Annu. Rev. Nucl. Part. Sci. 2003. 53:263–300
doi: 10.1146/annurev.nucl.53.041002.110431

DEVELOPMENTS AND APPLICATIONS OF HIGH-PERFORMANCE CCD AND CMOS IMAGING ARRAYS

James Janesick[1] and Gloria Putnam[2]
[1]Sarnoff Corporation, Advanced Sensors Group, 4952 Warner Avenue, Huntington Beach, California 92649, email: cmosccd@aol.com; [2]Eastman Kodak Company, 1999 Lake Avenue, Rochester, New York 14620-2010, email: gloria.putnam@kodak.com

Key Words charge coupling, hybrid sensors, digital photography

PACS Codes 85.60.Gz, 42.30.−d, 07.68.+m

■ **Abstract** For over 20 years, charge-coupled devices (CCDs) have dominated most digital imaging applications and markets. Today, complementary metal oxide semiconductor (CMOS) arrays are displacing CCDs in some applications, and this trend is expected to continue. Low cost, low power, on-chip system integration, and high-speed operation are unique features that have generated interest in CMOS arrays. This paper reviews current CCD and CMOS sensor developments and related applications. We compare fundamental performance parameters common to these technologies and describe why the CCD is considered a mature technology, whereas CMOS arrays have significant room for growth. The paper presents custom CMOS pixel designs and related fabrication processes that address performance deficiencies of the CCD in high-performance applications. We discuss areas of development for future CCD and CMOS imagers. The paper also briefly reviews hybrid imaging arrays that combine the advantages of CCD and CMOS, producing better sensors than either technology alone can provide.

CONTENTS

1. INTRODUCTION

In 1969, researchers at Bell Laboratories invented a technology that would change the course of imaging in applications ranging from astronomy to factory automation (1). That technology was a solid-state process called charge coupling, which literally overnight evolved into the charge-coupled device (CCD). At that time, the CCD was conceived as an electronic analog of the magnetic memory bubble device. In a CCD, a bit of information is represented by a packet of charges that is stored in the depletion region of a metal oxide semiconductor (MOS) capacitor. Manipulating voltages on the gates of the capacitors allows charge to move from one capacitor to the next—hence the name charge coupling. Charge is transferred to a charge detection amplifier or "sense node," at which an output voltage is generated in proportion to the charge packet size. For CCD imagers, charge packets are photoelectrically induced and collected beneath an array of gates that form individual picture elements or "pixels." Like the magnetic bubble device, the CCD is a serial device, i.e., charge packets are transferred and read one at a time.

Amazingly, charge coupling is completely noiseless and the transfer of charges between capacitors is almost perfectly efficient, yielding textbook performance. High-performance CCD imagers can read charge packets with only one-electron uncertainty (root mean squared, rms) and can transfer charge between pixels with 99.99995% efficiency. The combination of near-ideal performance and straightforward implementation in an application as important as imaging makes charge coupling a highly successful technology. The uses of CCDs range from imaging the living cell to imaging the edge of the universe.

Although the CCD is considered a mature technology today, advancements in performance, fabrication, and implementation continue. For example, astronomical applications continually require larger focal-plane arrays as telescopes grow in size (now up to 10 m in diameter, with 30-m mirrors on the drawing board!). The Canada-France-Hawaii (CFH) Telescope team has recently engineered a focal plane that incorporates 360 million pixels covering an area of 30 cm on the side. Solid-state professional photography has also come of age through large CCD imagers. Studio photography using 11-Mpixel digital backs for medium-format cameras has become standard, and in 2002 a 22-Mpixel digital camera back reached the high-end photography market. For example, Figure 1 shows an image taken from an 11-Mpixel color CCD sensor. As fabrication processes are refined and yields increase, larger and larger arrays can be manufactured at reasonable cost. Reduced cost and improved performance drive CCDs into new applications. The shrinking

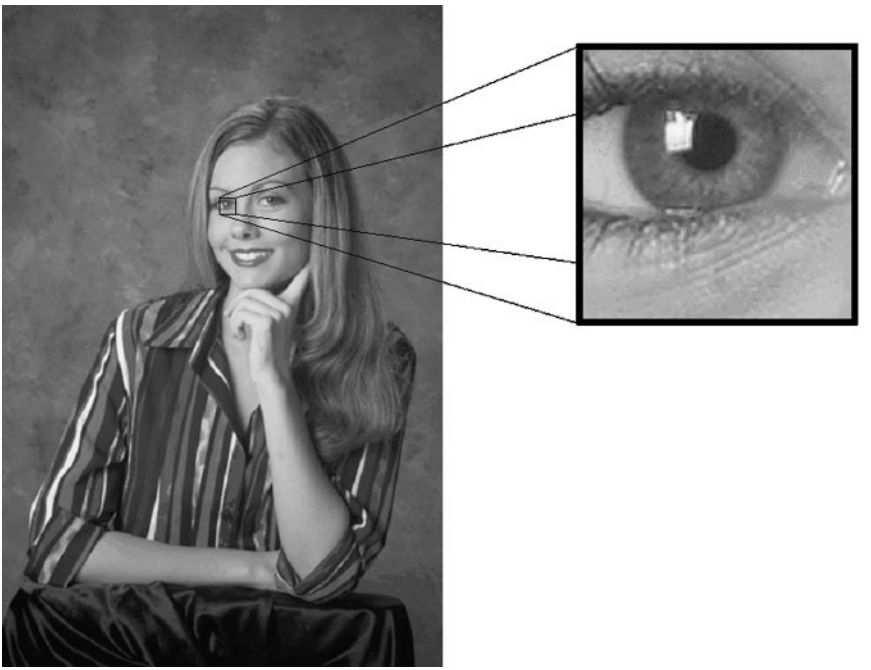

Figure 1 Portrait taken using a 35-mm format, 11-million pixel CCD.

of pixel size is also driving down cost and opening new markets. Because smaller pixels cannot provide the same image quality as large pixels, this development has primarily affected consumer products, where pixel sizes of 3 μm are typical today (in contrast to the 9-μm pixel often found in studio photography or the 15-μm pixel common in astronomy). The smaller pixel increases the resolution per unit area and therefore reduces cost. This development has enabled digital still-camera manufacturers to provide photographic image quality more closely matched to film.

Because the fabrication of CCDs requires a custom MOS fabrication line, most CCDs are made by high-volume manufacturers in Japan who supply sensors for scanners, camcorders, cell phones, and digital still cameras. The cost of a 1-Mpixel CCD bundled with the necessary readout for an entry-level consumer camera is typically under $10. Lower-resolution area arrays and line-scan sensors sell for less. A small group of manufacturers cater to lower-volume scientific and industrial markets. These manufacturers compete to provide the highest possible performance. A 4 K $\times$ 4 K back-illuminated CCD with excellent imaging quality can carry a price tag up to $120,000.

After dominating digital imaging for over 20 years, the CCD is beginning to see competition. About 10 years into development, a new imaging technology is emerging: complementary metal oxide semiconductor (CMOS) image sensors.

Because these imagers can be fabricated on a modified CMOS fabrication line, their production is not limited to manufacturers with specialized process capability. As a result, dozens of new CMOS imager suppliers have emerged. CCDs are still manufactured in much larger volumes than CMOS arrays, but as CMOS sensors continue to find their way into high-volume applications such as PC cameras, cell phones, and digital still cameras, the gap will narrow. As CMOS imaging technology continues to develop, new applications will further increase demand. It has been projected that the number of CMOS imagers manufactured will triple by 2006. CMOS sensor prices are driven, in part, by competition from CCDs—at present they sell for roughly the same price.

CMOS imager–based cameras are attractive because of their low power (10–100 mW), compact size, low part count, on-chip system integration, simple interfaces, "smart" image processing features such as random pixel-addressing capability, high radiation tolerance, and low cost. But today's CMOS sensors do not deliver the image quality that the CCD can provide. Over the next decade, improvements in sensitivity, noise, pixel cross talk, and intrapixel charge transfer will probably close the performance gap. Competition in recent years between CCD and CMOS technology has been fierce, and camera engineers have struggled to decide which technology is best for their applications. Now that CMOS sensors are available commercially and finding their way into commercial products, their real performance strengths and weaknesses are emerging. Contrary to some beliefs, CCD and CMOS imagers each have clear advantages in particular areas and are expected to coexist in the future.

Today, CMOS sensors are creating new applications and already replacing CCDs in existing applications. Automotive lane tracking, collision avoidance, and passenger monitoring are new applications that require the small size and low component count provided by CMOS imager–based cameras. CMOS sensors will also begin to find applications in medical x-ray, surgery, surveillance, biometrics, and motion analysis. Cell phones, in which CCD-based cameras were first introduced, are employing more and more CMOS sensors. CMOS sensors are also displacing CCDs in some consumer, or hobby, and professional photography cameras. Canon has spearheaded this change with the introduction of cameras featuring 35-mm format CMOS sensors ranging from 3 to 11 Mpixels. Kodak has also introduced a 35-mm camera based on a 14-Mpixel CMOS imager. Requirements for particular applications, in terms of performance, level of integration, and cost, will continue to drive the selection of the appropriate technology.

This paper is divided into five major sections. Section 2 reviews what is arguably the single-most important measure of imager performance: signal-to-noise ratio. The section will set the stage for our analysis of the current state of both CMOS and CCD imagers. Section 3 discusses performance parameters that are common to CMOS and CCD imagers and their corresponding performance limits. Performance comparisons are made in the context of the four operational tasks required of an electronic imager: charge generation, charge collection, charge

transfer, and charge measurement. Section 4 discusses recent developments in CCD technology in the areas of performance and functionality and reviews new applications. Section 5 reviews CMOS developments to date, demonstrating that CMOS imaging technology is not yet mature. We show that performance is primarily being driven by pixel design. Section 6 summarizes the technological challenges that will face image-sensor designers and manufacturers in the coming years. We discuss technologies now in their infancy, including hybrid image sensors that combine the best attributes of both CCD and CMOS arrays.

2. SIGNAL-TO-NOISE RATIO AND PERFORMANCE STANDARDS

The signal-to-noise ratio (S/N) is a key parameter that describes an imager's overall performance. Every imager and imaging system strives to achieve the highest S/N possible. Figure 2 shows a 100% contrast image with sinusoidal light intensity embedded in different noise levels varying from S/N = 1 to S/N = 20. Above ~10, the image is of good quality.

It is very important to note that all sensor performance parameters reflect on and contribute to S/N. This fact can be seen in the expression for S/N for a simple camera system that is stimulated with a flat-field light source,

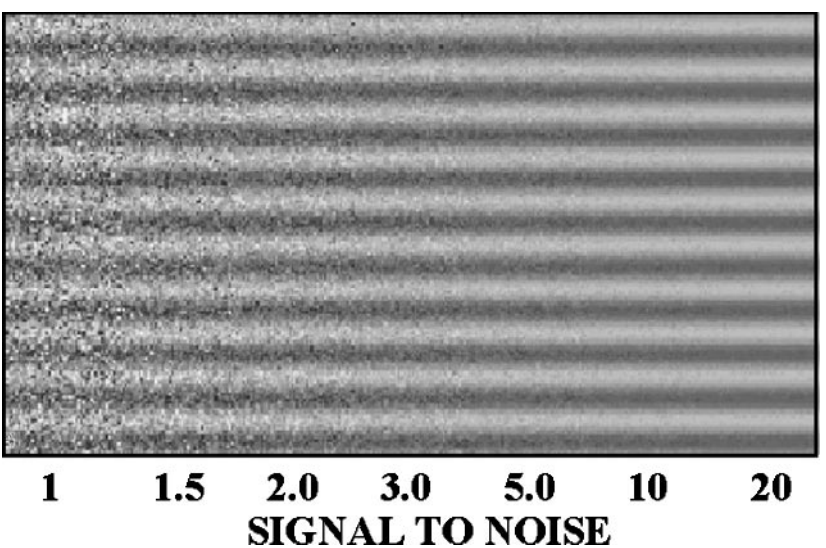

Figure 2 A sinusoidal image embedded in noise floors varying from S/N = 1 to S/N = 20.

$$\left(\frac{S}{N}\right)_{FF} = \frac{4 \times 10^{11} Q E_I P_A T_C T_L L_{UX} t_I (1 + 4f^2)^{-1}}{\left(N_R^2 + S_{SN}^2 + D_{SN}^2 + S_y^2 + Q_N^2 + R^2 + N_A^2 + P_{FPN}^2 + D_{FPN}^2 + O_{FPN}^2\right)^{1/2}}. \quad 1.$$

The numerator in this equation represents the average signal generated by the camera. QE_I is the sensor's quantum efficiency (i.e., number of interacting photons per incident photons/pixel), P_A is the pixel area of the sensor in centimeters squared, T_C is the transmission of the color filter, T_L is the transmission of the lens, L_{UX} is the light intensity given in lux (specified at 550 nm with a 10-nm bandwidth), t_I is the integration time in seconds, and f is the f-number of the lens that focuses light on the detector. The denominator is a sum of important noise terms. The first seven noise terms are temporal noise sources, whereas the last three terms represent fixed-pattern noise (FPN) sources. N_R is the pixel's sense-node reset noise, S_{SN} is the signal shot noise, D_{SN} is the dark-current shot noise, S_y is system noise, Q_N is the analog-to-digital converter (ADC) quantizing noise, R is read noise, N_A is the pre-amplifier noise, P_{FPN} is the pixel-to-pixel FPN, D_{FPN} is the dark-current FPN, and O_{FPN} is the offset FPN. Ideally, read and signal shot noise limit S/N performance for a high-performance camera system.

