J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*
Copyright © 1999 John Wiley & Sons, Inc.

IMAGE SENSORS

An image sensor is a device capable of converting the photon flux from an incident optical image into an electrical representation. Like photographic film, image sensors typically reside in the focal plane of an imaging system and record all or a portion of an object, or scene, focused onto its surface. Image sensors can, however, detect a limitless number of images, with each new image overwriting the previous one, and so they remain a permanent part of the imaging system. The majority of modern day image sensors are manufactured using semiconductor technology, rather than the vacuum tube technology of their predecessors (1), and are called solid-state image sensors. They are constructed so that selected regions of the integrated circuit are responsive to incident light. These optically sensitive regions convert photons into charge carriers and are divided into evenly spaced square or rectangular sections, called photosites or pixels. The photosites, in turn, are typically connected to a readout means consisting of *x*–*y* address lines or vertical and horizontal shift registers, depending upon the architecture, followed by an output conversion section that transforms the image charge into a voltage or current to be read by the imaging system.

Image sensors are available in many configurations, as shown in Fig. 1, for use in a wide variety of applications. One-dimensional image sensors, referred to as linear imagers, are commonly used in applications where the object or scene to be recorded is in relative motion with respect to the sensor. These devices capture one line of the object or scene image at a time and subsequently read out the image through a light-shielded shift register, while the next line is integrating in the photosites. Two-dimensional image sensors, or area imagers, are used in applications requiring an entire two-dimensional image of an object or scene to be recorded at one time. Block diagrams of both linear and area imagers are illustrated in Fig. 2.

Silicon is the most widely used semiconductor material for the manufacture of solid-state image sensors because of its ability to absorb photons efficiently in the visible spectrum (i.e., optical wavelengths between 400 nm and 700 nm). Infrared imagers for use in thermal imaging systems require the absorption of photons with wavelengths greater than 1.0 μm and use less prevalent materials formed from binary and ternary III–V semiconductors (2), such as indium antimonide (InSb) and platinum silicide (PtSi).

Image sensors for detecting color images are formed by depositing additional layers of either an absorption or an interference bandpass filter on the surface of selected photosites. These filters operate by absorbing or reflecting a high percentage of out-of-band photons, while transmitting in-band photons into the underlying photosite. The imager output is then reconstructed off chip to create a spatially accurate full-color representation of the original image, a process that requires an increasing number of photosites to ensure little or no perceptible loss of color (3) and spatial information (4).

The concept of a monolithic solid-state imager was suggested by Boyle and Smith (5) at AT&T during the late 1960s along with their invention of the charge-coupled device (*CCD*)—a semiconductor concept based upon the *MOS* (metal–oxide–semiconductor) capacitor wherein minority carrier charge packets, generated from an impinging photon flux, can be isolated and spatially transported by means of varying electrostatic potential wells. At that time, electron-beam scanning tubes (e.g., the orthicon and vidicon tubes) were the choice for image capture devices. Integrated detector arrays using an MOS switching concept for *x*–*y* addressing of photodetectors were being demonstrated (6) in the late 1960s, followed by the experimental verification of

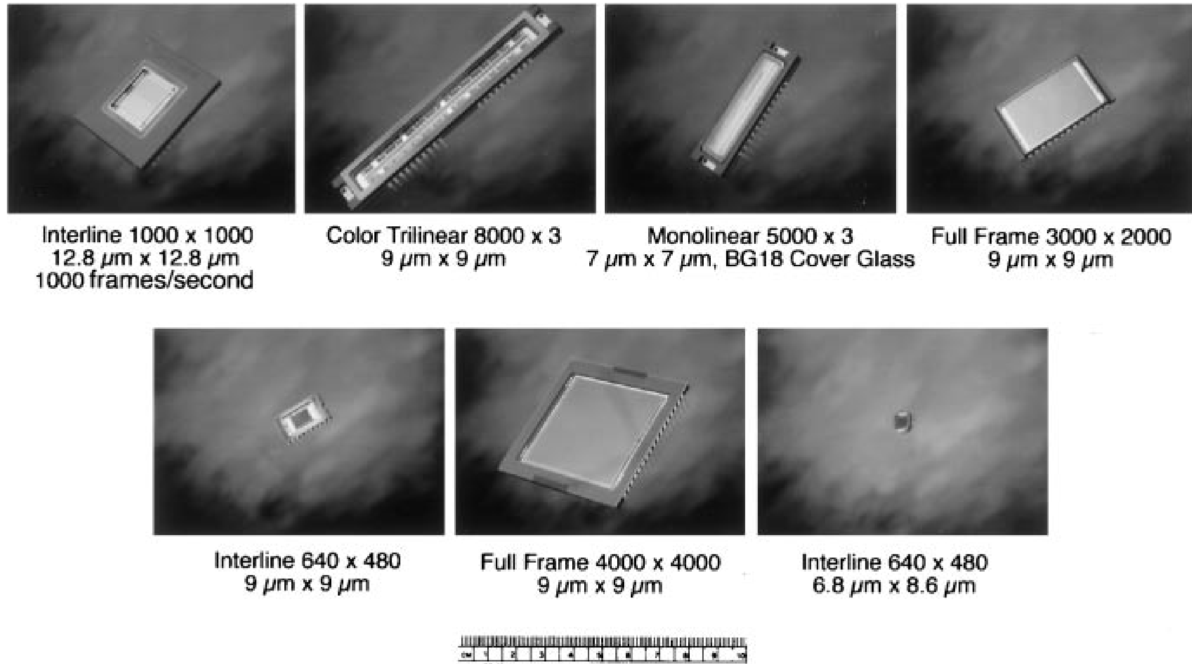

Fig. 1. Examples of solid-state image sensors.

the CCD concept in early 1970 (7). An adaptation of the MOS imager architecture called the charge injection device (*CID*) was also being developed in the early 1970s by General Electric (8). The rapid advances in silicon processing technology that occurred during the next decade yielded solid-state image sensors with performance capabilities rivaling those of the electron-beam tubes. The lower-cost, smaller, and more durable solid-state image sensors began appearing in professional equipment by the late 1970s, and by the mid 1980s, silicon-based image sensors were being produced in large quantities for both consumer and professional products. In the early 1990s, an image sensor architecture called the active-pixel CMOS sensor (*APS*) emerged. Using this technology, one can manufacture an imager with a process compatible with mainstream integrated circuit processing technology and equipment, resulting in a considerable cost savings over CCD processed imagers; hence this technology is attracting significant attention and development.

Solid-state image sensors are used in a wide variety of applications ranging from toy electronic cameras for young children to high-quality commercial broadcast cameras, and esoteric imaging systems for astronomy and space exploration. The cost can range from several dollars to several thousand dollars per unit, depending on the type and quality of the imager. Millions of linear and area imagers are used in consumer and professional products each year, and this quantity is expected to grow as the consumer digital still and video photography markets expand.

