives a pre-conditioned video signal from the gain stages which are offset compensated.

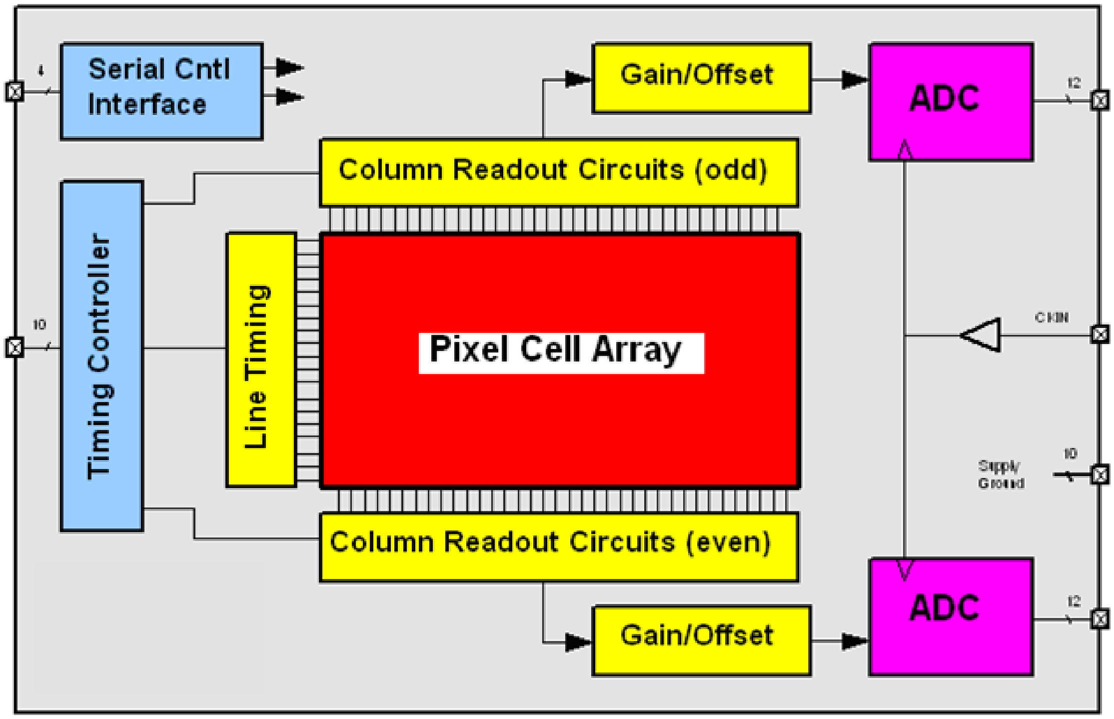

Figure 4 - Block diagram CMOS imager

The ‘LINE TIMING’ block addresses the pixel array in a row by row fashion. The addressing takes place during the line blanking. At the end of the line blanking the row signal is sampled at the column read circuits. Where they are multiplexed towards the ‘GAIN/OFFSET’ stage

The circuits are designed to run at at least 2x111MHz allowing for 1920x1080i180 raw. The raw data rate can be used to allow for true Correlated Double Sampling, [12]. It is used to achieve the broadcast signal-to-noise ratios and clean images.

The chipgraph is shown in figure 5. The highlighted functional blocks like “VERTICAL TIMING”, “COLUMN CIRCUITS”, “AMPLIFIER BLOCK’s” and ADC can also be found in the block diagram, figure 4.

The power consumption is about 1/3 of that of an equivalent CCD application.

High performance levels are reached through measuring the black level of each pixel and correcting the acquired image with it (known as correlated double sampling). Achieving a noise level of 11.5e overall and 4e for the pixel only. Due to the architectural choices this result is achieved without having to use soft reset, [3]. As a result no LAG occurs and no tweaking is needed for each scanning mode in obtaining a sufficient performance level.

The saturation level of the photodiode is at about 35kel. At 0dB camera gain setting 15kel is used to arrive at the 400%-500% overexposure. The latter numbers guarantee a CMOS imager that can be used in broadcast and Pro/AV applications.