S/N, as defined in Equation 1, can be measured by exposing an array of pixels with a calibrated light source. S/N is determined as the ratio of the average signal to the corresponding noise generated by the pixels. For example, Figure 3 shows the S/N as a function of integration time for different lux light levels [referred to as a lux transfer curve (2)]. The analyzed imager is based on a 9.6-μm CMOS pixel with an interacting QE of 35%, a read noise of R = 5 e^- rms, and an FPN of 0.7% (S_Y, N_R, N_A and O_{FPN} noise sources are assumed negligible). The dark-current generation rate for the pixel is 30 pA/cm^2 (300 K) with a corresponding dark FPN of 30%. The array is illuminated with a green 550-nm flat-field light source calibrated in lux (1 green lux = 4×10^{11} photons/cm^2-sec) without a lens or color filter. Three S/N regimes are labeled. For low-light levels, S/N is limited by read noise and increases proportionally with the signal. For midrange signals, the S/N is limited by shot noise and increases by the square root of the signal. For high signals, S/N is dominated by FPN noise. The dashed curves show S/N performance when pixel nonuniformity is removed through "flat fielding," thereby achieving full shot noise–limited performance (1). Note that S/N becomes constant when the dark or light FPN noise dominates. For example, the 0.1 lux curve barely maintains high S/N because of high dark-current noise at the operating temperature (320 K). Pixel-sensitivity FPN dominates the other lux curves at 20,400 e^- (i.e., $1/0.007^2$). Saturation takes place in the plots where the S/N suddenly drops. The useful imaging dynamic range (IDL) is also labeled for S/N > 10. For example,

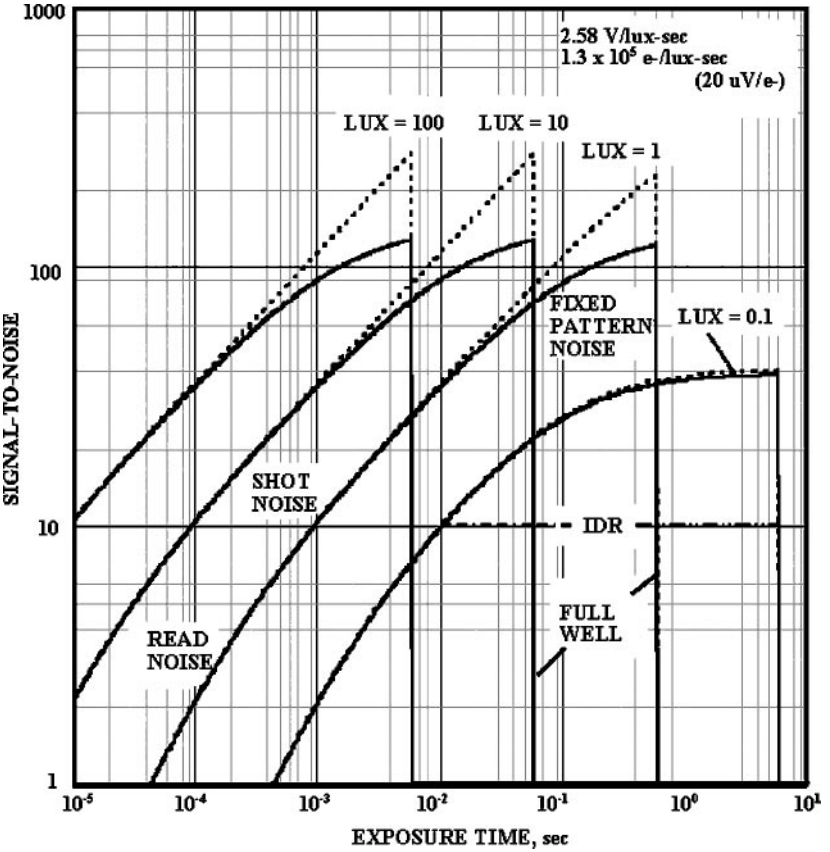

Figure 3 Lux transfer curve.

the 0.1 lux curve produces an IDL of 160 (44 db). The sensor analyzed exhibits a sensitivity of 1.38×10^5 e^-/lux-s (or 2.58 V/lux-s, assuming a sense node sensitivity of 20 μV/e^-).

Equation 1 is referred to as flat-field S/N because it corresponds to a sensor that is illuminated with a uniform light source. Real images exhibit lower S/N than is shown here. Low-contrast, high–spatial-frequency scenes are especially sensitive to S/N deficiencies. This characteristic can be demonstrated by calculating S/N for a sinusoidal image as in Figure 2. The signal is simply defined by the rms of the sine wave amplitude, $S_{RMS}(e^-)$ [i.e., 0.35 multiplied by the peak-to-peak amplitude, $S_{P-P}(e^-)$]. The noise term comprises three noise sources. The rms shot noise (σ_{SHOT}) of a sine wave (rms e^-) varies as the square root of the amplitude:

$$\sigma_{SHOT} = S_{P-P}^{1/2}. \qquad 2.$$

The FPN (σ_{FPN}) of the sine wave (rms e^-) also varies with amplitude as

$$\sigma_{FPN} = 1.2 S_{P-P} P_N, \quad 3.$$

where P_N is the pixel nonuniformity factor (approximately 0.01 for CMOS and CCDs).

The third component is referred to as pedestal noise (σ_{PED}). The pedestal is the offset light level on which the sine wave rides. Pedestal noise is composed of shot and FPN:

$$\sigma_{PED} = [S_{PED} + (P_N S_{PED})^2]^{1/2}, \quad 4.$$

where S_{PED}(e-) is the level of the pedestal (e^-).

From Equations 2–4, the overall S/N for a sine wave riding on a pedestal is given by

$$\left(\frac{S}{N}\right)_{SINE} = \frac{S_{RMS}}{\left(\sigma_{SHOT}^2 + \sigma_{FPN}^2 + \sigma_{PED}^2\right)^{1/2}}. \quad 5.$$

Figure 4 shows the S/N as a function of the sine-wave signal (rms) and pedestal level using Equation 5 and assuming FPN of 1%. The dotted lines represent S/N without FPN, whereas the solid lines include shot and FPN. In the absence of a pedestal, note that the S/N becomes limited at a value of 29 for high signal levels. Higher S/N is only possible when pixel nonuniformity is removed.

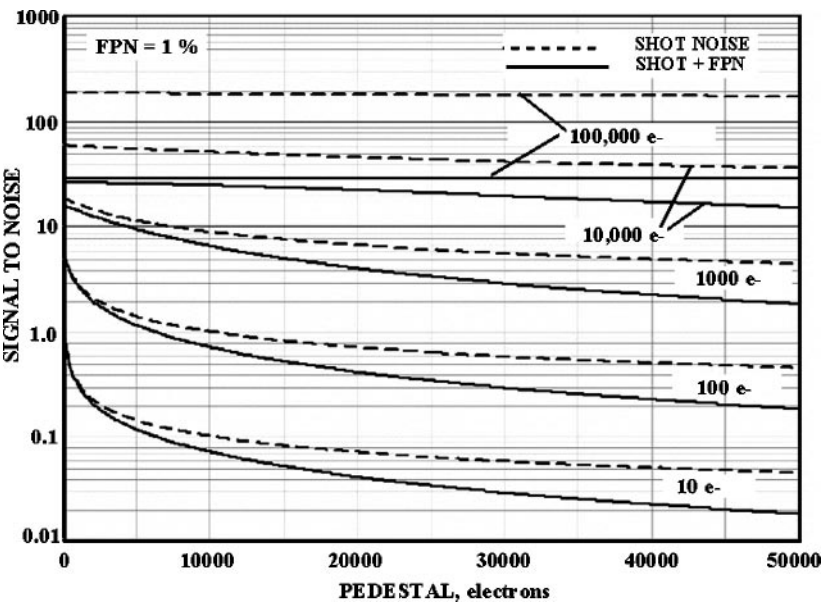

Figure 4 Sine wave S/N as a function of pedestal level and signal level.

For small pedestals, S/N is determined by the noise produced by the sine wave itself (i.e., varying shot and FPN). When pedestal noise becomes dominant, the curves become flat, yielding a S/N of

$$\left(\frac{S}{N}\right)_{SINE} = \left(\frac{S}{N}\right)_{FF} C_{SCENE} MTF(f) \qquad 6.$$

where $(S/N)_{FF}$ is the flat-field S/N at the pedestal level, and C_{SCENE} is the contrast of the scene defined as $S_{RMS}(e-)/S_{PED}(e-)$. MTF(f) is the modulation-transfer function related to the camera system at a spatial frequency of interest (cycles/millimeter). Defined this way, Equation 6 reduces to

$$\left(\frac{S}{N}\right)_{SINE} = \frac{MTF(f) S_{RMS}}{S_{PED}^{1/2}}. \qquad 7.$$

The S/N of a real scene is more accurately modeled by the sine wave on the pedestal than by the flat field. Whether imaging a star against the night sky or the vertical lines of a bar code, the scene content is on top of a background, and it is the contrast between the two that is of interest. The higher the spatial frequency of the content, the more difficult it is to separate the contrast from the noise. This is one reason why higher-resolution imagers provide better image quality.

3. PERFORMANCE FUNDAMENTALS AND LIMITS

As noted above, S/N is the figure of merit for imaging sensors. Individual performance parameters that contribute to S/N are evaluated through careful sensor characterization by using absolute test methodologies (e.g., Figure 3) (1). After a device is characterized and performance parameters are determined, one can make effective comparisons and projections to future performance. In this section, we associate performance parameters with the four operational tasks of generating an image: charge generation, charge collection, charge transfer, and charge measurement. Fundamental CMOS and CCD performance is compared for each task.

3.1. Charge Generation

Charge generation is the sensor's ability to intercept incoming photons and generate signal charges through the photoelectric effect. An ideal sensor would have 100% quantum efficiency (QE) at all wavelengths. To achieve a high response, sensor manufacturers must minimize three QE-loss mechanisms: absorption, reflection, and transmission. Absorption loss is associated with optically dead structures, typically located above, but also within, the pixel. Reflection and transmission losses are inherent to the physical properties of silicon. At certain wavelengths, reflection loss is significant. For example, at 2500 A, the reflection loss reaches a maximum of 70% for a raw silicon surface. Transmission loss occurs when incoming

photons pass through the sensor's photosensitive volume, a region typically 10 μm thick, without generating a signal charge. This problem is pronounced at very long (>7000 A) and very short (<2 A) wavelengths, i.e., the near-IR and soft x-ray section of the spectrum.

CMOS arrays experience greater absorption loss than CCDs because MOSFETs (metal oxide semiconductor field-effect transistors), which are for the most part optically dead, must be incorporated into each pixel for readout. Because CCD pixels do not require active transistors for readout, they can be constructed so that the entire pixel is sensitive, resulting in a 100% fill factor. CMOS sensors also require several metal layers to interconnect the MOSFETs. The metal bus lines are stacked and interleaved above the pixel sensors, producing an optical tunnel through which incoming photons must pass. The metal stack is typically several micrometers thick. The tunnel through which photons must pass can create a host of undesired optical effects for low f-number optical systems, including scattering, vignetting, cross talk, and diffraction problems. Using light shields above the pixel can control the problem somewhat, but this reduces the fill factor, which lowers the QE.

To improve QE performance, imaging detectors can be thinned and illuminated from the rear side, delivering spectral coverage from the soft x-ray to the near-IR (1–10,000 Å). Backside illumination eliminates absorption loss by producing a pixel with a 100% fill factor. Applying antireflection coatings eliminates reflection loss, leaving only transmission losses. Backside illumination has been used since the invention of the CCD and has produced the highest QE performance possible. Backside illumination also circumvents the CMOS optical-tunnel problem while permitting wider MOSFET bus lines for improved drive.