The most common application of area imagers is in still photography and video applications (e.g., digital cameras, telescopes, and camcorders). Video cameras typically use imagers compatible with NTSC, PAL, SECAM, or HDTV video standards, making it easy to interface to standard consumer video equipment. Cameras built around area infrared image sensors are used extensively in night vision and surveillance systems by both commercial and military markets. Area imagers also find use in medical applications such as surgery and dental x rays. Specialty area arrays capable of recording images at several thousand frames per second find

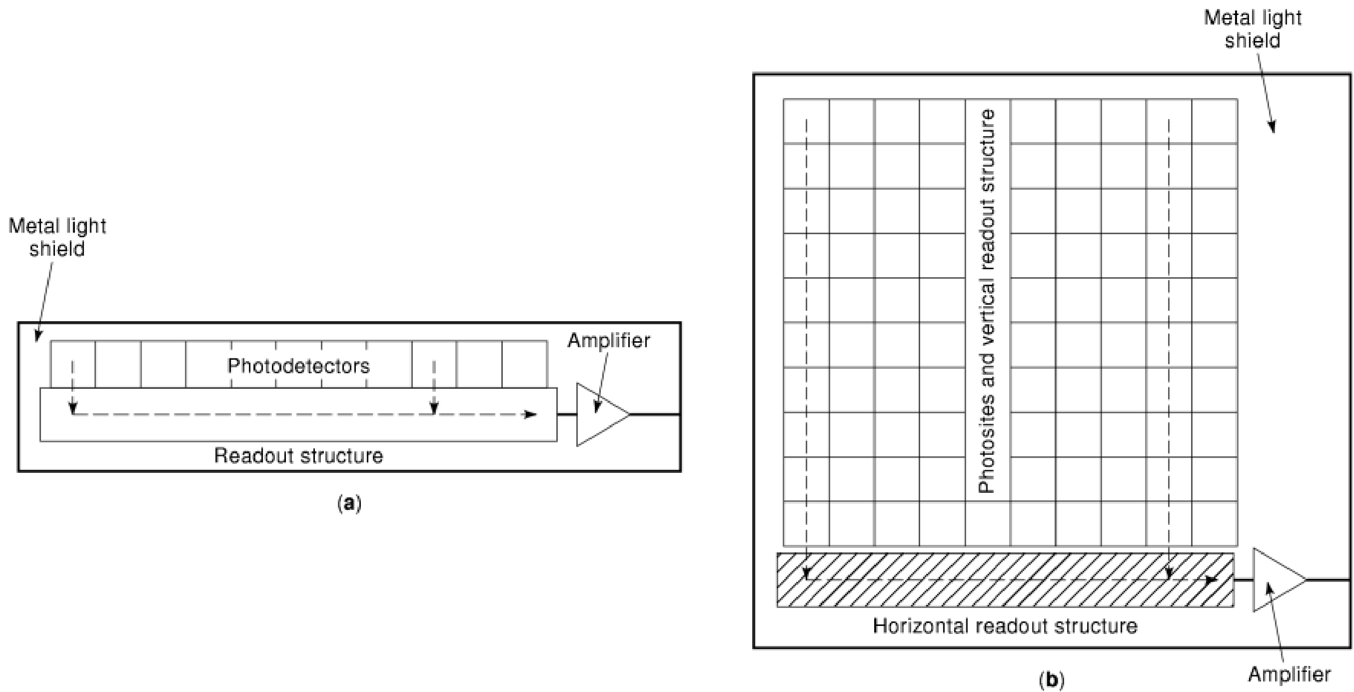

Fig. 2. Block diagrams of (a) linear and (b) area array image sensors (arrows indicate the flow of photogenerated charge). A metal light shield covers all but the photosensitive regions of the device.

application in motion study, where the recorded high-speed frames can be played back at much slower rates for detailed image analysis.

Linear imagers are used extensively in scanning systems such as fax machines, digital paper copiers, and drum and flatbed scanners. Although most airborne photographs still come from film-based cameras, solid-state image sensors, particularly linear imagers, play a critical role in the digitization of the processed photographs (9). Image sensors have also been used in earth orbiting satellites for many years (10). Operating in a space environment presents many challenges to solid-state devices, due to the elevated levels of high-energy radiation present (11). This fact, coupled with the drive to reduce overall spacecraft weight (and shielding) requirements, puts increasing radiation hardness demands on image sensors (12,13).

One can see from the discussion above that image sensors are found in many applications that previously used photographic film or perhaps were not even possible. As digital imaging continues to expand in commercial and consumer markets (fueled by the proliferation of the personal computer), image-sensor-based products are expected to become an integral part of both commercial and professional mainstream products.

Imager Architecture

When describing the operation of solid-state image sensors, the concept of electrostatic potential and potential energy diagrams proves indispensable. As depicted in Fig. 3, when isolated silicon atoms are brought together to form a crystalline lattice, the once discrete energy levels of the individual atoms expand to become bands of potential energy that electrons may occupy. The electrical properties of crystalline materials are determined primarily by the gap between the upper two energy bands, termed the valence and conduction bands (14). For a

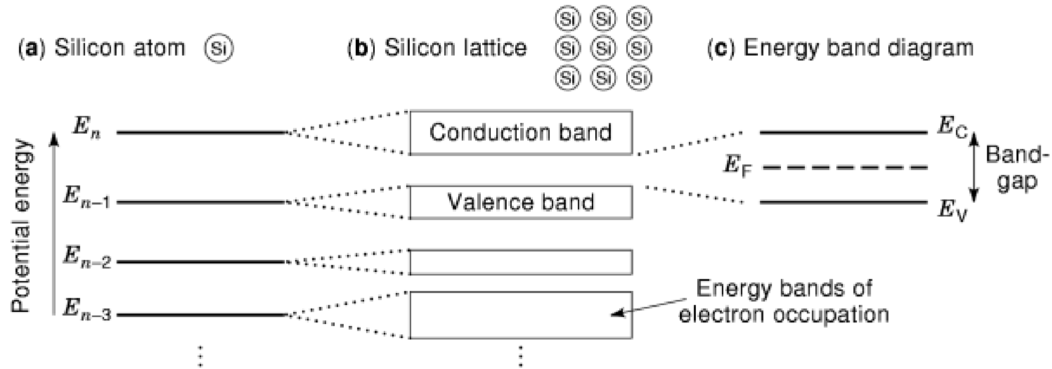

Fig. 3. Electron energy levels in (a) a single silicon atom and (b) a crystalline lattice of silicon atoms, and (c) the resulting energy band diagram.

semiconductor in thermal equilibrium, the gap between the valence and conduction bands is constant, and one needs only to specify the potential energy level of one band to describe the electrical state of the semiconductor in a particular region. By drawing the potential energy diagrams, or equivalent electrostatic potential diagrams (PE $= -q\ \Phi$), the areas where photogenerated charge may reside can be identified.

