Development Status of CMOS 1920 x 1080 Imaging System-on-Chip for 60p and 72p HDTV

By L. Kozlowski, G. Rossi, L. Blanquart, R. Marchesini, Y. Huang, and G. Chow

High-performance charge-coupled device imagers (CCD) with 2 million pixels in 1920 x 1080 format have been commercially available for more than ten years.[1] Nevertheless, today's HDTV cameras still do not simultaneously provide sufficient resolution and electronic film speed at the minimum progressive frame rates needed for creative motion imaging production. Overcranking, for example, is only available at ~1 million pixel (1280 x 720) resolution, and the maximum frame rate of today's 1920 x 1080 progressive cameras is 30 Hz. To truly offer a compelling alternative to motion imaging film, electronic HD cameras must provide a higher progressive frame rate along with additional bit depth to maximize production quality and creativity. In this paper, the technical background for the shift in technology from CCD imaging to CMOS imaging system-on-chip is discussed. Specifically, the sensor architectures are compared, and why CCD noise consistently increases with video frequency at about 3 dB/octave is explained. We conclude by briefly reporting the development status for a complimentary metal oxide semiconductor (CMOS) 1920 x 1080 image sensor for next-generation, affordable and compact HD cameras to support progressive frame rates at 60 Hz and beyond.

The unique combination of functionality, low power dissipation, and imaging quality offered by the latest CMOS-based sensors, which are made in deep submicron CMOS (0.25 µm) technology, has convinced camera designers to consider switching from mature charge-coupled device (CCD) sensors. For some modern products, such as digital still cameras for mainstream consumers, 2- to 5-megapixel CCDs remain the best choice. On the other hand, for a growing number of demanding applications where the CCD was previously unchallenged, CMOS-based technology has emerged superior, due to unrelenting migration to deep submicron lithography and payback of long-term investments in system-on-chip (SoC) integration.

Leveraging a key architectural advantage and ongoing improvements in CMOS technology, driven by a robust semiconductor industry that produces multi-GHz microprocessors at consumer prices, CMOS-based image sensors are now more capable than CCDs of meeting this challenge.

One battle already won by CMOS is the high-end digital still camera; another battlefield in the ongoing war is high-definition television. In extreme cases such as astronomy, where the highest possible detector and readout performance is needed, a second CMOS-based competitor, the hybrid focal plane array[2] was developed. This device, consisting of a CMOS readout attached to a silicon detector array with flip-chip packaging, has emerged as another excellent alternative to the aging CCD-multichip module production capability. Its costs are now decreasing rapidly via foundry availability.

Figure 1 compares development chronologies for CCD, MOS, and CMOS sensors, using the number of imaging pixels per sensor as the key figure of merit. Today's battles for image sensor supremacy span the range from 2 to 14 million pixels and are especially heated and intriguing. Although CCDs

with more than 10 million pixels were first introduced about ten years ago, commercial CMOS sensors having up to 14 million pixels recently became available by exploiting CMOS process technology at 0.35 µm. The latest CMOS sensors quickly conquered the high-end digital still-camera market—professional-grade CMOS digital SLR cameras from Canon, Kodak, Nikon, and Sigma produce the best still-image quality available for professional digital single lens reflex (DSLR) cameras. This surprising reality is likely not a fleeting anomaly; Kodak, a leader in producing large format CCDs, chose a CMOS sensor for their top DSLR camera, the DCS Pro 14n.[3]

Sensor Technology Primer

The mainstream CMOS and CCD technologies are reviewed in this section. The latter is still the most widely used technology in visible cameras. The former uses deep submicron CMOS technology to perform the signal processing that converts each pixel's analog photo-generated signal to video, whether the final form is analog or digital.

Monolithic CMOS

Monolithic CMOS imagers for visible cameras predominantly use pixel-based amplifiers. Figure 2 shows a representative imaging system-on-chip (iSoC) architecture with an amplifier at each pixel. By embedding extensive on-chip functionality, the CMOS iSoC can be a standalone multiformat camera needing only external power to produce NTSC, PAL, or HDTV video, among other protocols. The underlying CMOS technology currently supports, for example, the following on-chip functional blocks:

- Pixel-based amplifiers with analog or digital output.
- System-on-chip signal processing for suppressing kTC noise below limit of standard diode reset schemes used in CCDs[4] to generate video with at least 12-bit fidelity.
- System-on-chip signal processing for reducing dark current and fixed pattern noise.
- Programmable clock and bias generators.
- Analog and digital programmable gain amplifiers.
- Clamping of "black" signal level to rescribed or programmable reference value to preserve shadow detail.
- Use of pinned photodiode design to minimize dark current by effectively "burying" the photosensitive region. The surface pinning layer accumulates holes that passivate current-generating traps at the oxide-semiconductor interface.
- High-speed (5 MHz to f200 MHz) A/D conversion with 8- to 14-bit resolution.
- Low-speed (<2 MHz) A/D converter(s) with 10- to 18-bit resolution.
- Static or dynamic memory.
- Feedback circuits for automatic gain control (AGC) and/or white balance.
- Built-in system test (BIST).

Using the latest CMOS technology facilitates the highest possible optical fill factor and lowest noise amplification so that electronic speed and exposure latitude are maximized. Figure 3 shows the achievable as-drawn optical fill factor for implementing a conventional source follower per detector (SFD) pixel amplifier. It includes a photodiode along with three transistors per pixel using CMOS technologies from 0.8µm to 0.18µm that came and went during the last decade. The ratio of the light-gathering area, defined by the photodiode to the total pixel area is the fill factor. Fill factor >50% was difficult about five years ago but is now readily achieved at 5µm pixel pitch using 0.25µm lithography. A 5µm pixel pitch translates to 2/3-in. imaging format at 1920 x 1080 resolution for professional cameras; ~7.4µm pitch supports 1280 x 720 format. Figure 3 indicates that migrating to 0.18µm further increases the fill factor or reduces pixel pitch.

Advanced CMOS alternatively enables advanced pixel amplifiers offering lower noise, global shutter electronic imaging, or direct digital output[5,6] from each pixel. Scrutiny of Fig. 3 shows that 0.18µm technology produces fill factors at ~4µm pitch superior to frame interline transfer CCDs because these typically offer <50% fill factor depending on anti-blooming efficacy.

Electronic gating at each pixel's CMOS amplifier allows the signal bandwidth to be set at the optimum Nyquist frequency. The bandwidths for other key amplifiers in the signal chain can also be tailored to minimize noise without significantly degrading MTF. On the other hand, a key CMOS disadvantage only recently obviated is that each amplifier has different offset and, possibly, different gain. These mismatches otherwise generate fixed-pattern noise and/or dynamic texturing in electronic images. Fortunately, the ability to perform analog or digital signal processing in the SoC today mitigates this key issue.

Hybrid CMOS

Hybrid CMOS imagers also use pixel-based amplifiers and similarly offer imaging SoC functionality. The key distinction relative to both monolithic CMOS and CCD sensors is the unique potential for detector array optimization, independent of the supporting CCD or CMOS readout, as evident in Fig. 4. Consequently, trading for the higher production cost incurred in flip-chip packaging, hybrids offer the following benefits in addition to those listed for the monolithic sensor:

- Global shutter with near-100% fill factor in conventional package and no need for external mechanical shutter.
- Vertically stacked multicolor detection sans mosaic color filters.
- Quantum efficiency can be limited only by the specific anti-reflection coating.
- Pixel real-estate can be maximized to optimize the pixel amplifier within the full pixel area.
- Capability for optimized pinned photodetector operation to minimize dark current while maximizing quantum efficiency across the spectral band of interest (with useful silicon-based range from about 250 nm to just beyond 1000 nm.
- Visible sensor can use alternative detector material with optimum bandgap for color imaging to essentially eliminate dark current relative to silicon photodiodes.
- Capability for producing wafer-scale sensors with many tens of millions of pixels on 200 mm and 300 mm silicon substrates.
- Ongoing minimization of production cost due to commercialization of multi-chip module packaging via wafer-scale interconnect deposition.