When backside thinning is not possible, a microlens array can significantly increase the sensor's fill factor and QE. A microlens atop each pixel increases the light-collecting ability by focusing the incoming light onto the photo-active part of the pixel. Microlenses are effective when applied to a pixel with a fill factor of less than 100%; the smaller the fill factor, the more advantageous the microlenses become. An increase in QE of 2–3X is typical for a fill factor of 40%. Microlenses are patterned using standard photolithographic techniques and are applied to almost all interline CCDs and most consumer CMOS imagers. One disadvantage of microlenses is that they enhance the already troublesome problem of lens roll-off at the edges of the field of view. With a low f-number lens in front of the sensor, the edges of the field are significantly dimmer than the center. For CCDs, the problem is less serious because the sensors' overlapping polysilicon gates and microlens are placed very close to the silicon surface.

Figure 5 shows typical QE curves for a variety of sensors. QE performance varies widely depending on the fabrication process. Image sensors exhibit intrinsically high QE at visible wavelengths. However, the sensitivity rapidly decreases in the near-IR as the silicon becomes transparent. The detector's opacity in the IR can be increased by increasing the thickness of the active silicon layer. As an

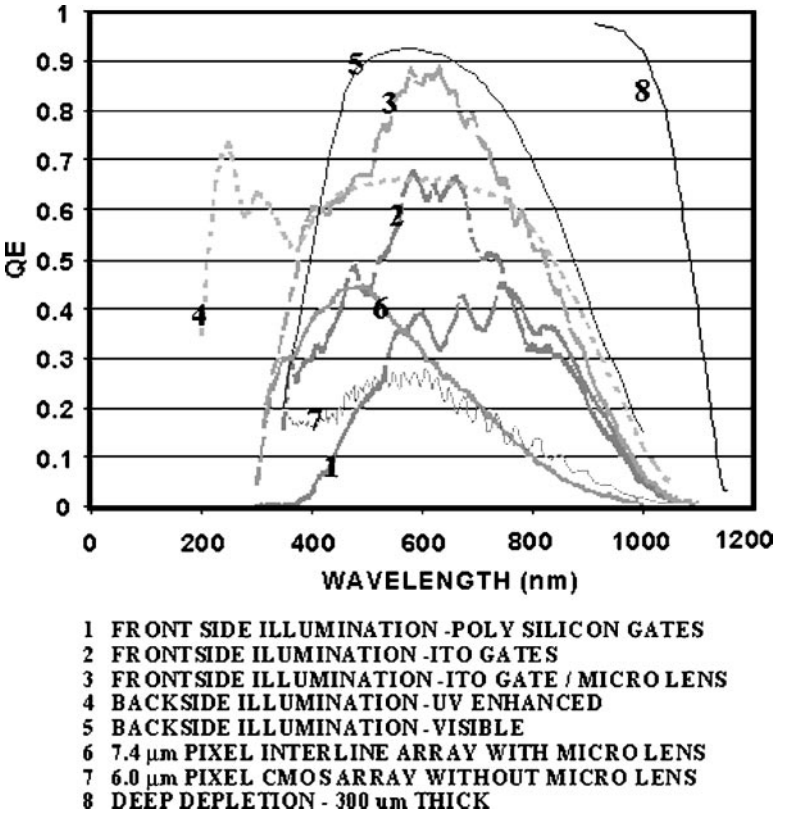

Figure 5 QE performance for a variety of CCD and CMOS sensors.

example, Figure 5 shows the QE for a deep-depletion, backside-illuminated CCD of 300 μm. A QE of 70% is achieved at a wavelength of 1 μm. The sensor has a near-IR antireflection coating (reflection for silicon is approximately 30% at 1 μm) (3).

3.2. Charge Collection

Charge collection, the second operational task of an imager, determines the sensor's ability to accurately reproduce an image. Four parameters describe this process: the number of pixels on the chip, the number of signal charges a pixel can hold (referred to as full well or charge capacity), the variation in QE sensitivity from pixel to pixel (i.e., FPN), and the ability of a pixel to collect electrons efficiently without loss to its neighbors [a parameter referred to as charge collection efficiency (CCE)]. Monolithic CCD arrays of greater than 81 million pixels have been produced. CMOS arrays with 14 Mpixels are available for

35-mm still-camera applications. Custom CMOS hybrid arrays on the drawing board have 400-Mpixel counts. Charge capacity for CMOS and CCD arrays is comparable: ∼5000–10,000 $e^-/\mu m^2$ of active area. The FPN problem originates from slight size variations in the pixel's geometry induced during fabrication. It is interesting to note that typical CCD and CMOS arrays exhibit an FPN of ∼1% of the average signal level.

CCE is a critical parameter because it defines the spatial resolution of the detector. Ideally, signal electrons generated in the silicon should remain in the target pixel. Thermal diffusion and weak electric fields within a pixel's active volume cause signal electrons to wander into neighboring pixels, creating cross talk and related modulation-transfer function loss, and poor color performance. Therefore, it is important that the charge-collecting electric fields below the photo region of a pixel be sufficiently deep to match the depth at which the incident photons interact. Pixel cross talk is most conspicuous for near-IR and soft x-ray photons that penetrate deep into the sensor, where very weak electric fields exist. Cross talk between pixels increases dramatically for reduced pixel size.

CCD manufacturers have minimized charge diffusion by using high-resistivity silicon wafers and high-voltage clocking to exploit the fact that the electric field depth varies as a function of the square root of resistivity and applied voltage. Electric fields typically extend to 7–10 μm, allowing visible and near-IR spectral coverage (4000–11000 Å). In contrast, CMOS arrays show relatively poor CCE performance because standard foundry processes use low-resistivity silicon wafers. Low-resistivity material is necessary to prevent cosmic-ray/radiation-event-triggered CMOS circuit latch-up and ground-bounce problems associated with support CMOS circuitry. Also, the low-voltage operation inherent to CMOS reduces electric field depth. This problem is becoming severe because operating voltages decrease proportionally to feature size. For example, state-of-the-art 0.18-μm CMOS processes operate at only 1.8 V. Typical electric-field depths extend only 1–2 μm, limiting spectral coverage to <6500 A (refer to Figure 10). Silicon resistivity is more critical for backside-illuminated devices because thinning processes are limited to a minimal thickness of ∼8 μm. Therefore, high-resistivity material is essential to generate collecting fields throughout the thinned material, especially for the blue portion of the spectrum (referred to as fully depleted pixels).

3.3. Charge Transfer

The third function, charge transfer, is critical to CCD operation. For very large CCD arrays, a small charge packet of a few electrons may need to transfer through several centimeters of silicon before reaching an output amplifier. The required efficiency can be achieved only when the signal channel is devoid of electron traps induced by flaws in design, processing, and the silicon itself. Charge transfer for some high-performance CCD applications must be 99.9999% efficient per pixel transfer. Such high performance makes the CCD extremely sensitive to high-energy

radiation sources (e.g., protons) that damage the silicon lattice and induce electron traps (1). CMOS pixels are directly addressable and read out, and thus, for the most part, are not as sensitive to bulk trap problems as CCD pixels. However, charge-transfer CMOS pixels can still encounter some difficulty. The ability of a sensor to transfer charge depends on the potential difference and electric-field strength between regions of transfer. Low-voltage operation has made this function challenging for CMOS, even though only a single transfer is typically required.

3.4. Charge Measurement

The last operation in generating an image is measuring the signal charge. Readout for CMOS and CCD imagers is fundamentally the same. Both imagers use a sense-node capacitor to convert charge to voltage and a source-follower MOSFET amplifier to buffer this output voltage. Designers have made the sense capacitance extremely small (on the order of a few femtofarads). The smaller the capacitance, the greater the S/N of the output-voltage. In addition, engineers optimally design the source-follower geometry to minimize noise from random fluctuations in the current flowing through the transistor. Source-follower read noise for both technologies is nearly the same (i.e., a few noise electrons). CCD cameras are usually limited by this noise source.

CMOS designers must contend with other noise sources not present in CCDs. For example, the majority of CMOS pixels, with their three-transistor readout architecture, are limited by reset noise, a large noise component that is generated when the pixels are reset. In contrast, the serial readout nature of the CCD allows reset noise to be completely removed by correlated double sampling (CDS) (1). Custom charge-transfer pixels that operate like CCDs are being designed to deal with this problem. Laboratory tests of CMOS charge-transfer pixels have demonstrated a readout of a few noise electrons (4). However, such noise levels are difficult to achieve in practice because low-noise, on-chip circuitry design is difficult. For example, CCD cameras reject white noise by off-chip capacitors that control the electrical bandwidth. CMOS designers do not have this luxury because adding filters would exceed the size limit of the chip. Hence, CMOS circuits usually work under open-bandwidth conditions, resulting in a noisier device. CMOS sensors must also contend with numerous electrical ground-bounce noise problems generated by on-chip timing logic and ADC circuitry. These system noise sources are difficult to control and often limit the sensor's noise floor above the reset noise level (tens of electrons). Optimizing on-chip ADC circuitry without quantizing or bit-weighting noise is also a problem (1). CMOS is currently limited to 12 bits of resolution. In contrast, CCD engineers can readily employ low-cost, off-the-shelf 16-bit ADCs to cover the necessary dynamic range without gain states or quantizing noise problems.

Thermal dark current, and associated noise, is also considerably higher for CMOS arrays. CCDs have employed custom processes, such as multipinned-phase (MPP) technology, to reduce dark current to very low levels (<10 pA/cm^2 at 300 K)

(1). CMOS sensors typically show levels of 100–2000 pA/cm^2. They require custom designs and fabrication techniques to compete with CCDs. Both CCD and CMOS arrays show thermally generated dark spikes, or hot pixels, which create "salt and pepper" in an image. The problem is most pronounced at high operating temperatures or when long exposures are taken. Metal contaminates in silicon wafers are one cause of dark spikes. For example, a single gold atom lodged in the silicon lattice can create a noticeably large bright pixel. Metal-contaminant dark spikes are more sensitive to temperature than normal forms of dark current.

For high-speed applications, CMOS arrays can achieve much lower noise levels than CCDs. The problem of single-channel readout for the CCD makes CMOS attractive for high-speed applications. For example, it is difficult for CCDs to maintain competitive noise levels for high-definition television (HDTV) imagers, which require pixel rates >70 Mpixels/s. In contrast, CMOS arrays can comfortably read lines at this rate with much lower noise. Some types of CCD employ multiple channels for low-noise, high-speed readout. However, this strategy is complex and expensive. In addition, multiple-readout amplifiers result in seam noise problems because gain and offset differences between channels show up in the image as a checkerboard pattern. This imaging artifact is very difficult to remove (1).

4. CCD ARRAYS

CCD development continues to advance device performance and functionality primarily for commercial applications. Several factors motivate these developments. Competitive pressure from CMOS sensors is driving CCDs toward lower power, smaller size, and higher functionality. Other advancements entail increasing QE sensitivity for the ultra-small pixels to reduce fabrication cost and to increase resolution for the growing digital still-camera market. A variety of pixel and readout architectures have evolved to meet the demands of a broad range of applications. Scientific CCD manufacturers continue to make improvements in both sensitivity and noise, but performance, for the most part, has reached theoretical limits. Below, we briefly review popular pixel architectures and how they are integrated into an imager for specific readout requirements. Performance, recent developments, and new applications for these architectures are also discussed.

4.1. CCD Pixel Architectures

Two primary pixel designs are currently employed in CCDs: (*a*) the traditional CCD pixel, which consists of multiple metal oxide semiconductor (MOS) gates, and (*b*) the pinned-photodiode (PPD) pixel.

4.1.1. TRADITIONAL PIXEL The traditional CCD pixel is not only the earliest implementation but also the simplest to manufacture and use (1). In order to collect and transfer photogenerated charge and maintain sufficient isolation from

neighboring pixels, the pixel is made of several MOS capacitors or "phases"—usually two to four. During image capture, one or more phases are biased high to collect charge (called the collecting phases), while a lower voltage is applied to the remaining phases (called barrier phases) to confine charge within the collection phases. The voltages to the gates are manipulated in sequence to transfer charge from pixel to pixel and to an output amplifier for charge measurement. Traditional CCD pixels are used in the full-frame and frame-transfer architectures described below.