If an image sensor is to generate an accurate electrical representation of an incident image, it must be able to localize packets of photogenerated charge within each photosite, and transfer these charge packets to an output circuit. The underlying structure that most solid-state imagers rely upon for charge isolation and transfer is based upon the MOS capacitor (15), shown with its potential diagram in Fig. 4. By pulsing this structure with the appropriate bias, a potential *well* can be formed within the silicon substrate that attracts and stores minority carriers (electrons in this case). When the device is built upon a uniformly doped substrate, it is termed a surface channel device. This is due, as shown in Fig. 4, to the fact that the potential maximum of the biased capacitor is near the device surface, at the silicon–silicon dioxide interface. By forming such structures, charge carriers can be spatially isolated or transported between sections of an imaging device, depending upon the potential of the adjacent regions. There are several means by which a localized change in electrostatic or well potential can be created, as illustrated in Fig. 5. Selectively doping areas of the sensor substrate with p (+) and/or n (−) impurities, scenarios 1 and 2 in Fig. 5, causes static changes in electrostatic potential, giving localized potential steps. The application of a local external bias voltage to the MOS structure, as discussed above and as shown in scenario 3 in Fig. 5, gives a dynamic change in potential, with the changes being linearly related to the applied bias voltage. A combination of impurities and applied biases, as shown in scenarios 4 through 6, can also be utilized to change the potential under all or part of an electrode, and is extremely useful in achieving directionality of the charge transport; it is used to simplify the clocking scheme necessary to operate some types of imagers. The formation of the potential wells is utilized for the storage and transfer of signal carriers in virtually all types of solid-state imagers, as will be shown in the following sections.

CCD-Based Image Sensors. The CCD structure is the dominant architecture of image sensors used today, in all types of imaging systems. It provides a simple, efficient, yet precise means of transporting electronic charge—a characteristic essential to high-resolution imaging. A CCD shift register is created by forming a succession of closely spaced MOS capacitors and sequentially clocking them. To illustrate this concept, the three-phase CCD structure illustrated in Fig. 6 will be discussed.

This structure is perhaps the simplest architecture to implement, and achieves charge movement by sequentially forming a potential well in front of a charge packet, then collapsing a well behind the packet. Part (a) of Fig. 7 shows the channel potentials and the transfer process at various stages of the clock timing, which

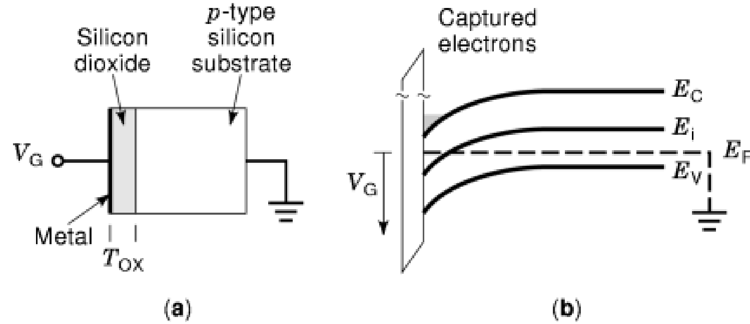

Fig. 4. (a) MOS structure and (b) energy band diagram for MOS structure under positive bias. Note the bending of the electrostatic potential lines near the silicon–silicon dioxide interface.

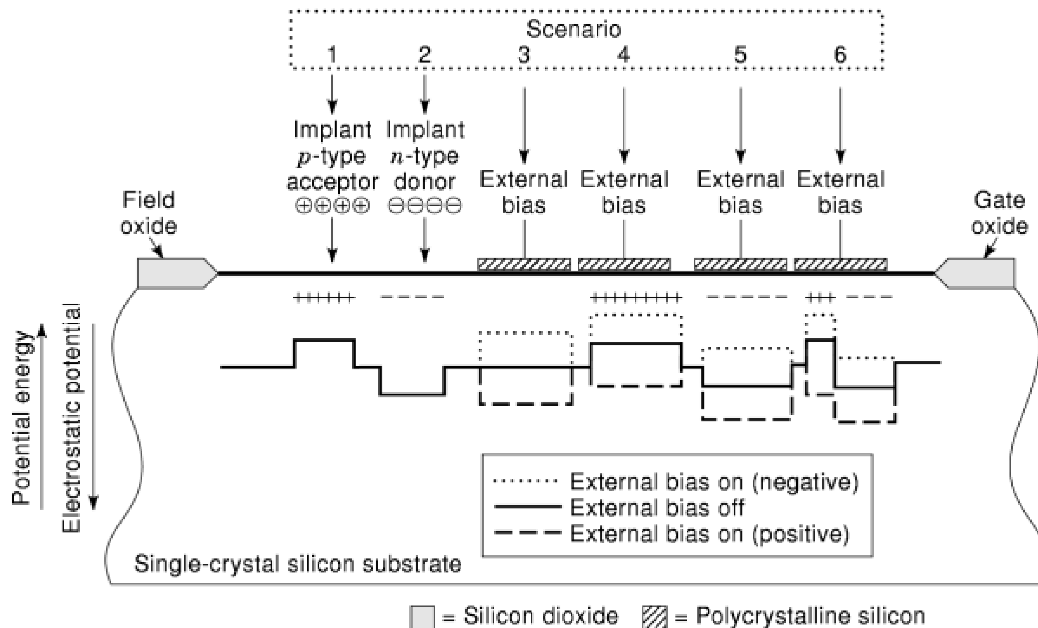

Fig. 5. Several scenarios for generating localized electrostatic potential differences within a semiconductor. The localized potential differences are used to confine packets of photogenerated charge within spatially isolated regions.

is shown in part (b) of the same figure. The degree to which charge can be completely transferred between potential wells is the fundamental measure of CCD performance, and is termed charge transfer efficiency (*CTE*). To maintain signal integrity, it is desirable for the CTE of a device to approach unity, meaning 100% of the charge is transferred between potential wells each time a transfer occurs. Realistically, this is not possible due to physical limitations of carrier transport and processing capabilities. In CCD-based image sensors, there is typically one conversion node for the entire array, and it is necessary to transport the packets of charge spatially from the site of origin to this charge-to-voltage conversion amplifier, which can be many clocking (or shift register) stages away. For a CCD having several thousand stages, transfer efficiencies of 0.99999 per stage must be achieved to maintain signal accuracy to within several percent.

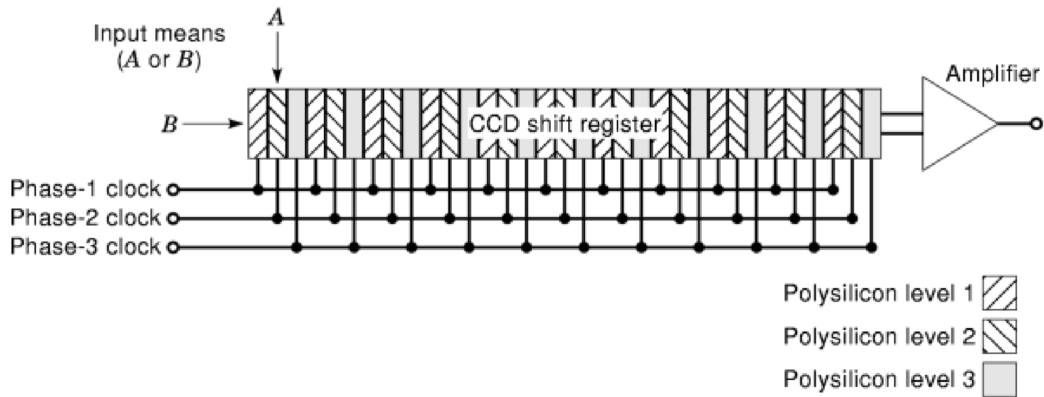

Fig. 6. Block diagram of a typical three-phase charge-coupled device. Each stage of the CCD is composed of three polysilicon gates. Charge can be injected into the CCD serially through B or in parallel (i.e., into all stages at once) through A.