The basic sensitivity of the imager starts high up at the sensitivity curve due to its fill factor of 56%.

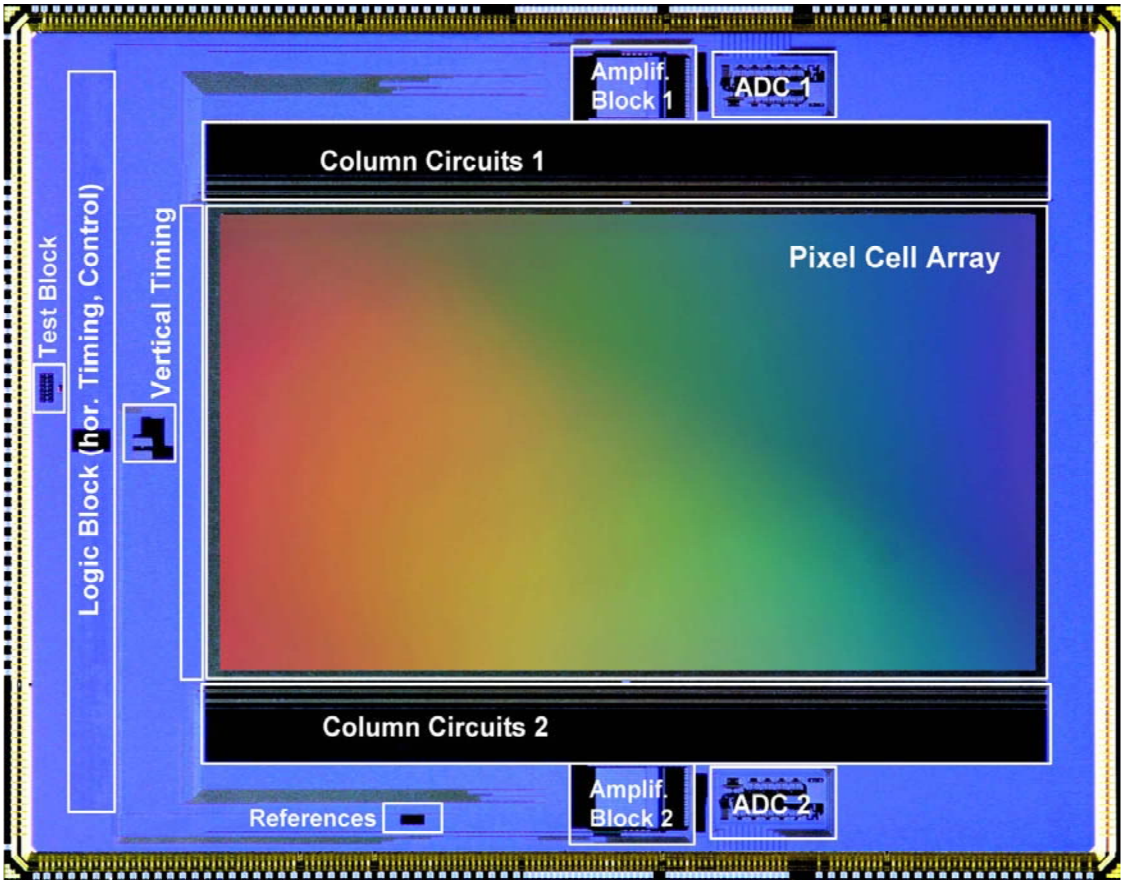

Figure 5 - *Chipgraph of Xensium*

The potential performance of this CMOS imager is best demonstrated by describing the scanning modes and its signal-to-noise for a fixed illumination of 2000lux; 3200K; 89.9% reflectance and SNR=54dB but with variable f-number, Table 3:

Mode	**f-number@54dB; 2000lux; 3200K; 90%**
1080p30	**f/11**
1080i60 Field mode	**f/8-f/10**
1080i60 Frame mode	**f/11**
720p60	**f/8-f/10**

Table 3 - Sensitivity versus some of the scanning rates.

Finally, in Figure 5, some preliminary high dynamic range images are shown to depict the exposure latitude a CMOS imager can have.

Figure 5 – Image down-left: normal operation, Image upper-right HDR image.