Traditional CCD pixels use doped polysilicon as the conductive gate electrode material in each phase. Polysilicon is transmissive in the green, red, and near-IR portions of the spectrum but absorbs significantly in the blue and completely in the UV part of the spectrum. Recent developments have focused on reducing the gate absorption for higher QE and good color performance. For example, backside substrate thinning and rear illumination eliminates the gate absorption problem and is the favorite process among scientific CCD manufacturers. Thin and poly-hole gate technologies increase QE sensitivity for frontside-illuminated CCDs (1).

A more recent innovation involves replacing the polysilicon with indium-tin oxide (ITO), a material used to manufacture displays. Although ITO is still highly absorbing in the UV, it is much more transmissive than polysilicon across the visible portion of the spectrum, especially in the blue. It also provides better index matching to silicon, reducing the amount of light lost, owing to reflection of the sensor surface. The advantages of increased blue response of ITOs in digital photography propelled Eastman Kodak Company in 1999 to commercialize a process in which ITO is used for one of the electrodes in each pixel (5). In 2002, Kodak introduced sensors in which light is focused onto the single ITO phase using a microlens array—the first use of a microlens on a traditional CCD pixel. This increased the QE to levels competitive with backside-thinned sensors in visible-light imaging. Today, CCDs incorporating a single ITO electrode are widely used in applications ranging from professional photography to digital radiography. Figure 5 shows the increase in QE achieved by this new architecture.

4.1.2. PINNED-PHOTODIODE PIXEL Although the traditional CCD pixel was the first, it is not currently the standard. Approximately 95% of all image sensors manufactured today incorporate a more complicated pinned-photodiode (PPD) pixel. The PPD pixel was developed for video applications, and these devices replaced broadcast cameras during the 1980s and have replaced vidicon tubes in camcorders today. This pixel is used exclusively in CCDs with the interline-transfer (IT) or frame interline-transfer (FIT) readout architectures. They are still used in video applications but are also the dominant technology in consumer digital still cameras. The PPD design has also found its way into CMOS arrays.

The PPD pixel is based on a p^+ npn structure, as illustrated in Figure 6. The p^+ layer gives the pixel its name. This highly doped but very shallow implant pins the surface of the diode to a fixed potential (typically 0 V). The pinning implant is

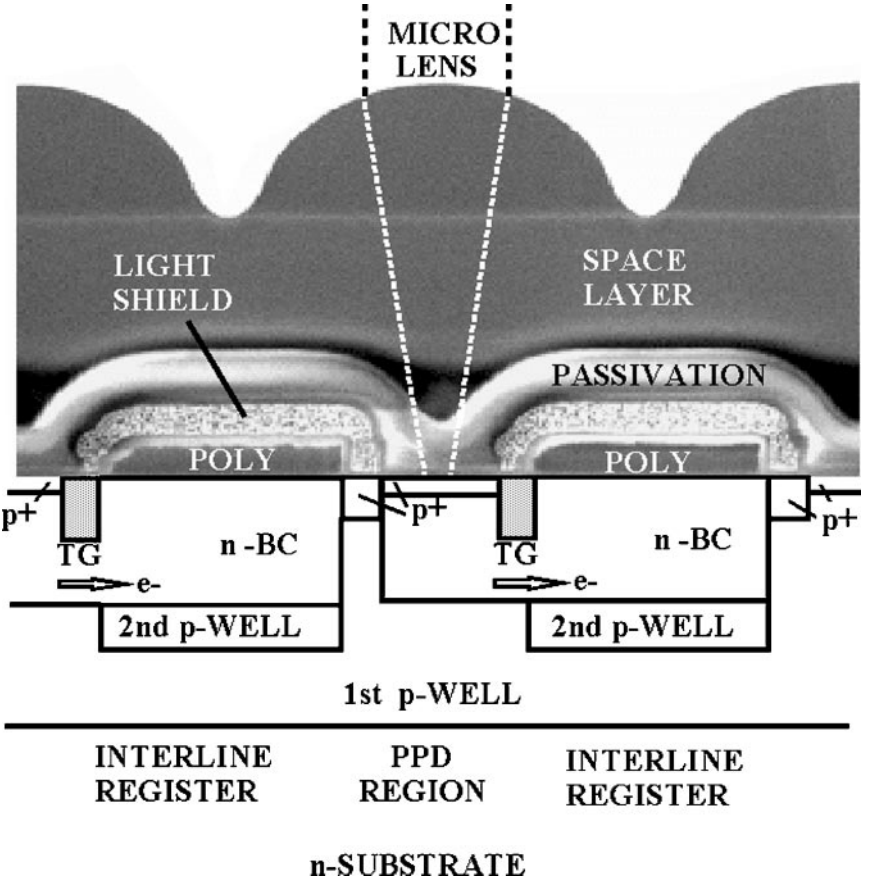

Figure 6 Scanning electron microscope cross-section of a 7.4-μm PPD pixel of an interline-transfer CCD with microlens.

necessary because the diode potential beneath the pinning implant must not change as charge is transferred by neighboring clocked gates, or severe charge-transfer inefficiency problems will result. The next layer, the n-region, provides a "buried channel," which collects and transfers signal electrons. The p-region below the buried channel provides a potential barrier to the n-substrate. The height of the barrier is adjustable and is set by the voltage applied to the n-substrate. It determines the buried channel's charge capacity (i.e., full well). As Figure 6 shows, excess charge can pass this barrier into the substrate when saturation is exceeded. This anti-blooming structure is called a vertical overflow drain because the excess charge flows vertically (as opposed to lateral anti-blooming designs typically employed in traditional CCD pixels) (1). The barrier can also be temporarily removed by applying a positive voltage to the substrate, emptying the charge in the PPD. The

ability to clear all array pixels simultaneously provides an electronic shuttering capability.

Once image capture is complete, the collected charge is moved under a transfer gate and into one of the phases of the vertically oriented CCD (VCCD) register (a readout architecture referred to as interline transfer, which is discussed below). The register portion of the pixel is similar in design to traditional pixels in transferring charge. Because this portion of the pixel is not used for image capture, it is covered with a metal light shield, as shown in Figure 6. Once signal charge has been relocated to this region, a process that requires only a few microseconds, the PPDs can begin to collect charge for the next image, providing an uninterrupted video stream. A p-doped stripe under the vertical register isolates charge from the n-substrate, so that the register does not empty when the electronic shutter pulse is applied. Because there is no poly gate covering the PPD, it does not exhibit absorption loss as does the traditional CCD pixel. A microlens is usually employed to focus incoming light onto the PPD to reduce the QE loss associated with the opaque vertical register (as shown in Figure 6).

The development of smaller and smaller PPD pixels is driven by the combination of high resolution and low cost required by the fast-growing digital still-camera market. In the next few years, pixels approaching 2 μm on a side will appear in commercial CCDs. In order to maintain image quality, reductions in pixel size must be accompanied by improvements in sensitivity. CCD manufacturers have a proven track record in this area. Whereas a 30,000-lux scene was required to capture a quality color image in 1975, today less than 1 lux is required—an improvement of over 10,000 times. Gains in sensitivity have been focused in three main areas: (*a*) increased light collection efficiency, or higher QE, (*b*) increased charge-to-voltage conversion gain, which lowers noise, and (*c*) reduced thermally generated dark current. Over the past five years, developments have focused primarily on increasing QE. Significant improvements have been demonstrated by optimizing the shape and position of the microlens above the PPD, reducing the gaps between the microlenses of adjacent pixels, shaping layers under the microlens to act as a second lens to better focus light into the photodiode aperture, and reducing the thickness of the pigment material used for the color filter array. A typical 4-μm pixel today, with a color filter and an optimized microlens, achieves a peak QE in the green channel approaching 50%.

It is also possible to substitute a phase of a traditional pixel with a PPD-like structure. Virtual-phase and open-pinned-phase (OPP) CCD arrays are constructed in this manner (1). Like a PPD, these pixels are defined by several implants that control the potential of the signal-carrying channel. Usually half of the pixel is pinned without a gate electrode, yielding good UV response. For applications that require high UV sensitivity, these pixel types provide an alternative to more costly backside-illuminated sensors. Virtual-phase CCDs are being used extensively in reticle and wafer inspection applications that require sensitivity at 248 nm and will continue to be useful as inspections move to 157 nm.

4.2. Readout Architectures

Four popular readout architectures have evolved since the invention of the CCD: (*a*) the full-frame CCD (FF CCD), developed primarily for scientific applications with long exposure times; (*b*) the frame-transfer CCD (FT CCD), a modified version of the FF CCD that provides some shuttering functionality; (*c*) the interline-transfer CCD (IT CCD), which provides full electronic shuttering as required for video capture and still captures of moving objects; and (*d*) the frame-interline transfer CCD (FIT CCD), which is used primarily in broadcast cameras.

4.2.1. FULL FRAME (FF) The FF CCD is the sensor of choice for applications that require the highest possible QE performance without electronic shuttering. A FF CCD provides 100% fill factor without a microlens. A backside-illuminated FF CCD can achieve the highest QE and broadest spectral range (1–11,000 A). A multitude of backside applications are currently used in the near-IR, visible, UV, extreme-UV, and soft x-ray regions of the spectrum. The magnitude and wavelength of the QE peak depend on the specifics of the antireflection coating. Rear-illuminated sensors optimized for visible-light imaging typically achieve a peak QE of 90%, as shown in Figure 5.

In astronomy and many other scientific imaging applications, minimizing noise is more important than fast readout. On-chip amplifiers designed for FF CCDs contribute as little as 1 e^- noise at a data rate of 50 kpixels/s. Also, the high dynamic range required for astronomical imaging necessitates large pixels—a 24-μm pixel can hold over one million electrons. In order to keep the voltage swing at the output of the device reasonably small for large signals, the device is paired with an output amplifier with low charge-to-voltage conversion factors, typically 2–4 μV/e.

FF CCDs are also popular for applications that use a pulsed light source for still-image captures. For example, in direct digital radiography, patients are exposed to pulse x-rays and the transmitted x-rays hit a phosphor screen where they are converted to visible-light photons, which are detected by the CCD. The x-ray source is only powered during image capture, so there is no smearing during readout. Because the patient being x-rayed is usually much larger than the CCD, fiber tapers and other optical methods are used to match the desired x-ray image size to the CCD size. The inevitable result is transmission loss through the system. This makes high QE sensitivity critical to minimizing patient x-ray dose.

The FF CCD is also used in studio photography, where smearing during readout is eliminated by the use of a mechanical shutter. Because cooling is usually not available in a digital still camera, photography also benefits from the very low dark current of FF CCDs. The lowest dark currents in the industry have been achieved with FF CCD technology—as low as a few pA/cm^2 at 25°C, which for a 9-μm pixel corresponds to a dozen electrons per second. Medium-format digital cameras use FF CCDs with resolutions up to 22 million pixels. These cameras retail for between $10,000 and $35,000, but the digital recording

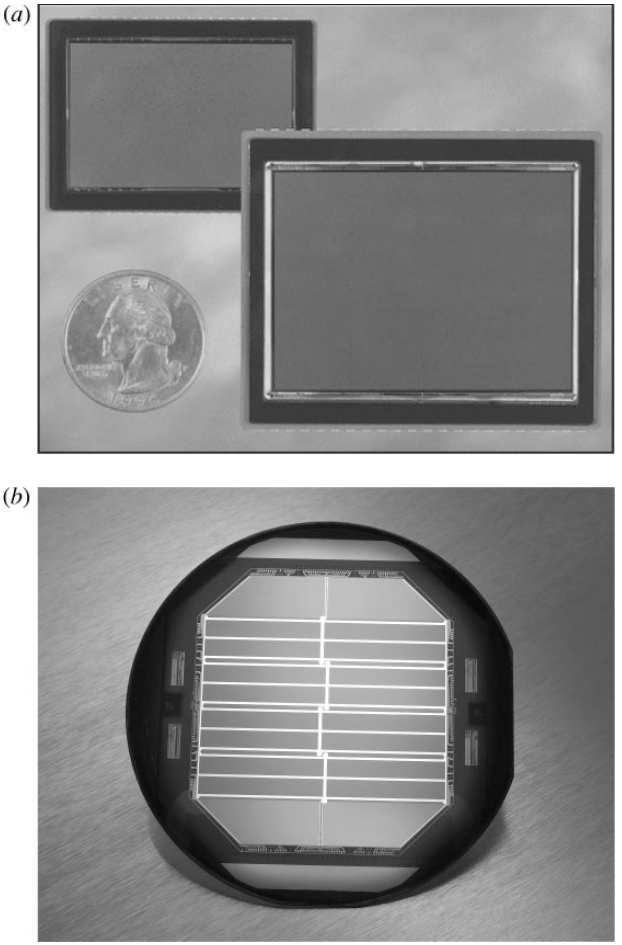

Figure 7 (*a*) An 11-Mpixel, 35-mm–format interline-transfer CCD and a 22-Mpixel, full-frame CCD used in medium-format camera backs. (*b*) Four-inch wafer-scale, backside-illuminated frame-transfer CCD with 5 million pixels.

greatly increases productivity compared to film. Figure 7*a* shows the 22-Mpixel FF CCD used in medium-format cameras alongside an 11-Mpixel, 35-mm format IT CCD.