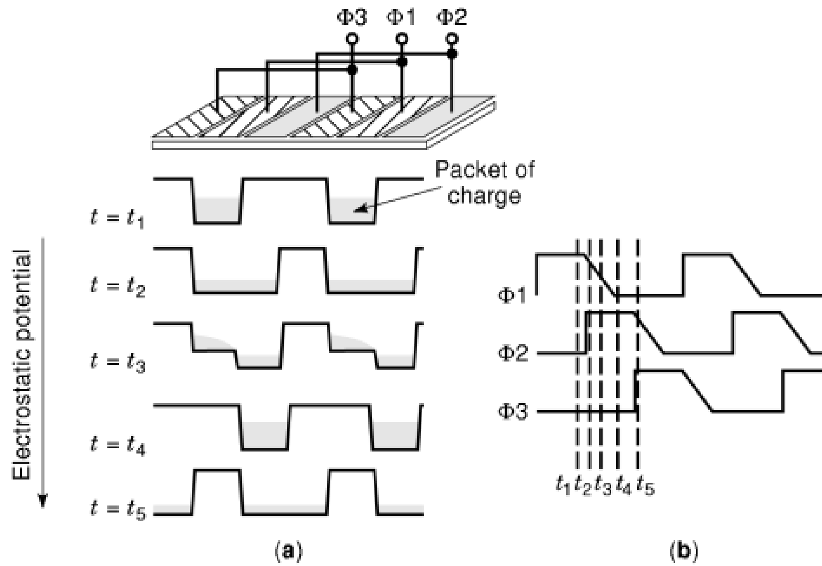

Fig. 7. (a) Charge coupling and (b) required timing for a three-phase CCD. Note the flow of charge from left to right within the CCD as the timing proceeds.

Charge transport in CCDs is governed by three phenomena: diffusion, self-induced drift, and field-aided drift. Diffusion transport arises from the gradient of the carrier distribution across the transferring and receiving electrodes and is proportional to the thermal diffusion coefficient of the signal carrier in the material. The charged nature of the carriers, again coupled with the gradient in the carrier distribution, can develop a charge gradient giving rise to an electric field. As an empty potential well is formed adjacent to a full well, the carrier gradient can become very steep and acts to help move carriers in the direction of transfer. This mechanism, referred to as self-induced drift, is a function of time and carrier density (16). As the transfer proceeds and the charge equilibrates between the potential wells, the self-induced fields vanish.

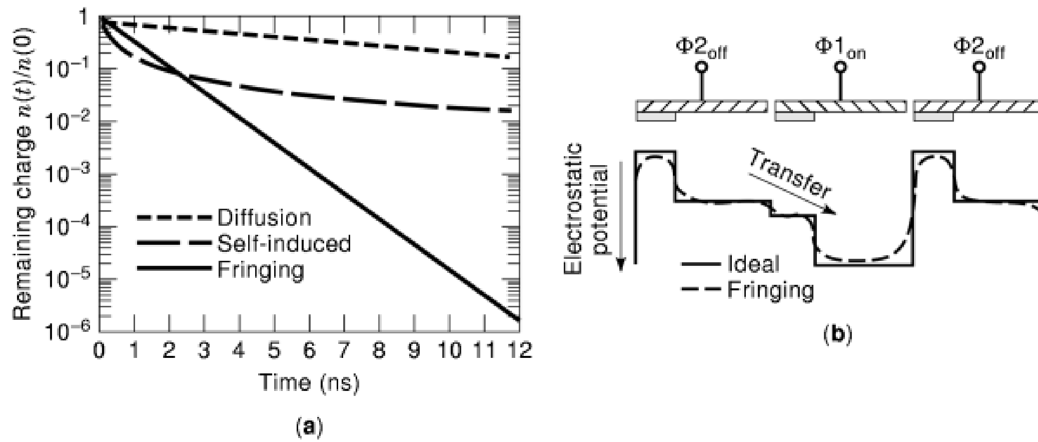

Fig. 8. (a) Two-phase CCD channel potentials and (b) relative charge transfer components (4.5 μm phase length). Note the significant reduction in time needed to transfer a charge packet due to fringing electric fields between and within individual CCD gates.

Barbe (17) presents useful approximations for the calculation of the self-induced fields and the resulting charge transport. Field-aided transport results from the fringing electric fields generated by a difference in the electrostatic potentials beneath adjacent electrodes or portions of an electrode, as in the case of a two-phase device.

The electrode arrangement and channel potentials for a two-phase CCD structure are shown in Fig. 8(a). The idealized channel potentials are shown as the solid lines, while the dashed lines indicate the actual potentials including fringing effects. The channel potential difference between the forwarding and receiving electrode regions is controlled by channel doping levels and the applied gate voltages and is typically several volts; hence the field and corresponding transport at the gate edges are high, but taper off toward the electrode center. The overall transfer efficiency, and, hence, the maximum clock rate, is ultimately limited by these center regions, and is strongly dependent upon the phase length. In the case of short-gate-length devices, two-dimensonal effects can change the potential profile across a significant portion of a gate electrode, giving rise to an increase in electric fields and marked improvements in transfer. The relative contribution of each transport mechanism is dependent upon the cell geometry and processing conditions, and can best be evaluated by calculating the components separately and comparing their values as functions of the transfer period.

The total transfer equation can then be approximated by superposition of the describing equations over separate intervals. Figure 8(b) shows the remaining charge ratio versus transfer time of each component for a 4.5-μm-length CCD cell. The fringing field component dominates in the latter stages of transfer when the self-induced fields have collapsed. This can be seen in Fig. 8(b), where the crossover to fringing-field-dominated transport occurs after the first 2 ns of transfer. Moreover, fringing field transport would predict up to a 50-MHz two-phase clocking rate while maintaining 10^{-5} efficiency per transfer, whereas diffusion limited transport would predict only 5-MHz operation. Precise solutions can be obtained through the use of two-dimensional simulation programs for the processing (e.g., SUPREMIV) and electrostatics (e.g., PICIES).