Monolithic CCD

CCD technology is mature with respect to all key design aspects, production yield, and performance—improvements have been incremental over the past ten years. Photo-generated carriers are typically accumulated in a pinned photodiode (hole-accumulating layer at surface to suppress dark current generation) and successively transferred via slow (vertical) and fast (horizontal) buried-channel CCD shift registers to the output node, where the charge is subsequently converted to a voltage. The key CCD noise sources for most applications are the reset noise (i.e., kTC) of the sense capacitance(s) that converts the charge to a voltage, the thermal noise in the output amplifier,[7-9] and the external camera noise. While noise generation within the CCD is very low, including temporal noise and fixed pattern noise, the relevant noise bandwidth of megapixel CCD imagers is several orders of magnitude higher than for a CMOS-based sensor to produce video with minimum electrical crosstalk. The read noise of a large-format CCD operating at tens of MHz is dominated by the quadrature sum of white noise from the output amplifier and the camera electronics— especially after correlated double sampling (CDS) is applied to remove the reset noise, as this process doubles the noise bandwidth.

CCDs require several external functional blocks to produce video. Modern cameras, however, are supported by application-specific integrated circuits (ASIC) that combine several functions onto sophisticated CMOS ICs. In effect, advanced CMOS sensors optimally combine the image formation with the back-end ASIC and eliminate the need for low-noise analog design wizardry on the printed circuit board where the image sensor resides. Key supporting blocks that are external to the CCD include correlated double sampler (CDS) to reduce reset noise; noise filter with Nyquist bandwidth defined by maximum video rate; clock generator via programmable gate array (FPGA), CMOS programmable logic device (CPLD), or dedicated ASIC; clock drivers; bias generator; analog programmable gain amplifier with limited range and resolution to amplify signal and provide gamma; automatic gain control; and video signal processor.

Performance Comparison

Read Noise and S/N Ratio

Figure 5 shows CCD noise data in volts and electrons from CCD and camera manufacturer datasheets along with approximate camera S/N ratio for HDTV with line-mixing. Various CCDs for disparate cameras encompassing astronomy through professional video are represented. Line-mixing refers to the practice of co-adding adjacent lines during each field of interlace video to double the S/N ratio (i. e., increase S/N ratio by 6 dB) at the expense of vertical definition. Similarly adding, or co-adding, adjacent pixels in the horizontal direction can further double S/N ratio in trade for horizontal resolution. The standard test condition for measuring sensitivity and S/N ratio for HDTV

comprises 2000 lux illumination, 89% scene reflectivity, 30 Hz or 60 Hz frame rate, and f8 lens aperture. For 2/3-in. optical format cameras operating at 60 Hz interlaced with line mixing (or 30 Hz progressive without line mixing), typical SNR is 54 to 56 dB. Progressive CCDs operating at 60 Hz provide SNR below 45 dB due to higher noise and inability to perform line-mixing.

At lower video frequencies, CCD noise and SNR varies over several orders of magnitude due to variation in sensor performance and design, camera design, semiconductor fabrication, and the practical challenges encountered when trying to minimize white noise with electronic circuits having high impedance amplifiers. At higher frequencies, CCD noise increases at about 3 dB per octave.[7] This increase is fundamental to CCDs because the output amplifier and its bandwidth largely set the minimum noise level. These data show a practical lower limit for CCD noise due to the composite white noise of the sensor and the camera electronics labeled "empirical white noise limit." The resulting trend line graphically demonstrates the typical 3 dB increase in CCD noise per octave in video frequency.

Of utmost importance to camera manufacturers and broadcast engineers is, consequently, the impact of noise on S/N ratio as demands on resolution and refresh rate together increase the necessary frequency for video production. The upper limit set by CCD noise embodied by the empirical white noise limit trend line is increasingly unappealing as sensors move well into the megapixel regime for video applications. So, while a CCD can have extremely low noise and high performance, these attributes are not achievable at the higher video rates to which the world is moving. The most popular CCD workaround is to increase the number of video taps to reduce the video frequency and noise at each tap. This brute force workaround invariably serves to increase camera power consumption, size, and cost.

Designers of CMOS-based sensors reduce high-frequency noise and thus eliminate this classic noise versus resolution trade, faced by CCD designers because the CMOS pixel amplifier bandwidth better matches the signal sampling frequency. The CMOS output buffer's noise is usually rendered negligible by appropriately setting the transfer gain or sensor transimpedance. It is consequently an inherent characteristic of well-designed CMOS sensors that the noise is constant versus video rate, as the CMOS pixel amplifier sets the noise bandwidth, not the video amplifier.

CMOS manufacturers have now corroborated this basic advantage.[8,9] Noise measurements up to 200 MHz with a new CMOS sensor specifically developed for producing HDTV content is included in Fig. 6. An imaging system-on-chip sensor having various on-chip circuits including 12-bit A/D conversion and signal processing to support HDTV production at up to 1080/72p/72i; its minimum read noise is ~10 electrons at up to 200 MHz video frequency using programmable settings optimized for professional and broadcast cameras. The measured camera S/N ratio approaches 60 dB with two-line mixing in extended sensitivity (ES) 720/60p mode wherein the 1920 x 1280 format is directly down-converted on-chip to 1280 x 720. Both 1920 x 1080 and extended sensitivity 1280 x 720 are 60 Hz progressive operating modes corresponding to composite 150 MHz video rate. The measured read noise for an earlier 1920 x 1080 CMOS SoC sensor operating at full 1920 x 1080 format was <25 e- for 60 Hz interlaced and 30 Hz progressive operation at 75 MHz video rate.[10] Takayanagi[11] also recently described an 8.3 million pixel CMOS sensor for UHDTV with 10-bit performance at composite data rate of ~500 MHz.

The relevant noise bandwidth for CMOS imagers is typically programmed via register settings such that there is no need for external filtering, e.g., CCDs use external 30 to 33 MHz filtering for HDTV. This lessens the noise bandwidth by several orders of magnitude from tens of MHz for CCDs to tens of kHz for CMOS. The read noise for CMOS iSoC sensors thus remains low and constant, even at very high video frequencies. Optimally setting the noise bandwidth well ahead of the video amplifier and using system-on-chip signal processing clearly minimizes sensor noise.

Electronic Film: Equivalent ISO Speed Rating

Straightforward methodologies exist to directly compare the sensitivity of motion imaging or photographic film to electronic sensors.[12] ISO specification 12232 documents the industry standard. Effective sensor read noise can therefore be estimated by measuring the ISO speed of digital

still and video cameras. In general, per the definition of film ISO speed rating, accurate camera exposure is obtained when the shutter speed is set to the reciprocal of the ISO speed with lens aperture at f16.

Following Baer,[13] the theoretical ISO speed curves versus read noise and pixel pitch are plotted and calculated at 100% optical fill factor shown in Fig. 7. At a square pixel pitch of 3µm per side and 10 e- read noise, the theoretical film speed for an electronic sensor that collects all photons impinging on the $9\mu m^2$ pixel area is ~100 ISO. Accurate electronic film exposure at f16 and 100 ISO would therefore require setting the shutter to 1/100th sec.