CONCLUSIONS

The history of CMOS imagers in relation to the matured CCDs is discussed. This knowledge was put to use in deciding when and how to design our own CMOS imager. It was shown that CMOS imagers for full HDTV are becoming viable now.

With the reporting of Xensium the first full HDTV imager is presented that offers broadcast quality images.

The architectural choices of Xensium enabled the development of a camera that reaches broadcast and Pro/AV quality.

Based on the limits of physics one can conclude that with regard to the shot noise in 1080p50 at f/10 one will always have images on the edge of being noisy. If one wants to achieve the same noise impression as SD one either has to apply noise reducers or accept that f-numbers in the range of f/5.6 are needed as a 0dB setting for the camera.

Generating 1080p50 or 1080p60 from a 1080p300 source will have the same noise impression for the exposed parts. The darkened areas of the images will be too noisy until the readout noise (noise in black) is reduced further.

REFERENCES

[1] L. Kozlowksi, Performance Limits in Visible and Infrared Image Sensors, IEDM DEC 1999, pp 36.1.1-36.1.4.

[2] H. Wong, Technology and Device scaling considerations for CMOS imagers, IEEE ED Vol. 43, pp. 2131-2142, DEC 1996.

[3] M. Loose et al. 2/3-inch CMOS Imaging Sensor for High Definition television, IEEE workshop on charge-coupled devices and advanced image sensors, June 7-9, 2001.

[4] M. Schubin, Optics of Small-Format HDTV Acquisition, SMPTE 147th Technical Conference and Exhibition, Nov. 9-12, 2005

[5] A. Theuwissen, Solid-State Imaging with Charge-Coupled Devices, 1995, Kluwer Academic Publishers, ISBN 0-7923-3456-6.

[6] P.Centen et. al., Aspect Ratio switching with equal horizontal pixel count, Technical Papers International Broadcast Convention, pp. 1-7, Amsterdam, September 1994.

[7] P. Centen, T. Moelands, J.v. Rooy, M. Stekelenburg, A Multi-Format HDTV Camera Head, proceedings of the 142st SMPTE Technical Conference and Exhibition, Pasadena, October 18-21, 2000.

[8] Invited, P. Centen, 2/3”- Cameras Challenging the Latitude of FILM, NAB Conference and Exhibition, Digital Cinema Summit 12-13 April 2008, Las Vegas.

[9] Y. Ishihara, K. Tanigaki, A High Photosensitive IL-CCD Image Sensor with Monolitic Resin Lens Array, Technical Digest IEDM, pp 497-500, Dec 6-8, 1963.

[10] M. Blouke, J. Breitzmann, J. Hall; Three-phase, backside illuminated 500×500 CCD imager, Solid-State Circuits Conference. Digest of Technical Papers. Feb-1978, Volume: XXI, page(s): 36- 37.

[11] J. Gambino et. al, CMOS Imager with Copper Wiring and Lightpipe, Electron Devices Meeting, 2006. IEDM apos;06. International, Volume , Issue , 11-13 Dec. 2006 Page(s):1 – 4,.

[12] M. White et al, “Characterisation of Surface Channel CCD Image Arrays at Low Light Levels”, IEEE JSSC, Vol. SC-9, No. 1, pp 1-13, 1974.

[13] Y. Nishida et.al, Design concept for a low-noise CCD image sensor based on subjective evaluation; Electron Devices, IEEE Transactions on, Volume 36, Issue 2, Feb 1989 Page(s):360 – 366

[14] J. Janesick, K. Klaassen, T. Elliott, Charge-coupled-device charge-collection efficiency and the photon transfer technique, Optical engineering, October 1987, Vol. 26, No.10, pp 972-980.

[15] P. Centen, Applied Mathematics to Simplify Imager and Camera Analyses Proceedings of the IEEE International Image Sensor Workshop, 7-10 June 2007, pp 121-124.

[16] P. Centen et. al, A 2/3-inch CMOS Image Sensor for HDTV Applications with Multiple High-DR Modes and Flexible Scanning, ISSCC2007, San Francisco, 11-14 February 2007, pp 512-513.