The surface channel CCD described above is the simplest to manufacture, but limitations on the transfer efficiency in these devices have been shown to occur due to trapping effects caused by interface states at the semiconductor–oxide boundary (18). The limitations surrounding the performance of surface channel devices led to the development of the buried-channel charge-coupled device (19), commonly referred to as a *BCCD*, which confines the charge to a channel beneath the semiconductor surface. Such a structure is shown in Fig. 9(a), where an additional doping layer of opposing conductivity type (n-type in this case) is placed at the surface

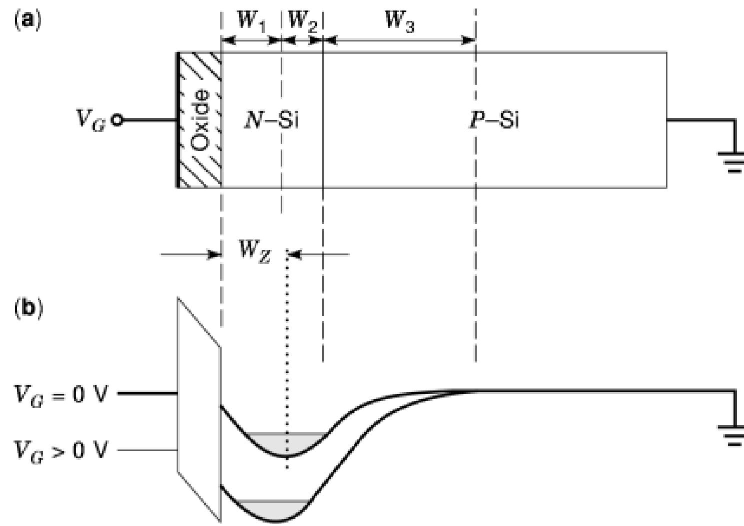

Fig. 9. (a) BCCD structure and (b) resulting channel potential profiles. Charge packets will now reside a distance W_Z below the silicon–silicon dioxide interface.

of the semiconductor. A depletion region, consisting of $W_2 + W_3$, is formed between the n-type region and the p-type substrate by application of a reverse bias across a contact to the n-type region and the substrate. With an additional bias placed on the gate of the structure, one can effectively create a second depletion region, W_1, similar to the surface channel device case, extending downward from the oxide–semiconductor interface, and merging with W_2, producing the potential profile shown in Fig. 9(b). The potential maximum for this structure is now located at a distance W_z below the semiconductor surface within the n-type region, and, provided the charge packet is not too large, the minority carriers in such a device can be transferred between potential wells without contacting the semiconductor–oxide interface, thereby averting surface trapping losses. Another benefit of this structure is the increase in the extension of fringing electric fields between adjacent phases. This allows a substantial improvement (nearly 10 times) in transfer efficiency for the BCCD, even at high operating frequencies.

Common implementations of CCD-type imagers are shown in Fig. 10. Interline CCDs consist of an array of photodiodes, a series of charge-coupled shift registers that are used to transfer the charge packets serially across the array, a transfer gate that is used to isolate the signal charge from the shift registers during imaging, and an output circuit that is used to convert the charge to a voltage or current signal. In this arrangement, the shift registers are shielded from light and are clocked while a subsequent scene is being integrated in the photosites. This configuration allows for real-time imaging and is commonly used in motion applications such as the video camcorder. Also depicted in Fig. 10 is a full-frame imager, where the vertical shift registers are made of a transparent electrode material, and serve as both the photon detector and the means of vertical transfer. These devices are well suited for still applications, but require a shutter to capture a frame from a dynamic scene.

The ratio of the photoactive area to the total pixel area is called the fill factor and governs the sensitivity of the device. Fill factor varies depending on the architecture, with full-frame imagers having a 100% fill factor, while interline devices may have less than 50%. To compensate for the lower fill factor, microlenses (referred to as lenticular arrays or lenslets) can be fabricated over each photosite. These lenses refract into the photosensitive area photons that would otherwise reflect off the imager surface, yielding a net increase in sensitivity.

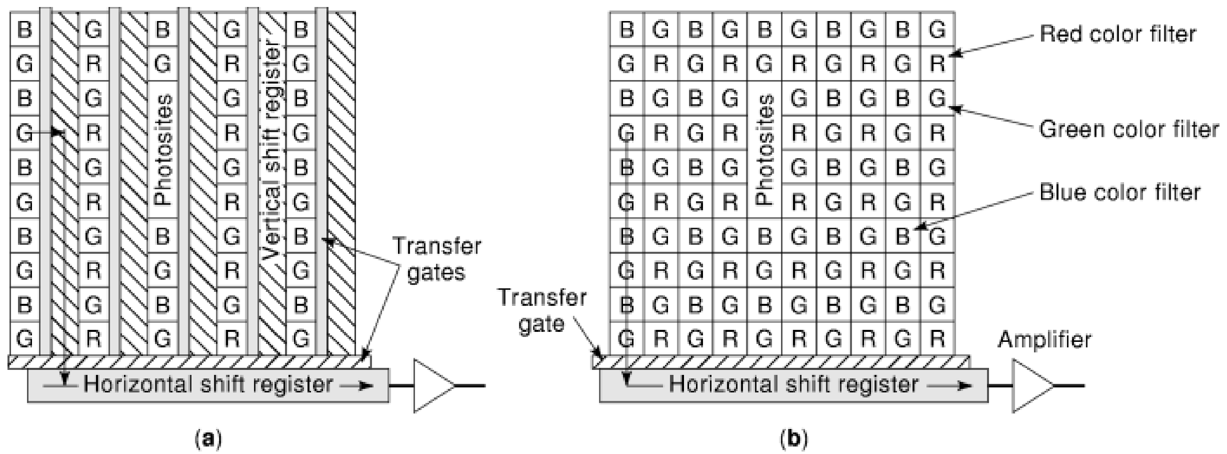

Fig. 10. Common types of CCD imager sensors: (a) interline area, 5 × 10; (b) full-frame area, 10 × 10. Full-frame image sensors are also referred to as progressive scan imagers. Note the repeating color filter patterns.

Other types of area arrays include frame-transfer (*FTCCD*) and frame-interline transfer (*FIT-CCD*) devices, which are variants of the above structures, and use additional light-shielded storage regions to enable quick scene capture with a slow readout rate.

Linear imagers require charge transport in one dimension only, and are typically smaller and simpler to manufacture than area arrays; hence their cost can be considerably less. They do, however, require movement of the image across the device to enable two-dimensional imaging, as is done with digital copiers or film scanners. Various types of linear imagers exist, as shown in Fig. 11, but most resemble a single column of an interline area array, having a row of photodiodes coupled to a single shift register through a transfer gate. Bilinear readout architectures are used in some designs to ease manufacturing tolerances, by increasing the shift register pitch and reducing the number of transfers a given charge packet must undergo. A popular implementation for color scanning is the trilinear imager, with three independent arrays integrated onto a single chip, each array having its own color filter stripe. A special class of linear image sensor, termed time-delay and integrate (*TDI*), actually consist of a two-dimensional array of photosites. The device is operated in a scanning mode (20), wherein the image motion is synchronized with the clocking of the rows of detectors, successively adding the charge from each line and thereby increasing the effective exposure by a factor equal to the number of TDI stages (rows). There are also various clocking configurations for CCD devices, including two, three, and four phases. The two-phase CCD structure, shown previously in Fig. 9, is a popular implementation, as it allows for the simplest timing and lowest noise, by requiring only two clocks (21). It requires the addition of impurities beneath a section of each phase, to modify the channel potential as shown previously in Fig. 5. Single-phase devices also exist, relying on implanted regions adjacent to the clocking electrode to provide an electric field in the appropriate direction, as were shown in scenarios 1 and 2 of Fig. 5 .