By measuring the ISO of an abundance of digital still cameras, the prior data was corroborated and the read noise was further established for state-of-the-art CCDs. The effective read noise levels for a representative sampling of today's digital still-cameras where the maximum video rate is typically about 28 MHz is compiled in Fig. 7. Considering the reduction in read noise due to camera signal processing, which is extensive in some cameras (up to 4x reduction in "grain" noise in image regions having little detail), CCD read noise is shown to be ~20 e- at 28 MHz. Indeed, the expected outcome that CCD read noise is higher at larger pixel pitch can also be seen. This is a consequence of requiring lower conversion gain to accommodate the larger charge-handling capacity so that the practically achieved white noise translates to a larger read noise in units of electrons.

The data embedded in Fig. 7 includes both CCD and CMOS sensors, although no CMOS ISO data is readily available at this time for pixel pitch below 4µm. Nevertheless, the superiority of CMOS sensor technology is clearly emerging at pixel pitches >6µm. Here, the cameras using CMOS sensors provide remarkably superior ISO ratings to the competing CCD-based cameras.

Qualitatively, the ISO speeds achieved with CCDs at pitch below 3µm decreases along with sensor pitch. Though the constraints imposed by physics apply to all sensor technologies, the practical read noise limit of CCDs is an engineering issue. The obvious path to achieving higher ISO speed as the pixel pitch pushes to 2m and less is to reduce read-noise. Based on the data presented in Figs. 5 through 7, this is clearly not the path that CCD technology follows.

Figure 7 also includes available ISO data for CMOS and CCD sensors, although the CMOS data is at over twice the CCD video frequency and thus favors the CCD. Nevertheless, focusing on progressive rather than interlaced operation and excluding line-mixing, CCDs for HDTV have ISO speed of about 300 at video frequency of about 40 MHz per video tap. HDTV CMOS (2.3-in. optical format with 5µm pitch) currently exhibit ISO speed from ~250 to 400 at up to 200 MHz frequency (including both temporal and fixed pattern noise). These data agree with the theoretical curves for read noise between 10 e- and 25 e-. HDTV CCD electronic film speed is typically increased to ISO 600 by operating interlaced mode and thereby enabling 2-line mixing with concomitant 6dB boost in S/N ratio.

Video Sensor Performance

As postulated and validated in the previous sections, CMOS technology is currently most advantageous at higher video rates, especially at 10 MHz and higher. However, as the pixel amplifier technology evolves along with the available lithography, the crossover frequency is expected to reduce to well below 1 MHz. This section reports the performance of CMOS sensor technology for HDTV. ATSC specifications for HD video at 1080/30p, 1080/60i, and 720/60p set the video frequency at 74.25 MHz.

Read Noise vs. Video Frequency

Figure 8 compares the validated empirical CCD noise trends (for 1 and 2 video taps) to the measured read noise of a recently developed CMOS sensor, developed for 1080-line HDTV. The HDTV CMOS sensor's noise for producing high-definition video is nearly constant versus video frequency at about 15 e- to the maximum frequency of 200 MHz. Specifically focusing on 1080/60p and 1080/72p video, Fig. 9 shows CMOS HDTV sensor noise as a function of signal level over 10-bit dynamic range (1024 digital numbers, or DN). The minimum noise is about 11 e- since the conversion factor is about 10 e-/DN (Fig. 10). Setting the sensor's black clamp level sets minimum video level in the absence of light to about 64 DN. Alternately tuning the sensor for low-light-level imaging by readjusting for 8-bit latitude reduces the read noise floor over this range to below 5 e-.

HDTV S/N Ratio

At illumination of 2000 lux and f8 aperture with 89% reflectivity, a 1920 x 1080 CMOS imaging system-onchip sensor recently achieved the signal-to-noise ratio plotted in Fig. 11. Specifically at the signal levels corresponding to the standard illumination condition for HD video, i.e., 100% signal, the S/N ratios at 60p and 72p are both approximately 55 dB without line-mixing. In comparison, CCD S/N ratio at 30p is 48 to 50 dB. By using line-mixing at 1080/60i, CCD S/N ratio is increased to over 54 dB.

Conclusion

This paper reviewed the basic theoretical differences of CCD and CMOS sensors along with supporting data and showed why CMOS technology is emerging as the preferred sensor technology for HDTV.

As in the ubiquitous CCD, CMOS-based imaging sensors are made in silicon and use photodiodes to detect light. The similarities between the two technologies quickly end in the way the resulting signal is handled. In CCDs, the photo-generated charge is noiselessly shifted from charge-handling bucket to charge-handling bucket until the packet reaches the output structure—the point where all the noise is generated at the broad electronic bandwidth needed to accurately generate video. The supporting video ICs remove some noise, specifically the reset noise from the "sense" capacitance that converts the charge into a voltage, via correlated double sampling and then digitize the video stream. The remaining white noise, however, limits camera S/N ratio and increases with frequency at ~3 dB per octave.

In the lowest noise CMOS sensors, the photo-generated current is instead quickly converted to a voltage at relatively narrow electronic bandwidth. Successive stages of analog and digital processing in the imaging system-on-chip (iSoC) suppress the temporal and spatial noise to the level set by the pixel design and the underlying CMOS technology. This noise level is significantly lower than the practical levels achieved with CCDs for HDTV. Furthermore, the lower noise levels are being achieved at video frequencies significantly higher than with competing CCDs.

Reference

1. L. Thorpe, F. Naguno, and K. Ishikawa, "A Comparison between HD Hyper-HAD CCD Camera and Color Film for Television Program Production," *SMPTE J., 103*:364, June 1994.
2. L. J. Kozlowski and W. E. Kosonocky, "Infrared Detector Arrays," *Handbook of Optics,* Ch. 25, McGraw-Hill: New York, 1995.
3. G. Meynants, B. Dierickx, A. Alaerts, D. Uwaerts, D. Scheffer, and S. Noble, "A 35mm 13.89 Million Pixel CMOS Active Pixel Image Sensor," *PICS 2003.*
4. U. S. Patent 6493030, Dec. 10, 2002.
5. B. Fowler, A. El Gamal, and D. Yang, "A CMOS Area Image Sensor with Pixel-Level A/D Conversion," Solid-State Circuits Conference, 1994, *Digest of TechnicalPapers, 41st ISSCC; 1994 IEEE International, 16-18, pp. 226-227, Feb. 1994.*
6. S. Kleinfelder, S. Lim, S. Liu, and A. El Gamal, "A 10,000 Frames/s CMOS Digital Pixel Sensor," *JSSC,* 36 (12), Dec. 2001.
7. K. Mitani, M. Sugawara, and F. Okano, "Experimental Ultrahigh-Definition Color Camera System with Three 8M-pixel CCDs," *SMPTE J., 111:*148, Apr. 2002.
8. L. J. Kozlowski, D. Standley, J. Luo, A. Tomasini, A. Gallagher, R. Mann, B. C. Hsieh, T. Liu, and W. E. Kleinhans, "Theoretical Basis and Experimental Confirmation: Why a CMOS Imager is Superior to a CCD," *Proc. SPIE, 3698:*388-396, July 1999.
9. L. J. Kozlowski, A. Tomasini, and J. Luo, "Performance Limits in Visible and Infrared Imager Sensors," Electron Devices Meeting 1999, *IEDM Technical Digest. International, 5-8*:867-870, Dec. 1999.
10. M. Loose, L. J. Kozlowski, A. M. Joshi, A. Kononenko, S. Xue, J. Lin, J. Luo, I. Ovsiannikov, J. Clarke, and T. Paredes, "2/3-in. CMOS Imaging Sensor for High-Definition Television," 2001 IEEE Workshop on CMOS and CCD Imaging Sensors, June 2001.
11. I. Takayanagi, M. Shirakawa, S. Iversen, J. Moholt, J. Nakamura, and E. Fossum, "A 1 1/4 in. 8.3M Pixel Digital Output CMOS APS for UDTV Application," *ISSCC* 2003.
12. Kodak CMOS Image Sensors ISO Measurement, Application Note MTD/PS-0234, July 24, 2001.
13. R. Baer and J. Holm, "A Model for Calculating the Potential ISO Speeds of Digital Still Cameras Based Upon CCD Characteristics," *IS&T PICS* 1999 Conference Proc., pp. 35-38 1999.