4.2.2. FRAME TRANSFER (FT) Full-frame sensors function well when no on-chip shuttering is required because integration times are very long compared to readout times, or the illumination is controlled by pulsing or with a mechanical shutter. When more shutter control is required on chip, the full-frame architecture is modified to create a frame-transfer (FT) CCD. This is beneficial when the application

requires 100% fill factor or very high near-IR sensitivity, which the PPD pixel cannot provide because near-IR photogenerated charges diffuse into the substrate. In this architecture, half of the sensor closest to the horizontal-shift register is covered with a metal light shield. Charge is collected in the uncovered region and quickly transferred to the opaque-metal–covered storage region. Once the charge is in the storage region, it can be read out slowly while the next frame is integrated. The smearing of the image depends on the ratio between the time required to move the image to the storage region and the integration time. This smear can be further reduced by incorporating two outputs, one at the top and one at the bottom of the sensor, creating storage regions in the top and bottom sections of the imager and splits the image vertically into two halves for readout. This variant of the frame-transfer architecture is called split frame transfer. Some FT CCDs incorporate vertical overflow drains or gated lateral overflow drains, which can be used to clear the array of charge before an image capture begins.

FT CCDs are sometimes employed for remote sensing, security, and radiography but are not the dominant technology in any particular application. For example, the backside-illuminated, split frame-transfer CCD shown in Figure 7*b* is used for ground-based deep space surveillance. The device incorporates 5 million pixels with buried lateral overflow drains to maintain 100% fill factor.

4.2.3. INTERLINE TRANSFER (IT) The most widely used CCD architecture in volume applications is interline transfer, which incorporates the PPD pixel discussed above. Today, IT CCDs are dominant in a variety of applications including digital photography, machine vision, fluoroscopy, and microscopy. In recent years, the digital still-camera market has driven improvements in IT technology and has motivated the move toward smaller pixels for higher-resolution photography.

Improvements in functionality are also necessitated by increases in resolution. As digital photography employs higher resolution imagers, the need for subsampling for high frame-rate live preview has led to innovations in IT CCD architectures. It is relatively simple to subsample a color image in the addressable architecture of a CMOS image sensor but less so in the serial-readout architecture of a CCD. To subsample vertically, the gate electrodes controlling different rows of pixels can be separated and controlled independently, similar to the way in which interlaced video is read from an IT CCD. Although subsampling a color image can provide reasonable quality for live preview on a small LCD, high-quality video capture requires binning of pixels on chip in such a way as to preserve color information.

Competitive pressure from CMOS sensors has also driven significant power reduction in IT CCDs. Power consumption has been decreased in part by reducing the voltage swing required to operate the horizontal-shift register to below 5 V, in some cases 3.3 V. The combination of superior image quality, competitive pricing, and low power consumption has helped IT CCDs maintain a significant share of the cell-phone–camera market, which was expected to be dominated by CMOS sensors.

The use of IT CCDs in scientific applications is driving other improvements. In most consumer applications, the CCD is read out very quickly—pixel rates around 30 MHz are typical. For scientific applications, images are often read out more slowly in order to minimize read noise from the sensor output. But if the CCD is generating too much dark current, the benefits of slow readout can be negated. This problem is addressed in full-frame architectures by clocking schemes that invert the n-signal–carrying channel [referred to as MPP operation (1)]. In the inverted state, holes accumulate at the surface and suppress dark current. This technique can decrease the dark current by more than two orders of magnitude. In FF CCDs, the p-type substrate is a source for holes to invert the surface. Because IT CCDs are built on an n-type substrate, this source of holes is not available. Holes must come into the pixels through the p+ channel stops from a ground contact at the end of the CCD. This is also the case for back-thinned FF CCDs. The motion of holes in and out of the sensor along this high-resistivity path causes "p-well bounce," reducing the effective full-well capacity of the sensor (1). For this reason, inverted clocking was not used for IT CCDs until very recently. Bounce can be avoided by using coincident edge clocking. If one gate swings from low voltage to high voltage at the same time its neighbor swings the other way, holes need to only move from one phase to another, rather than in and out of the CCD. IT CCDs that can be clocked in this way were introduced in 2002 and demonstrate reductions in VCCD dark current by as much as 630 times at 40°C (6), reducing the need for cooling in applications such as microscopy and astrophotography.

The IT CCD architecture is also being used in new applications, including biometrics and intelligent traffic systems. In finger, palm, and iris scanning for personal identification, the shuttering capability allows an accurate capture, even when the subject is moving. Traffic cameras for monitoring intersections exploit the IT CCD's ability to take two images in quick succession, providing a measurement of vehicle speed. These systems use near-IR illumination, and the cameras block much of the visible light to reduce glare from license plates. Near-IR sensitivity is a key requirement. Some IT CCD manufacturers have increased the near-IR sensitivity of their devices by process changes that deepen the vertical-overflow drain point in the photodiode. The near-IR QE achieved by these sensors is still much less than can be achieved with an FF CCD; therefore, applications that need the highest possible sensitivity and do not require shuttering will remain loyal to FF CCDs.

4.2.4. FRAME-INTERLINE TRANSFER (FIT) Smear is a key performance parameter for IT CCDs. Because the image charge is moved into the VCCD very quickly—typically in a few microseconds—very little smear signal is generated. But if light is still incident on the CCD while the image charge is being transferred along the VCCD, some signal will leak from the photodiodes into the VCCD. This can be caused either by light finding its way under the metal light shield or by charge diffusing into the VCCD. The result is bright vertical streaks surrounding bright objects in the image or, in the case of an evenly illuminated field, an overall decrease in contrast. Digital still cameras use mechanical shutters to eliminate smear. But

streaks around bright objects are visible when a digital camera is used in preview mode, during which the shutter remains open. Smear is not a problem in most machine-vision and inspection applications because they use strobed light sources. In microscopy and traffic applications, where the sensor is constantly illuminated, smear signal can cause problems. The amount of smear signal generated by an IT CCD is specified by a number that indicates the portion of signal that leaks from the photodiode to the VCCD when the sensor is illuminated with a spot that covers 10% of the VCCD rows and the integration time is equal to the readout time. Typical specifications range from –80 dB to –100 dB. This small amount of signal can noticeably reduce the contrast in bright-field microscopy applications, where the integration time can be much less than 1 ms, or cause streaking around headlights and taillights in traffic images (see Figure 8).

The FIT CCD was developed to reduce this smear signal. The architecture is similar to the frame-transfer (FT) architecture described above. Half of the sensor closest to the horizontal-shift register is covered with a metal light shield. Once the image is captured and transferred into the VCCD, it is very quickly transferred along the VCCD into the covered storage region from which it can be read out at a slower rate. This approach minimizes the amount of time smear signal is collected and typically reduces smear to approximately –140 dB. This type of sensor is used in broadcast cameras to reduce streaks caused by bright objects and reflections in the scene.

Figure 8 Traffic-camera image showing smear-related streaking from bright lights.

5. CMOS ARRAYS

This section focuses on developments behind popular CMOS pixel architectures and their readout. As with CCDs, fundamental performance is governed by important design details of the pixel, which dictate nearly every performance parameter responsible for high S/N ratio. Unlike CCDs, CMOS pixel designs are not mature, and intense development is currently ongoing. An underlying goal for most CMOS designs is to achieve performance equal to that of the CCD.

5.1. Pixel Architectures

Four basic CMOS pixel architectures are under development: photodiode (PD) pixel, pinned-photodiode (PPD) pixel, photogate (PG) pixel, and charge-share (CS) pixel. Figure 9 illustrates the architecture of each pixel type. The PD is considered a passive pixel because signal charge is collected and read from the same region of silicon (i.e., charge is not transferred). PPD and PG architectures are called active pixels because charge is transferred from a collecting region to a readout region. The CS pixel is semiactive because charge is transferred but shared between the collecting and readout regions. Note that the PPD, PG, and CS pixels are extensions of the PD pixel. That is, the three MOSFETS used for

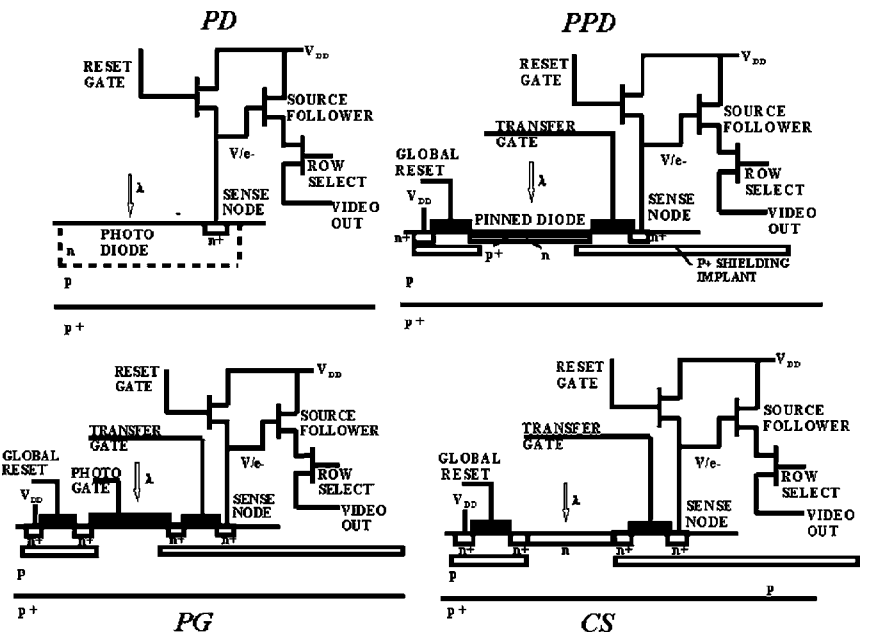

Figure 9 Photodiode, pinned-photodiode, photogate, and charge-share CMOS pixels.

TABLE 1 Performance of CMOS and CCD pixels ($V_{DD} = 3.3$ V)

Parameter	PD pixel	PPD pixel	PG pixel	CS pixel	CCD
pixel pitch, μm	8	8	8	8	8
dark current, pA/cm^2 (300 K)	200–2000	50–500	200–2000	200–2000	3–50 MPP
dark current FPN, % rms	10–30	10–30	10–30	10–30	10–30
sensitivity, V/e$^-$	15	25	25	15	5
read noise, e-rms (progressive scan)	60 rolling shutter	5	70	150	2
charge capacity, e$^-$ (hard reset)	100,000	50,000	75,000	100,000	80,000
dynamic range	1667	10000	1070	666	40,000
fill factor, %	75	65	65	65	100
interacting QE, % (400 nm)	35	35	4	35	90
pixel FPN, % rms	1	1	1	1	2
MTF, % (900 nm) (Nyquist)	10–25	10–25	10–25	10–25	50
nonlinearity, % (full well)	7	7	7	7	1

PD readout are also required to read out charge-transfer pixels. For comparison, Table 1 presents general performance parameters for each pixel type and includes a scientific backside illuminated CCD that can be used as a reference for near perfect performance.

5.1.1. PHOTODIODE (PD) PIXEL Most CMOS imagers are based on the PD pixel because of design simplicity and conformity to conventional CMOS processes. As Figure 9 shows, the PD pixel comprises four major components. The PD region is responsible for generating and collecting signal charge. Three MOSFETs are required to measure the charge on the diode: (*a*) a source-follower MOSFET, which converts signal charge to an output voltage; (*b*) a reset MOSFET, which resets the photodiode before charge is integrated on the diode; and (*c*) a row-select MOSFET, which selects a line for scanned readout.