MOS and CID Image Sensors. Other image sensor structures, including MOS-selected photodiode arrays and CID imagers, have a similar physical layout to CCD-based imagers. In contrast to CCD imagers, these devices do not require a charge shift register; rather, the integrated signal in each photosite is coupled to or decoupled from a sensing line using a MOS transistor switch, controlled by address lines as shown in Fig. 12(a). For photodiode arrays, the signal is integrated on a photodiode structure, consisting of a *pn* junction that has been set to a reference bias level and is then left floating; the voltage across the diode decays due to the local photocurrent during the integration period. The signal level is detected by addressing the photodiode

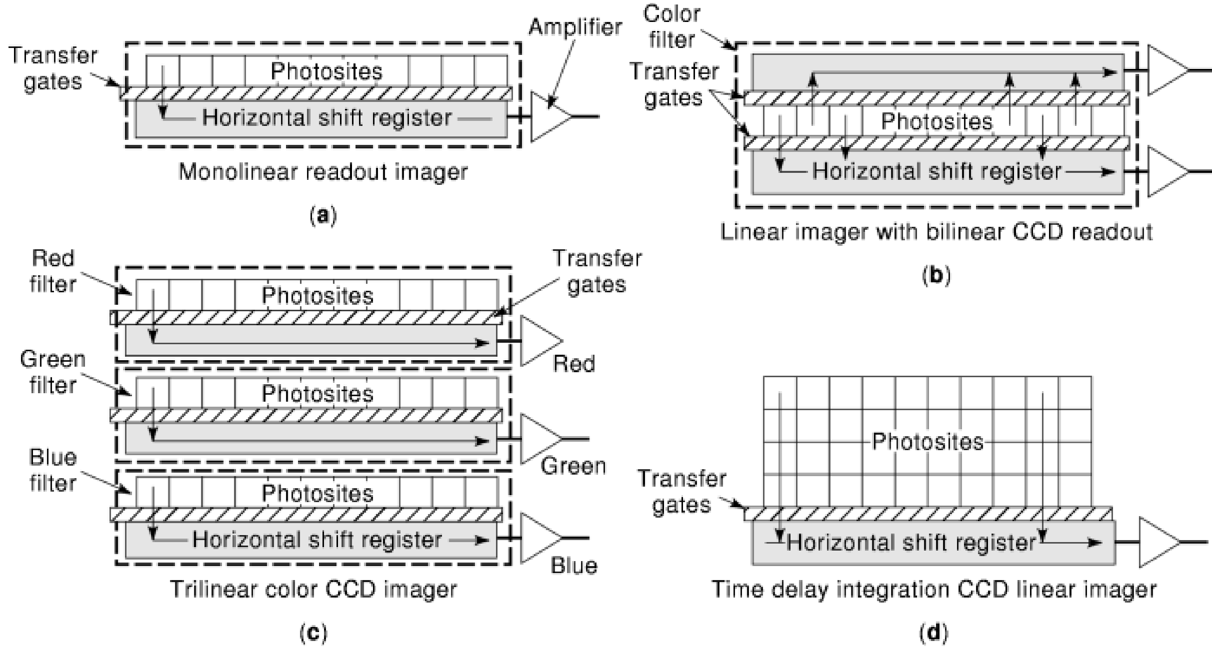

Fig. 11. Linear CCD image sensor configurations (clock lines and light shield removed for clarity). Increases in readout speed can be easily achieved by adding more outputs per CCD.

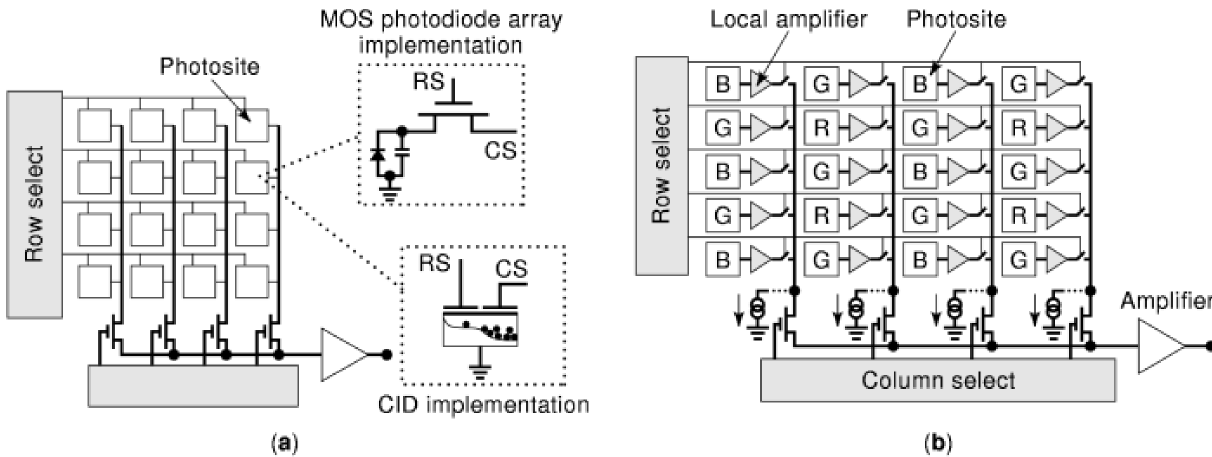

Fig. 12. Block diagrams of (a) 4 × 4 MOS photodiode and CID area two-dimensional image sensors and (b) 4 × 5 active pixel area image sensor (ACI).

through a MOS transistor switch and observing the displacement current required to reset the diode back to its reference level.

CID arrays utilize MOS capacitors as the photodetectors. The capacitor is biased to form a potential well much like the CCD, and photocurrent is collected and stored in this well during the integration period. The signal is then detected by selectively injecting the integrated charge into the underlying substrate or

localized well structure using MOS switches and measuring the displacement current, as is done with the MOS photodiode arrays (22). Some CID imager architectures allow photosites to be addressed in a random fashion, permitting subsections of the recorded image to be readout at higher frame rates. More recent implementations enable the readout of individual photosites in a nondestructive manner (23), where instead of injecting the charge into the substrate, the charge is injected into a column capacitance for sensing, and is then switched back to the storage capacitance to retain the signal. This feature enables adaptive exposure control and provides the option to reduce temporal noise by averaging multiple readings from each photosite. Both MOS and CID devices require the pixel to have high reset capacitance; therefore they suffer from high reset (kTC) noise, which restricts the achievable dynamic range of these devices and has limited their use in imaging applications.