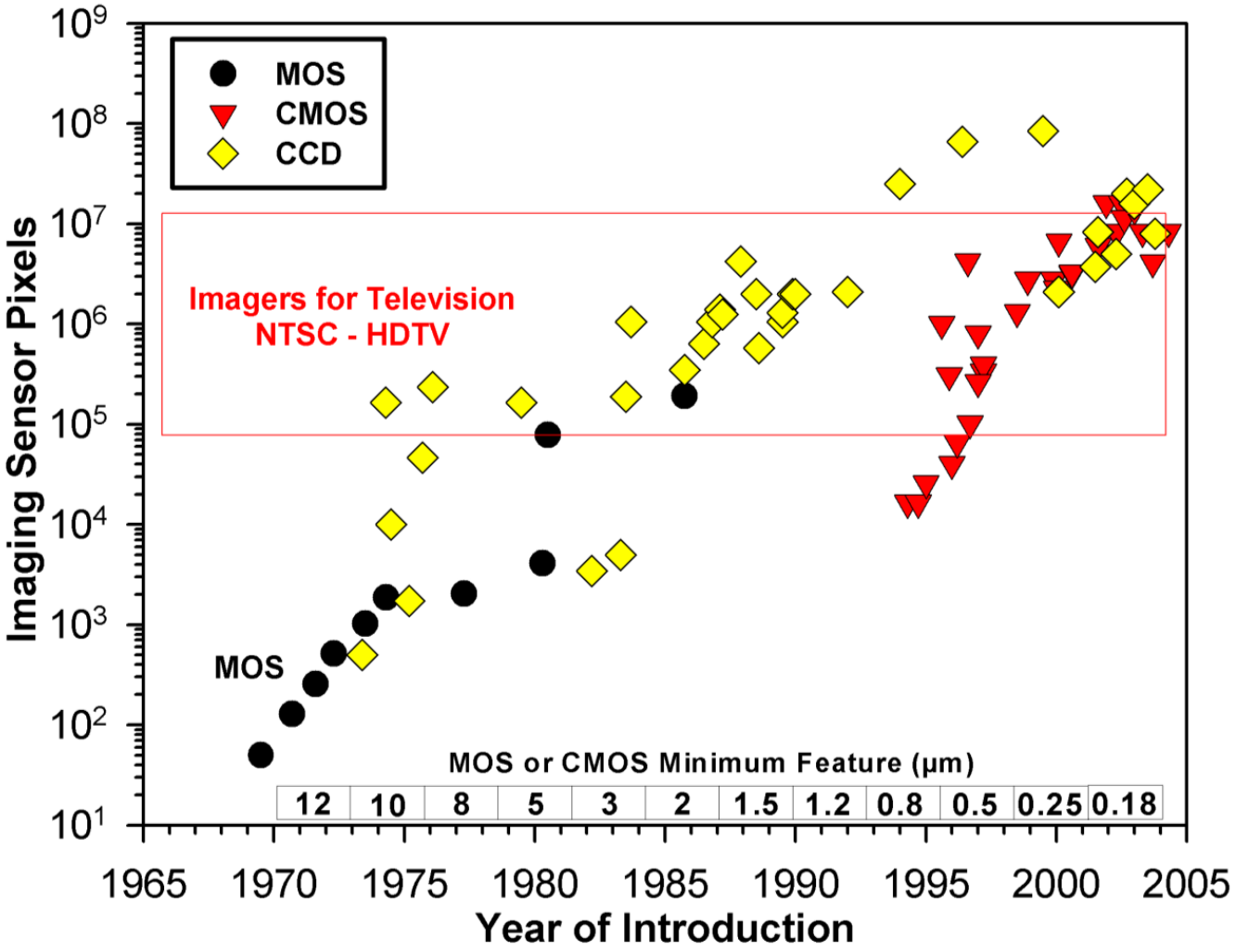

Figure 1. Chronology of visible imager development for CCD, MOS, and CMOS.

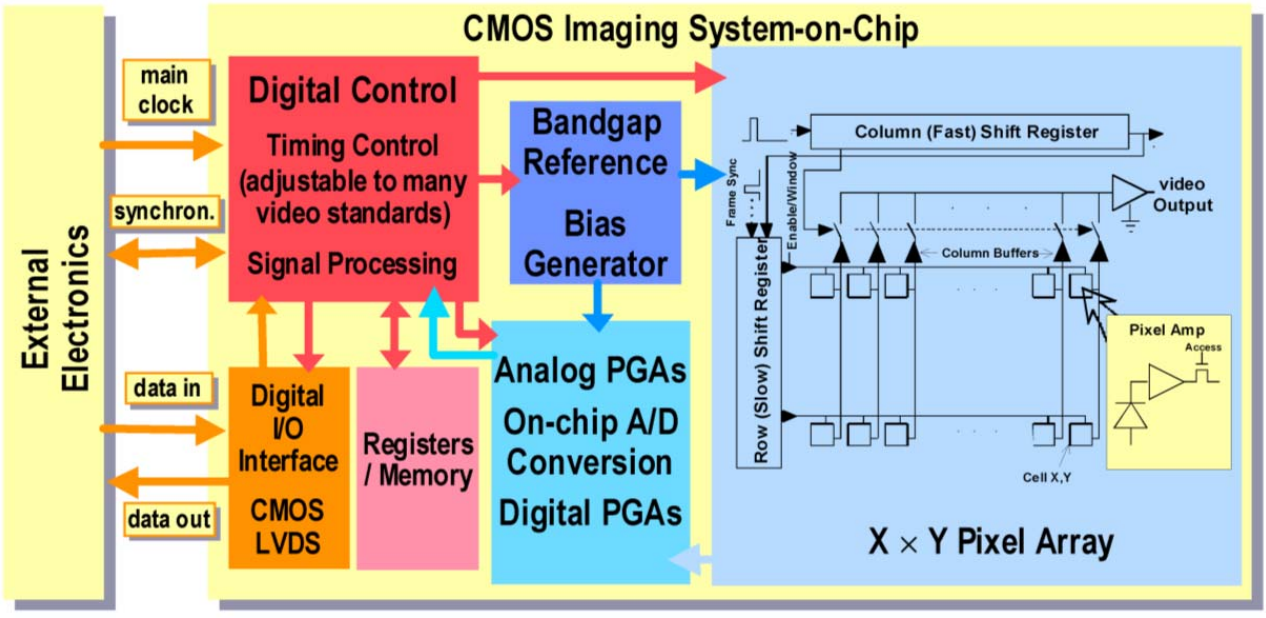

Figure 2. Architecture of representative CMOS imaging system-on-chip.