The PD pixel has many desirable qualities, including high UV QE, good fill factor, low cross talk for visible wavelengths, high full well, anti-blooming, and absence of image lag. High read noise is a weakness because the architecture is incompatible with correlated double sampling (CDS) signal processing. As discussed below, the PD pixel uses "rolling shutter" readout, which produces a

noise level of 20–70 e^-, depending on the size of the pixel. PD pixels exhibit dark-current generation rates that range from 0.2 to 2 nA/cm^2, considerably higher than the CCDs. Dark FPN and shot noise associated with the average dark floor can have a significant affect on the S/N ratio at high operating temperatures. Many CMOS applications (e.g., automotive, machine vision, cell phone) are required to operate at temperatures of 85°C and greater. At this temperature, dark noise exceeds 100 e^-, assuming 0.2 nA/cm^2 and a 10-ms integration time. Significant reductions in dark noise are now being achieved through custom designs that isolate the PD from interfaces that tend to generate dark current. PD dark-current levels of <0.1 nA/cm^2 are being achieved experimentally with some penalty in fill factor and charge capacity. However, as the average dark current is reduced, random hot pixels become more noticeable. The problem is sometimes associated with the n^+ ohmic contact to the diode, a region where interface damage occurs. Silicon defects and impurities are also responsible for the dark spikes. Spikes become noticeable over the reset noise level at exposures of $\sim$30 ms at 300 K.

The PD pixel, like all CMOS pixels, is vulnerable to image lag if not properly reset. Image lag occurs when the reset gate is only clocked high to the supply voltage of the chip, V_{DD} (referred to as "soft reset"). During this time, the potential under the reset gate is less than the sense-node voltage by the MOSFETs' threshold voltage, V_{th} (i.e., $V_{DD} - V_{th}$). In this state, charge from the sense node can thermally cross the reset gate barrier to the V_{DD} drain region. As charge escapes, the sense-node voltage increases and results in a signal variation with time at the output of the source follower. Image lag is most noticeable when a dark scene follows a bright scene. A residual image is seen as the sense node discharges under dark conditions. Image lag for the PD pixel can be eliminated by over-driving the reset MOSFET with a clock voltage greater than $V_{DD} + V_{th}$ ("hard reset"). This clocking action will force both the sense node and potential under the reset gate to V_{DD} potential, thereby eliminating the field that would cause electrons to leave the sense node. This mode of operation is also very advantageous because it increases the pixel's well capacity by allowing the sense node to vary over a greater range (typically by a factor of two). The reason for image lag is discussed in more detail below because this phenomenon represents a major problem for CMOS imagers in general (unlike CCDs).

To achieve low cross talk and high modulation transfer function, the PD pixel must collect signal charge without sharing it with neighboring pixels. As noted above, this problem occurs when photons interact below the PD depletion region. One or more high-energy n-implants form the n-side of the diode, 4–5 μm deep (a deep n-well). Also critical is the p-depletion depth of the diode. Figure 10 shows the p-depletion depth as a function of silicon epitaxial resistivity and sense-node bias voltage. Also shown is the photon absorption depth for various wavelengths of light (absorption depth is the distance at which 63% of photons are absorbed). For example, 640-nm light penetrates 3 μm, the longest absorption depth for the visible applications. Two curves are shown for soft and hard reset operation. Note that for hard reset and 2-Ohm material (typical CMOS silicon resistivity), the p-depletion depth is only 0.8 μm. Therefore, for the depletion region to match a

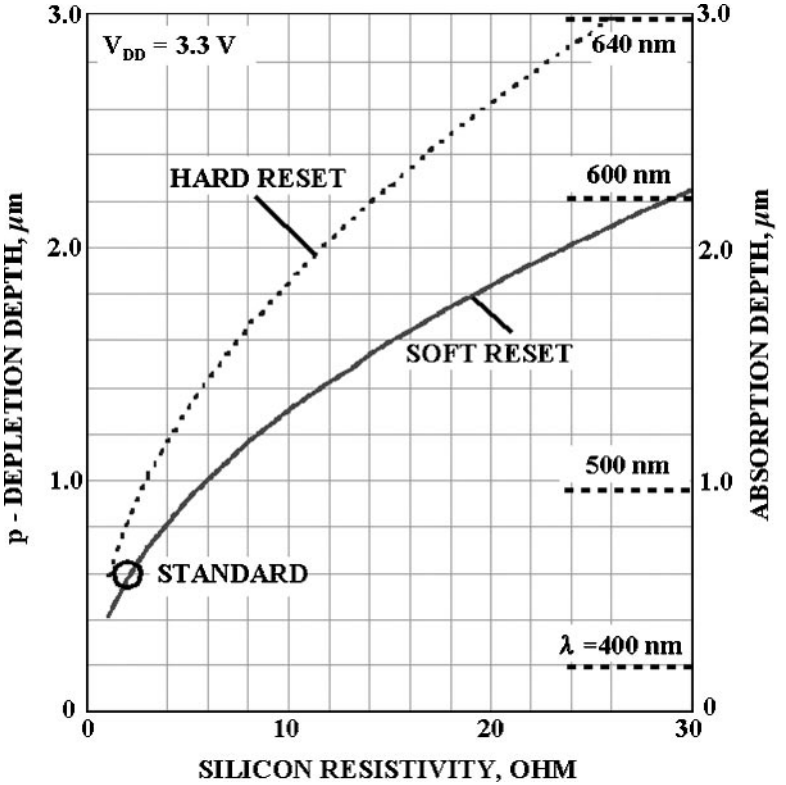

Figure 10 Depletion depth as a function of silicon resistivity for two different pixel resets.

photon absorption depth at 640 nm, the n-region of the diode must extend to at least 2.2 μm, which is readily achieved by deep n-well processing. It should be noted that many CMOS applications require near-IR performance, which demand greater depletion depths. Unfortunately, the depletion problem is becoming more serious as CMOS circuitry, pixel size, and operating potentials are further reduced.

All pixel designs, including PD pixel architecture, provide anti-blooming protection. That is, when a pixel becomes saturated by a bright light source, charge is prevented from escaping to neighboring pixels. Instead, the charge is attracted through the reset MOSFET and removed via the V_{DD} supply line. The anti-blooming feature is intrinsic to CMOS imagers and is the envy of CCD designers, who must provide such protection through custom designs and fabrication processes.

5.1.2. PINNED-PHOTODIODE (PPD) PIXEL The PPD pixel (see Figure 9) has become the choice charge-transfer CMOS pixel. The pixel is based on charge coupling, invented for CCDs. The transfer gate, when activated, transfers signal charge from

the pinned region to the sense node without image lag. A global reset gate can be used to control exposure and prevent sense-node antiblooming. For certain readout modes, reset noise can be completely eliminated by employing true CDS processing. Because the PPD region is decoupled from the sense node, very high sensitivities can be achieved (100 μV/e$^-$ is possible). This attribute in turn can potentially produce very low read noise.

The PPD pixel exhibits low dark current because holes from the pinning implant passivate traps at the Si–SiO_2 interface. Dark-current rates of a few tens of pA/cm^2 have been demonstrated. Like the PD pixel, a PPD pixel exhibits high blue QE, depending on the fill factor. For example, a PPD 9-μm pixel layout can exhibit ~80% fill factor using current CMOS design rules. For short UV wavelengths, reflection is the main QE loss mechanism. At 4000 A, ~40% QE can be achieved for the light-collecting area of the pixel.

Compared with the PD pixel, the PPD pixel has low charge capacity, which limits its use to large-pixel applications. Also, the PPD requires 2–3 custom implants, similar to the pinned CCD pixels. The implants must be accurately aligned to avoid charge-transfer problems. For example, a trap problem often develops when the p^+ pinning implant encroaches under the transfer gate during high-temperature processing, and this decreases the potential at the leading edge of the gate and obstructs the electron transfer. The fringing fields between the transfer gate and sense-node region typically do not overcome this barrier because the potential difference is very small. Although CCDs exhibit similar trap problems, they are not as vulnerable as CMOS because clock and bias voltages and corresponding fringing fields are greater and normally prevail over the trap. The process problem can be corrected by slightly offsetting (or angling) the pinning implant from the transfer gate (<0.4 μm).

5.1.3. PHOTOGATE (PG) PIXEL The PG pixel shown in Figure 9 is similar to the PPD pixel except that a PG replaces the pinned diode. Also, the gap between the PG and transfer gate is implanted with n^+ material to conductively link the region for charge transfer. The PG pixel is fabricated using conventional CMOS processes, which is an advantage over PPD technology. Charge transfer to the sense node takes place by clocking the photo and transfer gates, as in CCD operation.

Unlike the PPD, the PG pixel suffers from image lag. This disadvantage is inherent to the PG pixel because of n^+ material located between the photo and transfer gates, a region that contains a supply of free electrons. Figure 11 illustrates the timing of the charge transfer from the PG to the sense node. The potential of the n^+ region is initially defined by the transfer-gate potential. In the high state, electrons from the n^+ region can thermally cross the transfer-gate potential to the sense node, charging it negatively. The illustration shows the n^+ discharge for time periods up to 100 ms. The n^+ region leaves a potential pocket that acts as a charge trap, greatly degrading charge-transfer performance for low light-level imagery. Figure 12*a* shows the output voltage of a PG pixel, emphasizing the long time constant involved. Figure 12*b* shows a 1000-e$^-$ square wave response demonstrating how

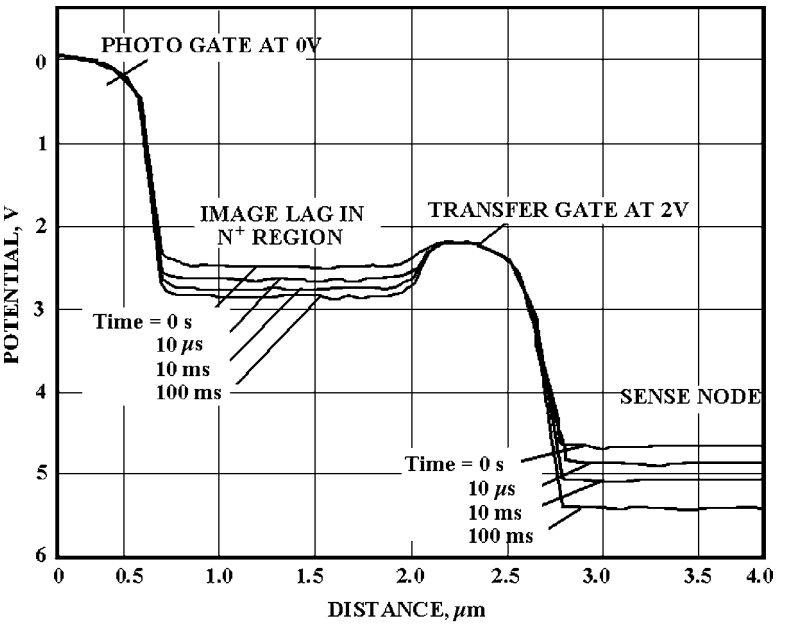

Figure 11 PISCES modeling illustration showing n^+ region responsible for image lag.

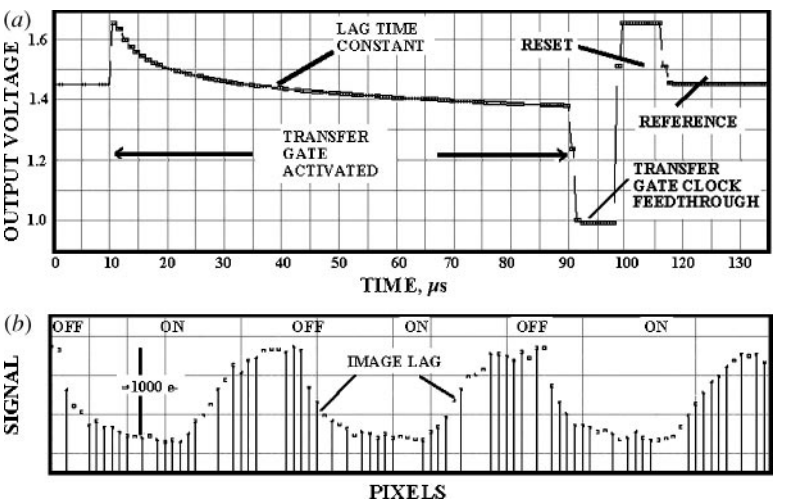

Figure 12 (*a*) Photogate, image-lag time constant and (*b*) corresponding output video in reaction to a changing light source.