Technology advances leading into the 1990s in the area of device scaling brought about the realization of a new imager architecture, the active-pixel sensor (*APS*). Like MOS and CID imagers, APS imagers do not require the use of a CCD shift register (24). Instead, a charge-to-voltage conversion amplifier is fabricated adjacent to each photosite using advanced CMOS processing. Individual photosite voltages are read from each photosite using a multiplexed x–y addressing scheme as shown in Fig. 12(b). As with CID imagers, APS devices can be addressed in a random manner. The presence of a local conversion/buffer amplifier at each pixel yields a much lower reset capacitance, reducing the detector reset noise, at the expense of increasing the pixel-to-pixel temporal noise. Each column typically has a bias network and an additional buffer; hence column-to-column fixed pattern noise can also exist. Nevertheless, the compatibility with standard CMOS processing and the availability of extensive standard cell libraries in this technology allow the implementation of noise cancellation or correction schemes not available on CID or MOS photodiode arrays. Since much of the world's integrated circuit processing is based upon CMOS implementations, there is significant worldwide manufacturing capability for these devices in terms of capacity, technology development, and low-cost manufacturing.

Charge Detection and Readout. The charge packets constituting the recorded image are typically converted into a voltage before being output from the image sensor to be compatible with conventional electronics. For CCD arrays, the readout structure consists of a shift register terminated into a charge sense node, typically a resettable floating diffusion formed by a reverse-biased *pn* junction, followed by a buffer amplifier, as shown in Fig. 13. The charge-to-voltage conversion factor η can be expressed as $\eta = q/C_\text{f}$, where C_f is the floating diffusion capacitance, and q is the electronic charge. Common values of η range between 1 μV and 20 μV per electron, corresponding to floating diffusion capacitance values in the range of 8.0 fF to 160 fF. For the two-phase CCD structure depicted in Fig. 13, the conversion is performed on the falling edge of the $\Phi 2$ phase when charge is transferred over the barrier set by the output gate (OG) bias potential.

The conversion node is typically followed by a one- or two-stage amplifier to provide power gain. Source follower amplifiers, such as that composed of the Q1–Q2 and Q3–Q4 transistor pairs of Fig. 13, are typically used, as they provide a high input impedance, low output impedance, and good linearity, and are compatible with the CCD device processing. The conversion node is cleared of charge, or *reset,* while $\Phi 1$ is in the not asserted (off) clocking state, through a transistor that is pulsed on by a reset clock signal. The clocking signals and the resulting output waveforms are shown in Fig. 14. The valid portion of the signal occurs during the latter part of the $\Phi 2$ off state. The reset signal is strongly coupled to the output waveform, due to the low capacitance of the floating diffusion node, and adds an artifact to the waveform in the form of a feedthrough voltage. The output signal settles shortly after the reset pulse is turned off. The true magnitude of the image signal is given by the change from the settled, or *clamp,* level to the maximum negative voltage, as indicated by ΔV in Fig. 14.

APS devices do not require shift registers for charge readout, but rather incorporate a charge-to-voltage conversion structure, similar to Fig. 13, at each photosite. For these devices, the shift register phases $\Phi 1$ and $\Phi 2$ and the output gate OG of Fig. 13 are replaced by a single transfer gate coupling the photodetector to the conversion node as shown in Fig. 15. The reset and amplifier sections are similar to those of the CCD imager; however, the drive transistor Q1 of the first stage of the amplifier is contained within each photosite region, giving rise to the name active pixel. This transistor may have a second gate, which allows for selecting

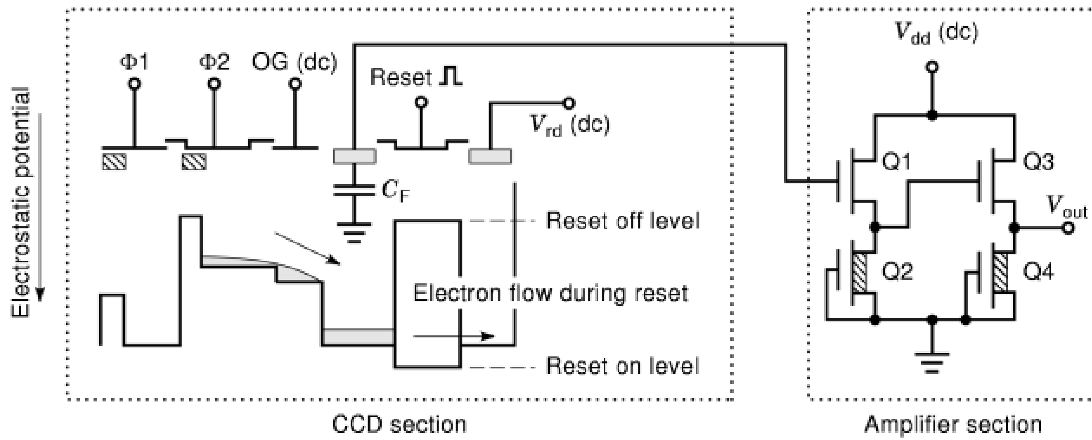

Fig. 13. Electrical schematic of a typical image sensor output structure and amplifier. As charge flows onto the floating diffusion capacitance, a change in gate voltage at Q1 is induced.

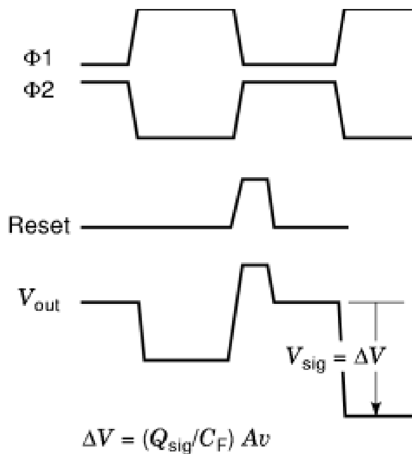

Fig. 14. CCD output structure timing and resultant output waveform for a two-phase CCD-based imager.

or deselecting the photosite. APS imagers are *x–y*-addressed during readout, and this second connection can be common to all of the drive transistors for a row of photodetectors, thereby becoming the row select (*RS*) connection. Each column may have its own amplifier section providing the output signal voltage to a second column selecting the multiplexing circuit.