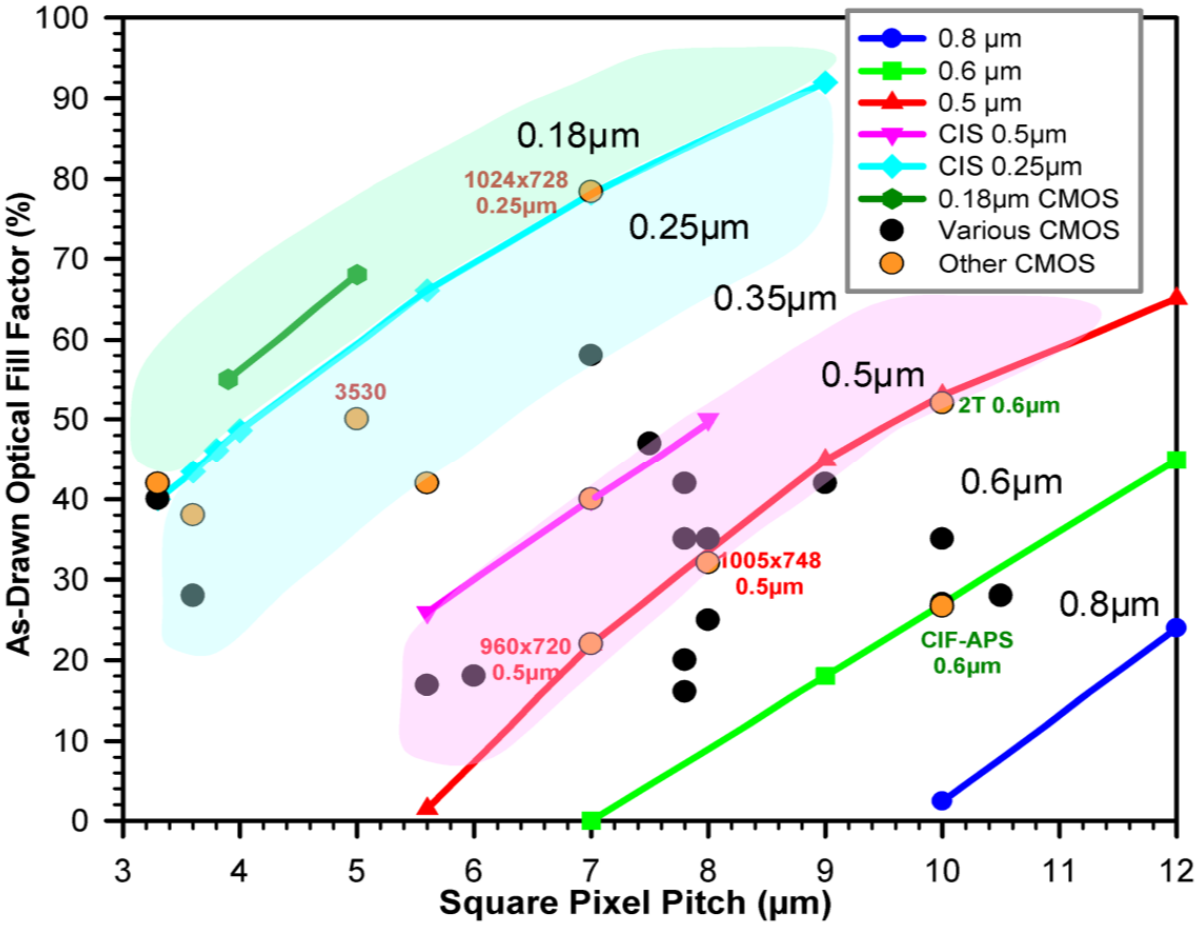

Figure 3. Optical fill factor vs. pixel pitch for monolithic CMOS with source follower amplifier.

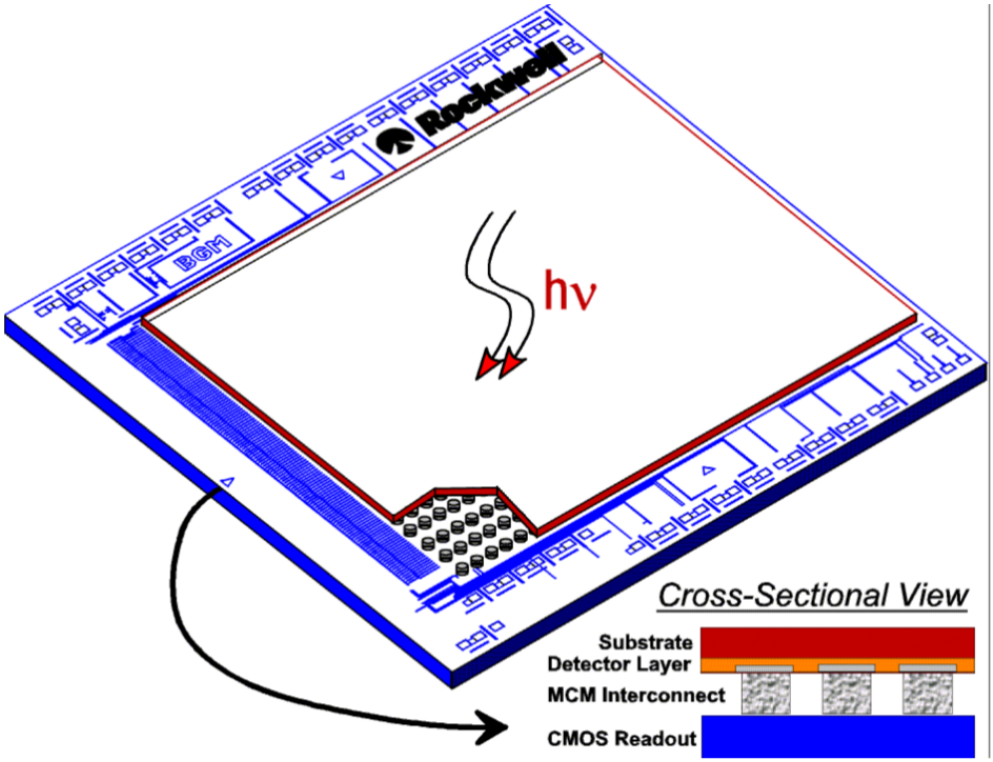

Figure 4. Hybrid FPA methodology for visible and infrared sensors.