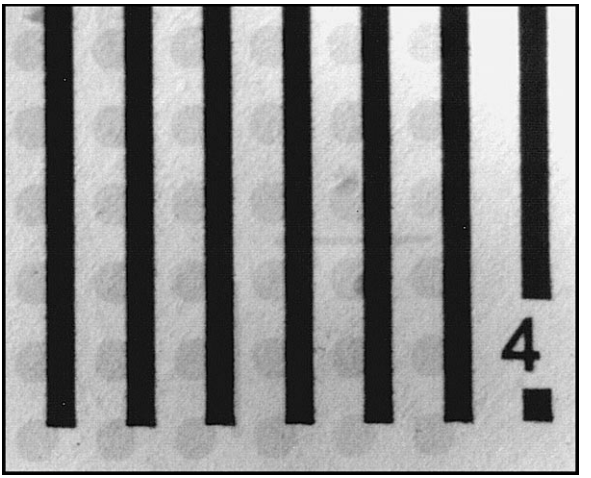

Figure 13 Image taken with a CMOS sensor with an image-lag problem.

the trap charges and discharges. The amount of trapping depends on the physical size of the n^+ region and its relation to the transfer gate. The image in Figure 13, taken with a CMOS imager, exhibits an image-lag problem. First, an image of dark round circles was taken; next, the target was switched to a bar target, generating the image shown. Image lag is present where discharge within the round circles has taken place.

Image lag is unacceptable for most applications and must be controlled. It can be eliminated for the PG pixel by simultaneously performing a hard reset to the reset, transfer, and photogates. After a hard reset, the PG and transfer gates are set to proper potentials for charge collection and subsequent charge transfer for readout. However, this operation results in high reset noise, depending on the capacitance associated with the PG. High read noise is reflected in Table 1 for this reason.

Unlike the PPD pixel, the poly-silicon PG is absorbing, which represents a major QE loss, especially for blue/UV wavelengths (only a few percent QE is achieved at 4000 A). QE performance can be enhanced by employing PGs that partially cover the photo region (similar to poly-hole CCDs). The QE problem can also be circumvented by backside illumination.

5.1.4. CHARGE-SHARED (CS) PIXEL The CS pixel (Figure 9) is similar to the PPD pixel except that the photo region is a simple photodiode without pinning. As in the PG, this arrangement makes the pixel vulnerable to image lag because of undepleted n^+ material. The problem can be eliminated, as for the PG pixel, by performing a hard reset with a penalty of reset noise associated with the capacitance

of the diode. Also, as its name implies, signal charge in the CS pixel is not totally transferred but is shared between the PD and sense-node regions. Charge sharing effectively lowers the sensitivity (V/e^-), producing additional noise (e.g., CS read noise can be >100 e^- for an 8-μm pixel). However, the CS pixel does offer high blue sensitivity, unlike the PG pixel. The noise impact becomes less worrisome as the CS pixel becomes smaller.

5.2. Readout Architectures

The three popular readout modes for CMOS pixels are rolling shutter, snap, and progressive scan. Read noise is highly dependent on the mode employed. As is shown below, progressive scan delivers the best noise performance and is, therefore, the choice readout mode for scientific applications.

5.2.1. ROLLING SHUTTER The PD pixel uses rolling-shutter readout. Charge-integration time, or the time period when the PD is allowed to discharge to photon input, is achieved by leaving the reset switch off for a specified number of line periods before a row is selected for readout. For example, minimum integration time occurs when the reset switch is inhibited one line before that row is selected for readout. Rolling shutter exhibits image smear and motion artifacts because charge integration takes place over the entire readout period (1/30 s for standard video rates). For example, Figure 14 shows a distorted image of a moving bus caused by the rolling-shutter readout.

Rolling shutter inherently exhibits high read noise because of its incompatibility with CDS signal processing. With rolling shutter, the reference and video levels are not synchronized to the same reset pulse (i.e., a reset occurs between the two samples). The resultant noise is, therefore, limited by reset noise. It is hoped that future development will bring a straightforward solution that eliminates reset noise without image lag (i.e., point images will be detected at a level equal to the read

Figure 14 Rolling-shutter image showing motion distortion.

noise achieved) (7, 8). Although true CDS cannot be achieved, it is still common practice to employ CDS processing with rolling-shutter readout to minimize FPN generated by source-follower threshold and column-to-column gain and offset differences. Because of this need, pixel reset noise increases by the square root of two after the double sampling.

5.2.2. SNAP Snap readout is used by charge-transfer CMOS pixels (i.e., PPD, PG, and CS). The readout scheme is initiated by resetting all pixels using a global reset gate or by resetting the sense node, transfer, and PGs together. Next, charge is integrated within the photo region for a given exposure period. During this time, the sense node is reset to eliminate dark current or signal charge build-up. The signal charge is transferred to the sense node for all pixels at the same time. After transfer, readout commences line-by-line using CDS-like processing (i.e., sample 1, reset, sample 2). The readout scheme is referred to as snap because exposures can be taken that are very short relative to the readout time. This allows freeze-action photography not allowed by rolling-shutter readout. As with PD operation, true CDS is not possible because there is always a reset between the two samples. This limits noise performance to reset noise.

Charge generated and collected by the sense node can cause image smear during readout, especially for moving scenes (referred to as sense-node smear). Metal light shields above the pixel prevent photons from interacting with the sense-node region. However, photoelectrons generated below the pixel's depletion region can diffuse to the sense node. The smear problem is most conspicuous for very short, near-IR exposures. Diffusion can be minimized by a shielding implant beneath the sense node and transfer gate, as shown in Figure 9. The implant gradient produces a self-induced electric field that directs signal charge into the photo region. The sense-node and read MOSFETs are normally placed in a p well that naturally reflects signal carriers.

For snap readout pixels, the sense node must also be protected from charge that leaks over the transfer gate during readout. Such "sense-node blooming" occurs when the exposure time is less than the readout time, as the sensor is still illuminated during readout. Sense-node blooming can be circumvented by biasing the global reset gate shown in Figure 9 with a potential lower than the transfer-gate potential, allowing excess charge to escape to a drain region. The problem is not experienced when a snap array is used in conjunction with a mechanical shutter or strobed illumination. Snap readout can take place as the next frame is integrated (similar to frame-transfer and IT CCDs). Here, the exposure and readout periods are equivalent.

5.2.3. PROGRESSIVE SCAN Progressive scan is another readout used by charge-transfer pixels. This mode typically employs a mechanical shutter (or pulsed light source). After exposure, charge is read line-by-line using true CDS readout (i.e., reset, sample 1, charge transfer, sample 2). The mode allows the read noise to be limited by the source-follower MOSFET. It is possible to achieve very low noise

performance, equivalent to that of the CCD in slow-scan mode (i.e., a few noise electrons) (4).

The problems of sense-node cross talk and sense-node blooming are much less serious for progressive-scan readout than for snap readout. This is because charge that diffuses to the sense node during readout is eliminated by leaving the reset switch on except when a line is read. In this case, the charge-diffusion problem translates to a QE loss, which, as mentioned above, can be minimized by a sense-node shielding implant. Progressive scan also results in low dark current because the sense node can be continuously reset until a line of pixels is read.

6. FUTURE DEVELOPMENTS

Scientific CCD development has slowed simply because CCD technology has matured. Many CCD performance parameters have reached theoretical limits. This mature state is reflected by the government's diminished interest in funding advanced CCD development and tests. It is clear that the economy of semiconductor manufacturing is pushing CCD manufacturers away from maintaining capabilities to fabricate specialty components. Hence, the number of custom CCD manufacturers is shrinking, especially in the United States. On the other hand, commercial CCD development is strong and productive. CMOS development is also very active. Dozens of CMOS design houses are developing imaging arrays for a multitude of new applications. And there are plenty of CMOS foundries to fabricate the designs. Hybrid imagers are an exciting emerging technology. Hybrid sensors combine the advantages of CCD and CMOS and outperform them both.

6.1. CCD

Consumer IT CCD development today focuses primarily on restoring performance lost to decreasing pixel size. Over the next few years, efforts to increase charge-to-voltage conversion are likely to merge with ongoing improvements in QE. Sensitivity is increased by reducing the capacitance associated with the output structure, which is enabled, in part, by decreasing feature size. Today's high-gain output amplifiers deliver around 30 μV/e-, a number that will probably double over the next few years as pixel sizes continue to shrink. In a sensor with a 2-μm pixel, corresponding to a $\sim$10,000-electron well capacity, and a 60-μV/e- charge-to-voltage, 60-dB dynamic range will be achievable at pixel readout rates around 25 MHz, making this pixel size a reasonable choice for a digital still camera. Efforts to reduce power consumption and increase functionality will also continue as CCDs face increasing competition from CMOS sensors.

Scientific full-frame CCD development will continue to focus on increasing QE and decreasing noise. Eastman Kodak Company has demonstrated front-illuminated CCDs with all electrodes composed of ITO (indium-tin oxide). The commercialization of these devices will provide QE in the visible spectrum similar to that of backside-illuminated sensors today, without the need for a microlens

array. Eliminating the microlens from the surface eliminates the angle dependence of the sensor response. It also allows the sensitivity improvement to be realized even when fibers are bonded to the surface of the sensors, as is done in many x-ray applications.

For low-light scientific applications, reductions in noise have been important. Today's scientific CCDs typically exhibit noise floors of 1–3 e^- rms. However, in cases where very few image photons are available for sensing, this noise floor cannot provide the S/N required for applications that require single visible-photon counting. The noise floor can be reduced by sampling each pixel many times, on-chip—as with "skipper" output amplifiers (1)—or off-chip, by averaging multiple frames. Another method uses charge multiplication in the CCD shift register (9). It was demonstrated many years ago that signal charge can be multiplied in the CCD shift register, before the introduction of noise by the on-chip amplifier. The multiplication physically results from impact ionization as charge packets are transferred through regions of high electric field. These regions are formed in a section added to the end of the horizontal-shift register by using a much higher drive voltage. The gain per transfer is relatively small, so this region usually consists of several hundred pixels, over which gains as high as tens of thousands can be achieved. With effective noise levels below 0.01 e^- rms, single-photon counting is possible. This technology will become useful in night-time surveillance, low-light fluorescence experiments in the biosciences, and any other application in which intensified CCD cameras are used today. Although two companies have commercialized full-frame CCDs incorporating this technology in the past three years, their widespread adoption has yet to take place.

6.2. CMOS

Significant development must take place before CMOS will exhibit performance as high as the CCD. Fundamental characterization of CMOS arrays has identified several performance difficulties. Most deficiencies are the same problems that CCD technology has solved through custom fabrication processes, pixel designs, and special operating modes. For example, custom photodiode implants will be used to meet depletion requirements and reduce pixel cross talk. Proper silicon wafer selection is also essential for proper depletion depths (i.e., resistivity and epitaxial thickness). As noted above, dark-current levels are considerably higher than in CCDs making CMOS application difficult in high-temperature environments. Observed high dark current has been linked to the damage done when metal traces are etched with high-energy plasma cycles employed during CMOS fabrication. Custom hydrogen-annealing cycles are required to passivate the damage induced. Special pixel designs that isolate high dark-current sources from signal charge-collection regions are required (11). Dynamic range for CMOS arrays can be increased through MOSFET threshold implants and high-voltage operation (e.g., hard reset clocking). Special shielding implants will reduce sense-node cross talk for charge-transfer CMOS pixels. Different CMOS processes will need to be integrated to

optimize pixel performance and support circuitry independently. For example, the tight design rules applied to reduce the size of support circuitry have a negative impact on pixel performance. This scaling mismatch is getting worse as CMOS design dimensions continue to shrink. Ideally, CMOS pixels should be fabricated with design rules that maintain high-voltage operation while support digital circuitry employs tighter rules for compactness, high speed, and low power consumption.

To compete with the CCD scientifically, high-performance CMOS imagers will need to transfer charge (i.e., charge coupling) and exhibit complete charge transfer (i.e., no image lag), and have low read noise (i.e., true CDS must be applied), low dark current (i.e., through surface inversion), low cross talk (i.e., full depletion), and high QE (i.e., backside illumination). For example, the advanced backside-illuminated CMOS PPD pixel shown in Figure 15*a* solves the depletion and QE problems (4). The pixel is based on a standard deep n-well CMOS process used to isolate MOSFETs located in a grounded p well from the substrate layer. This arrangement allows the substrate to be biased negatively ($-V_B$) and generate a depletion region that extends to the rear of the device. High resistivity and thick epitaxial silicon are also employed for very deep depletion operation, significantly increasing near-IR/soft x-ray QE performance. Figure 15*b* shows depletion depth achieved for three substrate bias conditions assuming 400 Ω-cm silicon and a photodiode bias of 3 V. Depletion depth is significantly greater compared to conventional pixels analyzed in Figure 6.