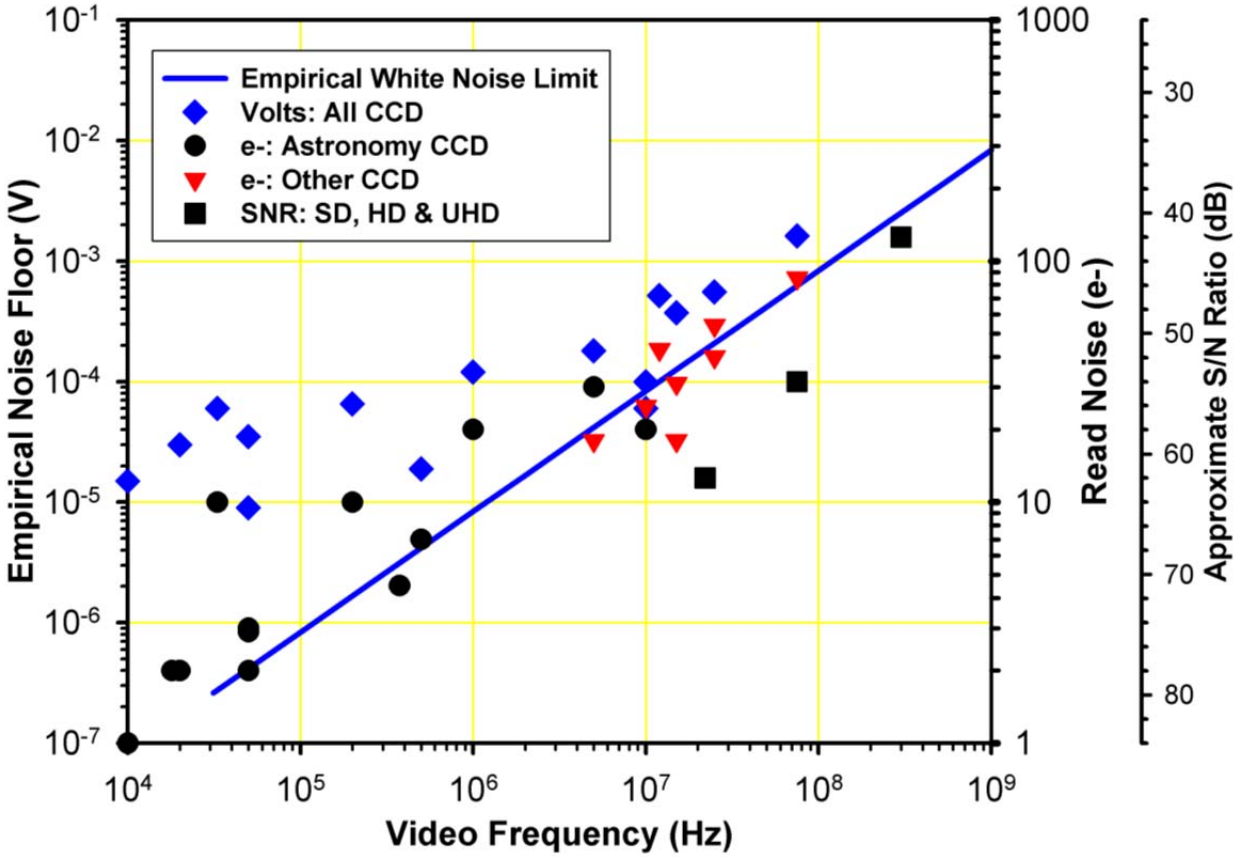

Figure 5. Observed read noise for CCD sensors.

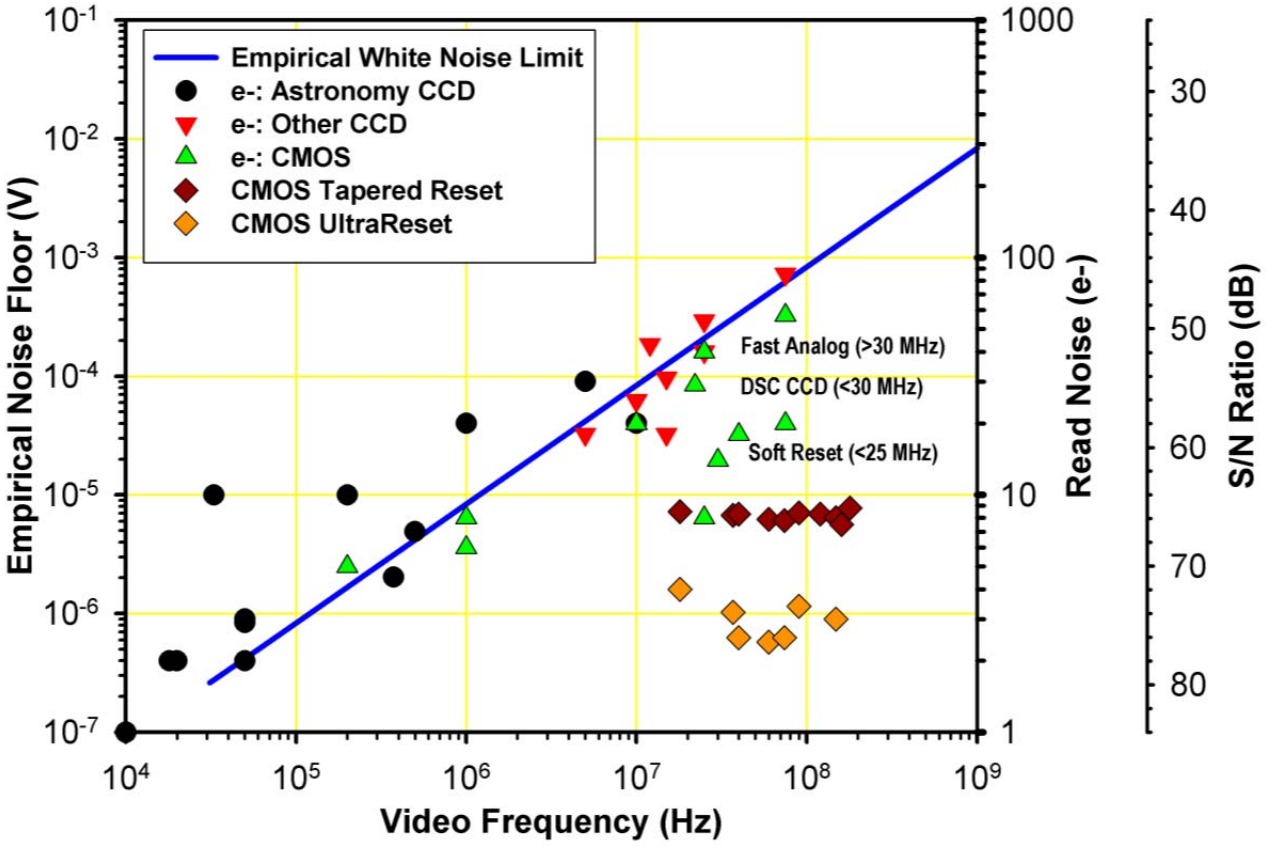

Figure 6. Measured read noise for CMOS and CCD sensors.

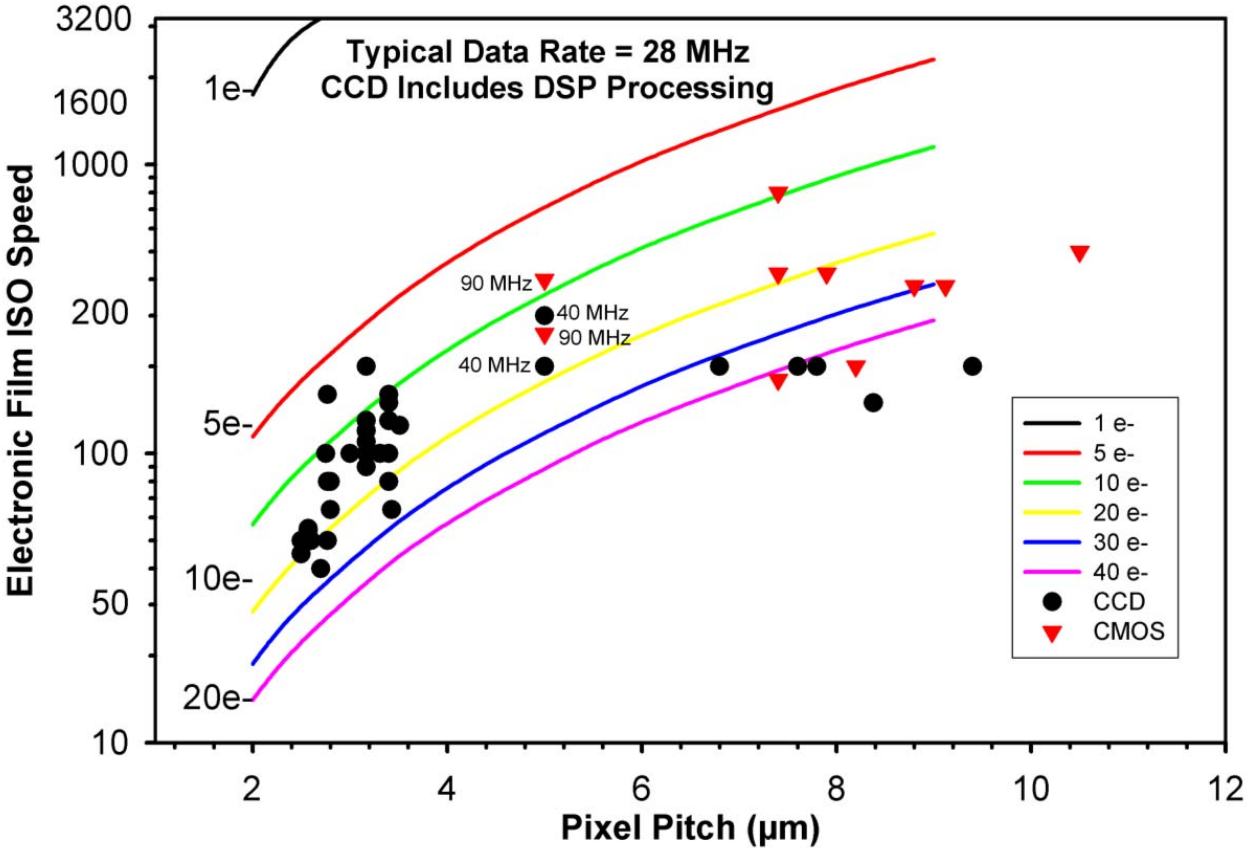

Figure 7. Electronic film speed for digital still-cameras ca. 2003.

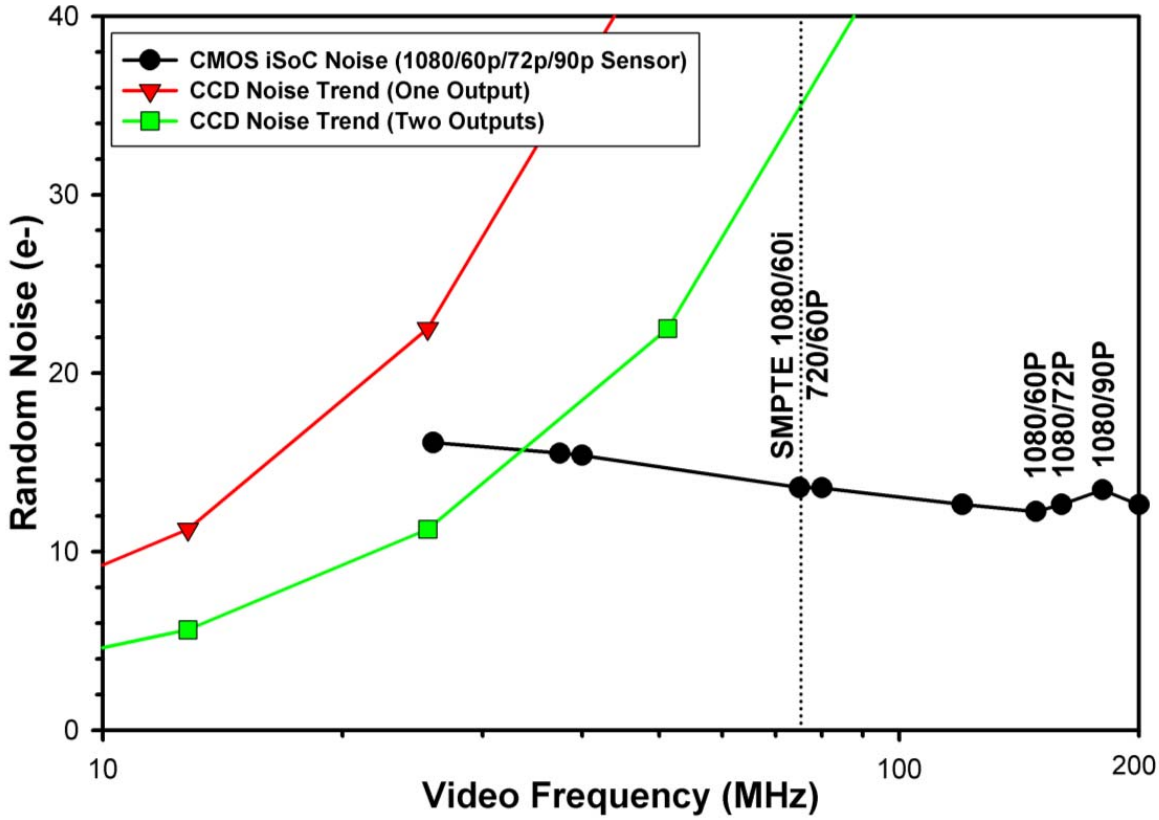

Figure 8. CCD noise trends at video frequencies vs. measured CMOS data.

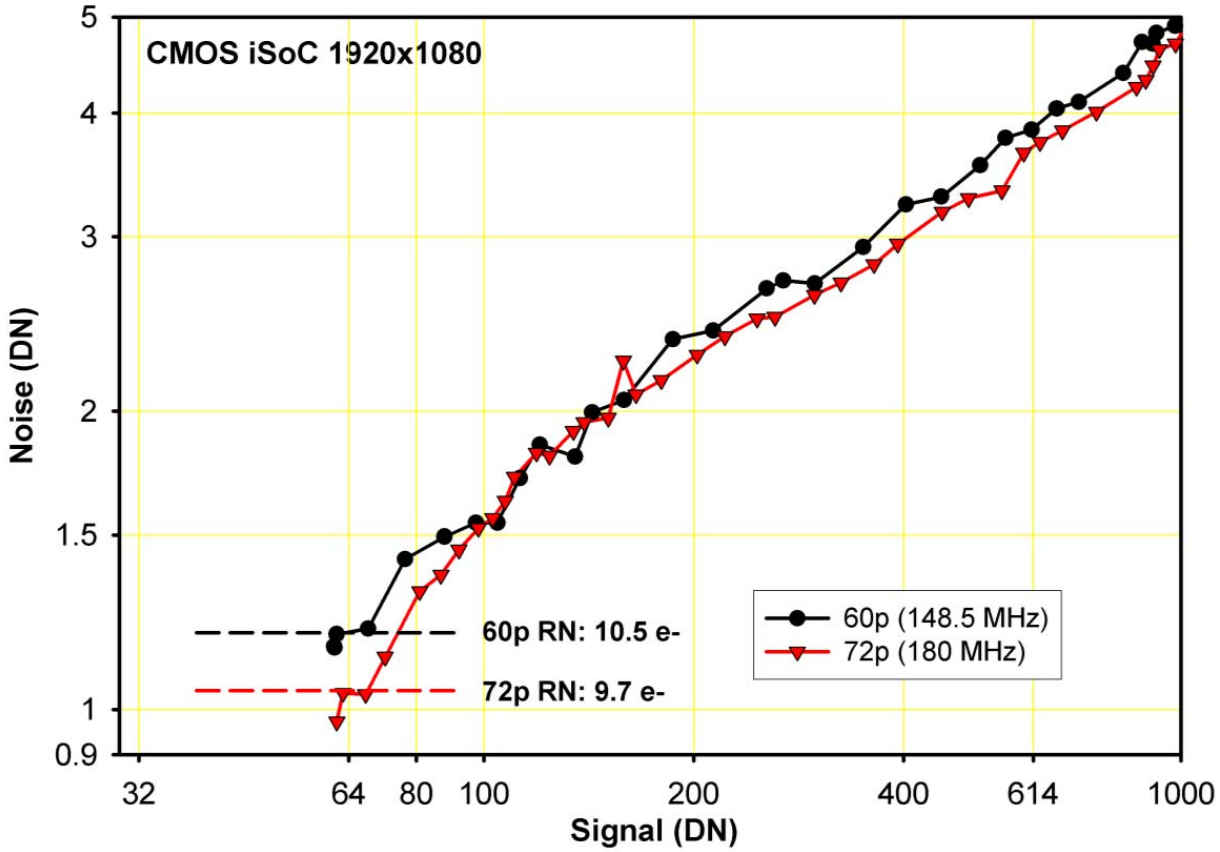

Figure 9. Measured noise vs. signal for HDTV CMOS at 1080/60p and 1080/72p.