Conventional PG CMOS pixels exhibit either low noise with image lag or high reset noise without image lag. Also, the pixel exhibits high dark current because the surface under the photogate is depleted, which maximizes dark-current generation (1). To eliminate image lag, the photo- and transfer gates can be "charge coupled" without an n^+ region between the gates (referred to as charge-coupled CMOS pixels). Several different CMOS pixel designs can perform this important function (4). For example, double polysilicon overlapping gates commonly used by CCDs have been implemented and characterized. Ultra-low noise can be achieved by reading the PG pixels by progressive scan readout and true CDS signal processing. Read noise floors of 2.6 e^- have been demonstrated for test arrays (4). Dark current can be significantly reduced by inverting the surface under the photogate, which requires a custom implant under the photogate. Dark-current generation levels <10 pA/cm^2 are expected (4). A PG pixel array can also be back-thinned and substrate-biased. The combination of these design features and process technologies can potentially produce a CMOS pixel that delivers performance equal to the CCD's performance (or CCDs). Active development along these lines is ongoing.

An exciting area of development is array "stitching." CMOS lithography is limited to fabricating arrays of 20–40 mm, depending on the manufacturer's alignment capabilities. For a 10-μm pixel, this limits arrays to 2–4 K-pixel formats, less support circuitry. As is true for CCDs, many high-end applications demand larger formats. Some CMOS foundries offer stitching capability, which means multiple blocks of pixels are connected to form very large arrays, theoretically

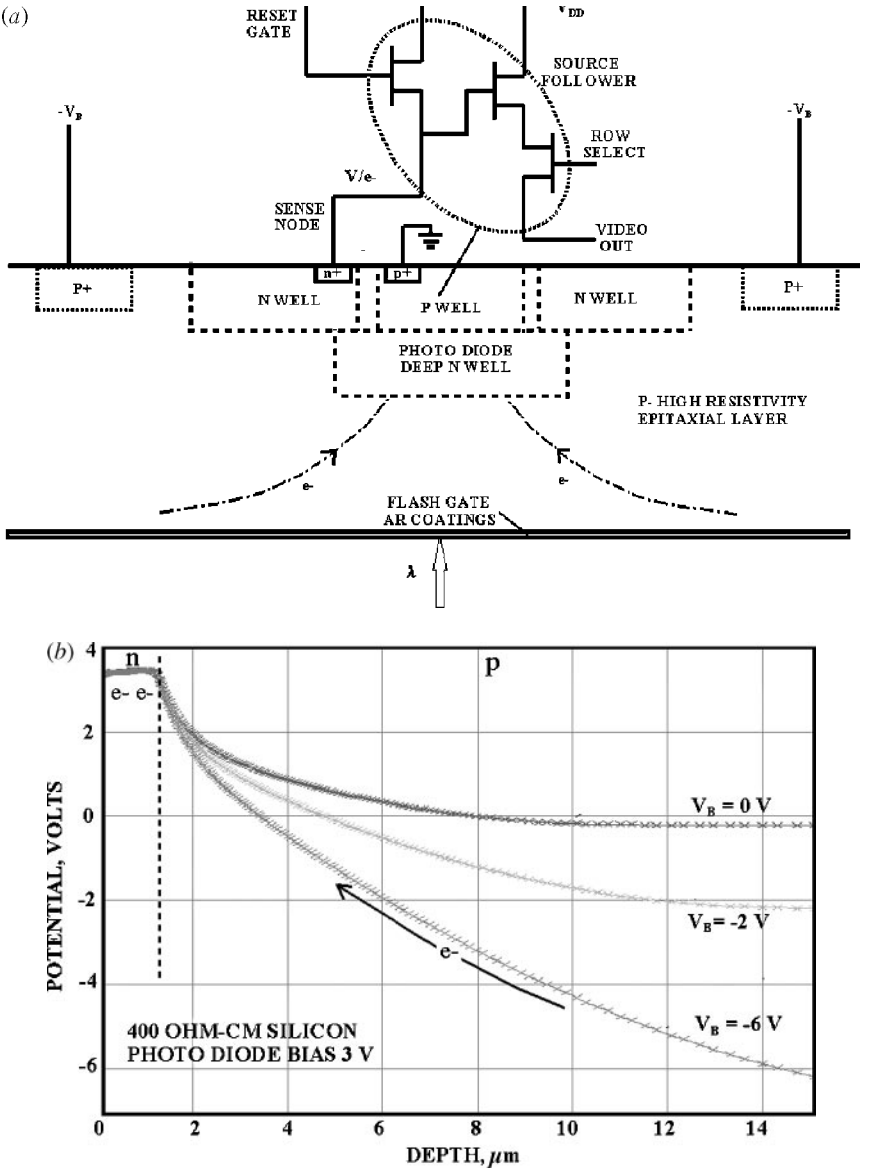

Figure 15 (*a*) Advanced CMOS pixel that is backside-illuminated and substrate-biased for high QE and deep depletion performance. (*b*) Potential plot for the advanced CMOS pixel showing different depletion depths with substrate bias.

limited in size by the dimensions of the wafer (20 cm is common). A seamlessly stitched CMOS sensor with a 3840 × 2160 resolution—four times the resolution of HDTV—and 7.5-μm pixels has been fabricated. A stitched 2 K × 2 K-pixel IR array with a size of 54 × 55 mm has also been made—one of the largest submicrometer imaging arrays produced to date.

6.3. Hybrid Arrays

CCD technology has shown textbook performance with near-perfect photon interaction and collection of signal carriers. CMOS technology has not demonstrated such high performance. On the other hand, CMOS arrays allow very fast signal acquisition through parallel processing with low noise, low power, and excellent electronic-circuit compactness. Imaging groups are taking these fine qualities from each technology and integrating them into a hybrid array. For example, a CCD-to-CMOS hybrid array can be constructed by bump-bonding a high-performance CCD to a CMOS signal-processing array. For example, Figure 16 shows the output region of a CCD array mated to a CMOS readout array. A source-follower amplifier detects signal charge at the bottom of each vertical register. Bump bonds are applied to interface to the CMOS processor. For low power dissipation, bumps can also be applied directly to sense nodes without source-follower amplifiers. The figure also illustrates how the devices can be physically mated. In the figure, a backside-illuminated CCD is bump-bonded to the CMOS array. The ultra-compact hybrid architecture results in high speed, low noise, high QE, and low power dissipation. For example, a 1 K × 1 K-pixel hybrid device can be read at 80 frames/s, assuming four ADCs are employed and read at 20 Mpixels/s (4). The hybrid can potentially achieve a read noise of 10 e^- at this rate.

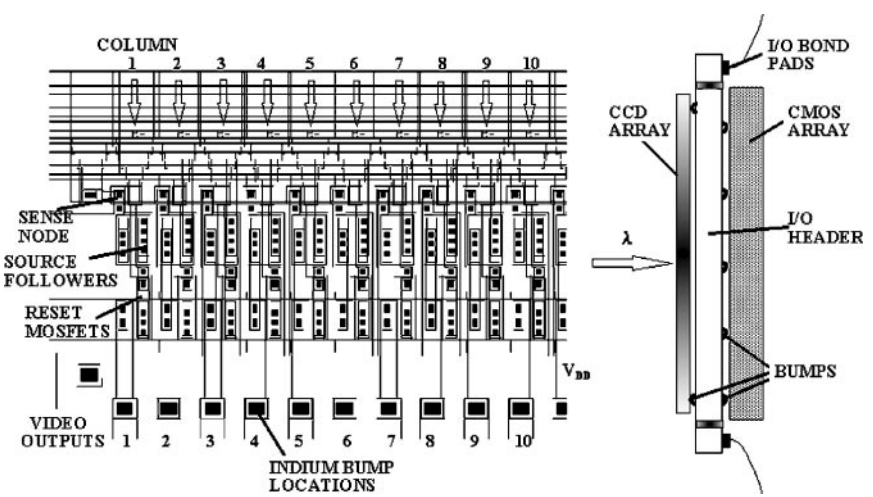

Figure 16 CCD to CMOS hybrid sensor.

It is also possible to mate a backside-illuminated CMOS pixel array to a CMOS readout array (4). The arrangement allows the CMOS pixel array to be optimized and fabricated independently from the CMOS processing array. Deep-depletion pixels (Figure 15*a*) are fabricated on a high-resistivity epitaxial silicon that is biased negatively. Also, the pixels and signal-processing circuits are electrically isolated for low noise performance. As with the CCD-to-CMOS hybrid, bumping can be done at the top and bottom of each array column. Bumping can also be done within the imager because individual pixels are buffered with source-follower amplifiers. This signal-processing arrangement offers even higher frame rates than the CCD-to-CMOS hybrid. For example, a 10 K $\times$ 10 K pixel CMOS-to-CMOS hybrid that is tapped every 512 pixels vertically can produce a frame rate of 100 frames/s (4). A net pixel rate of 10 Gpixels/s can theoretically be generated. With this hybrid design, there is no trade-off between noise performance and array size, and the hybrid can be read directly without transferring charge over long distances (i.e., pixels can be randomly accessed). This feature is advantageous for applications that must work in high-energy radiation environments (e.g., space) where CCDs are very vulnerable (1).

Development of CMOS imagers fabricated on SOI (silicon-on-insulator) wafers is in its infancy but is already achieving high performance compared with conventional bulk technology (10). SOI wafers include two silicon layers that are separated by an oxide insulator. Considered hybrid, the top and bottom silicon layers are referred to as the device and handle layers, respectively (Figure 17). The insulating layer is called the "box" layer. Each silicon layer can perform different functions with separate ground returns. The "device" layer incorporates the three-read pixel MOSFETs, whereas the "handle" layer is used for the photo region. Fabricating CMOS imagers on SOI greatly improves performance. For example, CMOS circuitry can be isolated from photo regions, allowing high-resistivity silicon and substrate bias to be employed. In addition, CMOS circuit ground returns

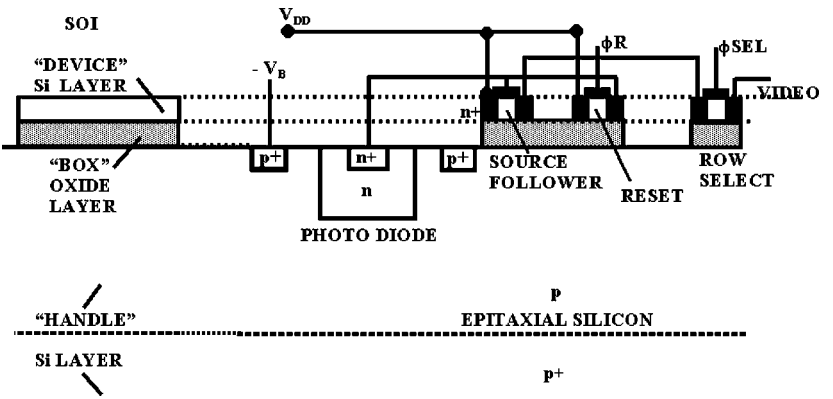

Figure 17 CMOS silicon-on-insulator hybrid photodiode pixel.

for SOI can be isolated, eliminating substrate bounce and transient coupling problems, to allow higher operating speeds and lower noise. SOI also makes it easier to passivate surfaces for low dark-current generation because of its planar structure. This advantage, in turn, makes the device much more resistant to high-energy radiation environments. Lower FPN and lower power consumption have also been reported for SOI.

The *Annual Review of Nuclear and Particle Science* is online at http://nucl.annualreviews.org

LITERATURE CITED

1. Janesick J. *Scientific Charge-Coupled Devices.* Bellingham, WA: SPIE (2001)
2. Janesick J. *Proc. SPIE* 4669A:232. Bellingham, WA: SPIE (2002)
3. Holland SE, et al. *IEEE Trans. Electron Devices* 50(1):225 (2003)
4. Janesick J. *Charge coupled CMOS and hybrid detector arrays.* Presented at SPIE Astron. Optics Instrum. Conf./Focal Plane Arrays for Space Telescopes (AM119), San Diego, CA, Aug. 3–8, 2003
5. Des Jardin W, Kosman S. *SPIE Proc.* 3649:74 (1999)
6. Losee PD. *SPIE Proc.* 5017:167 (2003)
7. Pain B, et al. *Analysis and enhancement of low-light-level performance of photodiode-type CMOS active pixel imagers operated with sub-threshold reset.* Presented at 1999 IEEE Workshop CCDs and AIS, Nagano, Jpn., June 1999
8. Pain B, et al. *Proc. 2002 Int. Electron Devices Meet., Dec. 8–11.* New York: IEEE (2002)
9. Mackay C, et al. *Proc. SPIE* 4306:289 (2001)
10. Zhang W, Chan M, Wang H, Ko PK. *Proc. 1999 IEEE Int. SOI Conf., Oct. 1999.* New York: IEEE (1999)
11. Cheng H-Y, King Y-C. *IEEE Electron Device Lett.* 23(9):538 (2002)

Annual Review of Nuclear and Particle Science
Volume 53, 2003

CONTENTS