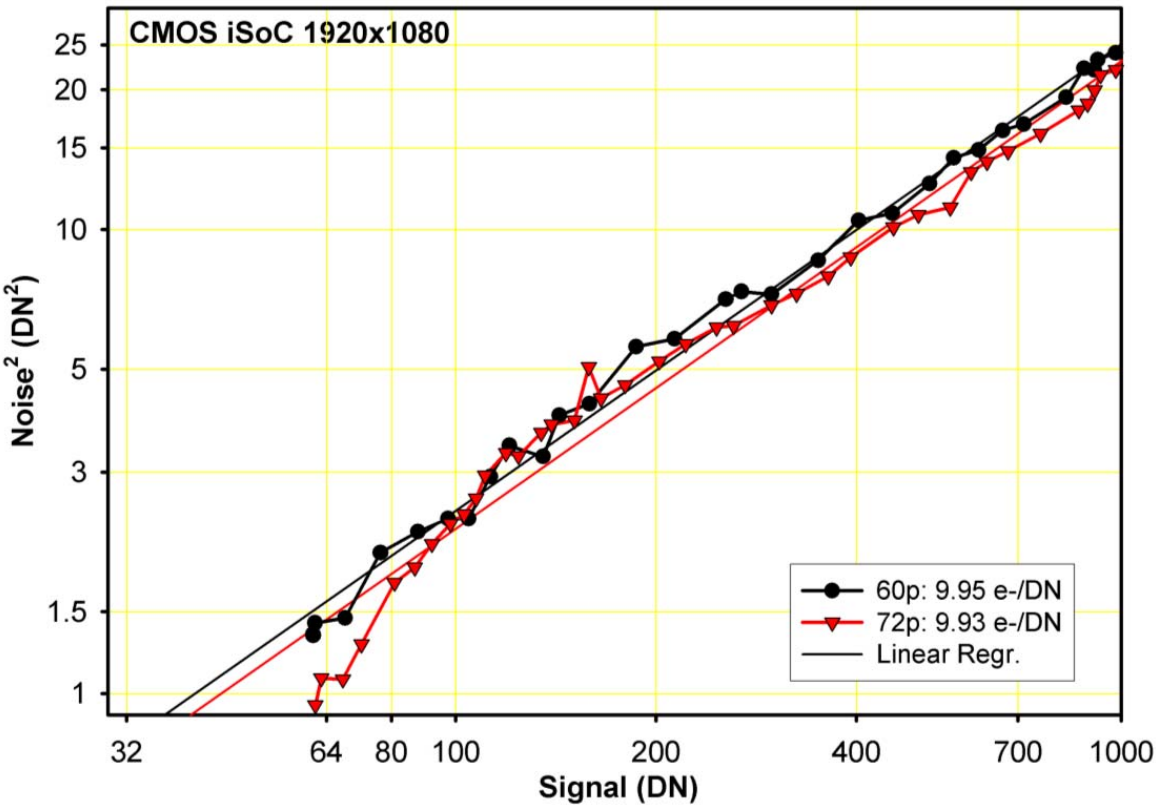

Figure 10. Measured noise^2 vs. signal for HDTV CMOS at 1080/60p and 1080/72p.

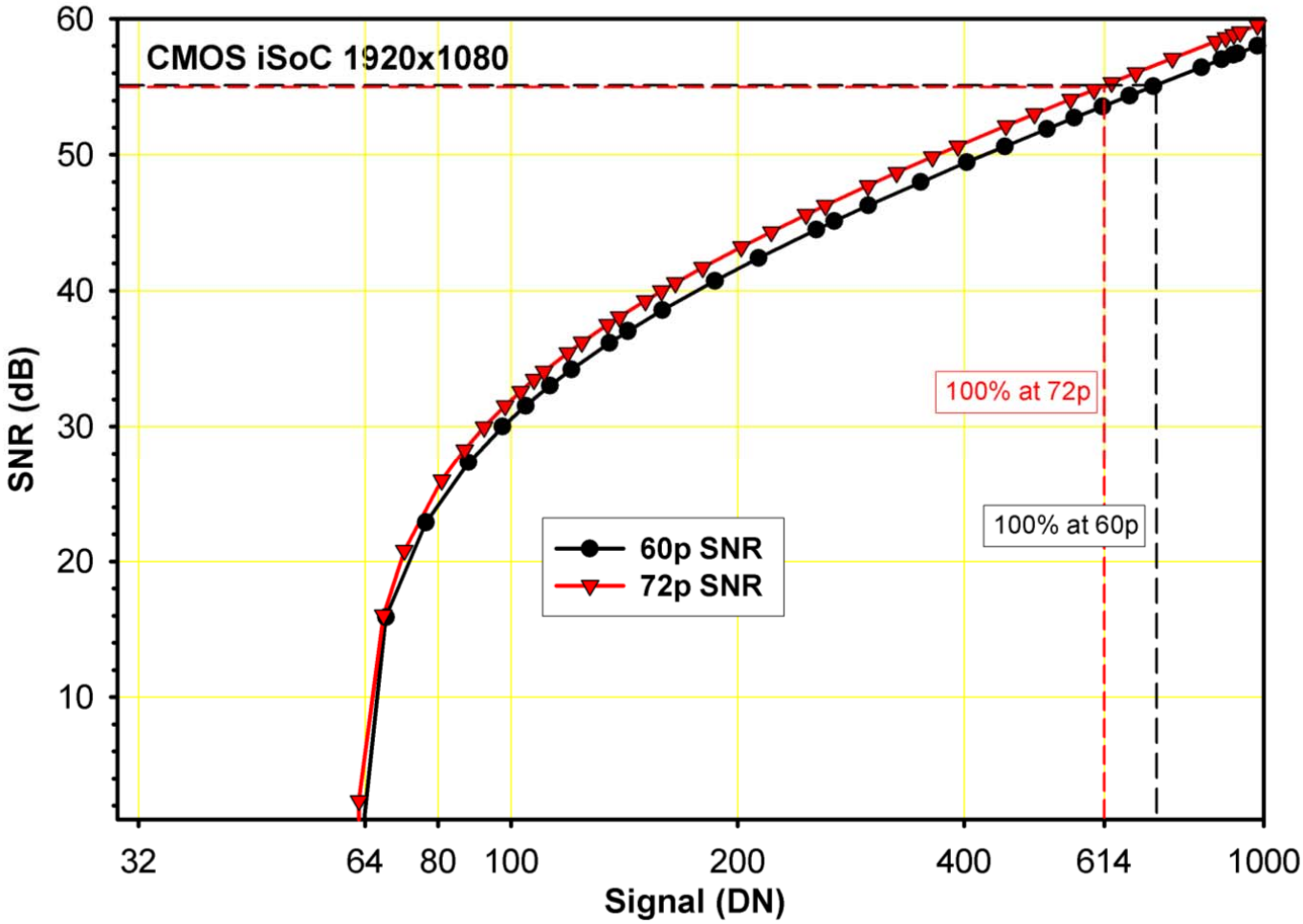

Figure 11. Measured S/N ratio vs. signal for HDTV CMOS at 1080/60p and 1080/72p.