Manufacturing imagers with a consistent value of η can be challenging. This problem is of particular concern in APS design and manufacture because each photosite has a dedicated single-stage charge-to-voltage amplifier. Failure to adequately match the thousands of amplifiers on an APS yields pixels having different gain, resulting in fixed pattern noise. Similarly, each column has its own amplifier, which can lead to a column-to-column pattern noise. This latter noise can be exceptionally objectionable, as the eye is very sensitive to spatially fixed patterns. For this reason, dedicated circuits are incorporated into each column readout to correct for gain and/or offset differences. These circuits can be quite complex; however, as mentioned earlier, APS processing is compatible with standard CMOS fabrication, and therefore these circuits are readily integrated

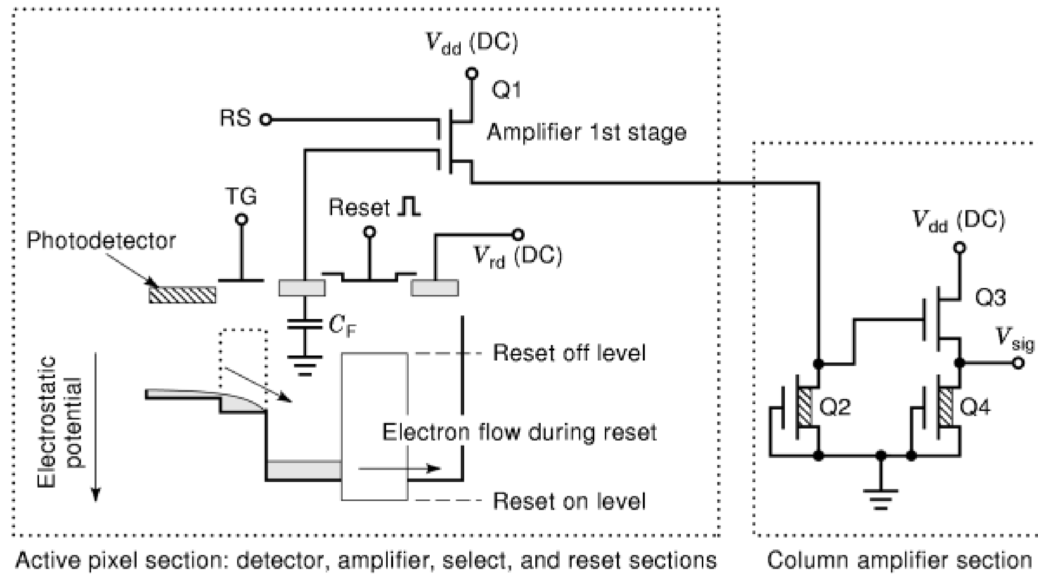

Fig. 15. Electrical schematic of a typical ACI detector and output structure. Note that each pixel contains part of the readout amplifier (Q1).

into the device. This capability enables APS imagers to rival the performance of CCD imagers for a number of applications, especially in low-cost, high-volume consumer applications.

Antiblooming and Exposure Control. When the maximum charge capacity of a photosite is surpassed, the excess charge, if not otherwise contained, can spill into adjacent structures such as neighboring photosites or shift registers. This condition is called blooming, and results in a corrupted image adjacent to the site of blooming. There are several means employed by image sensor designers to prevent or minimize blooming. One is to design exposure control structures adjacent to the photosites, which enable the user to control the period of time over which photogenerated charge will be collected. This period can typically be any fraction of the frame readout period. Another way of controlling blooming is to design specific antiblooming structures adjacent to the photosites. These structures direct the flow of excess photogenerated charge into the substrate in a controlled manner. It is also possible to combine the exposure control and antiblooming functions into one structure, as illustrated in Fig. 16(a) for a linear imager. If a sufficient positive bias is applied to the lateral overflow gate (*LOG*) in Fig. 16(a), photogenerated charge will be swept into the substrate through the adjacent drain. Hence, the effective exposure time can be controlled by varying the length of time the LOG is on during a readout frame. To prevent blooming, the off potential of the LOG is set higher than that of the adjacent transfer gate, thereby causing excess charge to spill over the LOG before spilling over the transfer gate. Antiblooming structures can handle overexposures in excess of 100 times the saturation level without blooming. Lateral exposure control and antiblooming structures are commonly used on linear imagers where the additional area required is more readily available without sacrificing fill factor. To minimize the additional pixel area, area imagers more commonly use vertical exposure and antiblooming structures, an example of which is shown in Fig. 16(b). Here, *vertical* means the overflow path is directly below the photosite. Exposure control can be accomplished by pulsing the substrate potential, effectively collapsing the barrier between the photosite and the substrate.

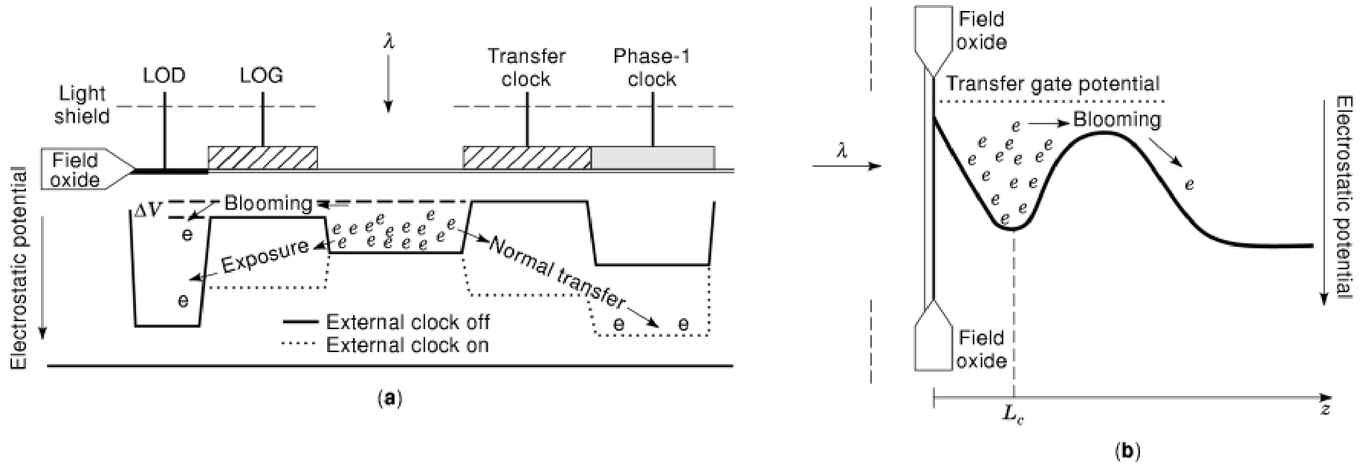

Fig. 16. Example (a) of lateral exposure control and antiblooming structure and (b) vertical overflow antiblooming structure. Protection against overexposures is critical in many applications, such as film scanners and astronomy.

Image Sensor Performance

Image sensors must capture impinging photons with a high degree of efficiency and spatial resolution and then convert the charge stored in each photosite into a signal suitable for readout off chip, all while maintaining a high signal-to-noise ratio. Some basic performance parameters used to quantify how well these tasks are performed are the quantum efficiency, modulation transfer function, and read noise.

Quantum Efficiency. Electrons are excited from the valence band into the conduction band by the absorption of incident photons with energy levels greater than the semiconductor bandgap, as described by the photoelectric effect (25). For silicon, the bandgap E_g is 1.12 eV; this implies only photons of wavelength λ less than 1100 nm can be absorbed, since $\lambda_{max} = hc/E_g$, where h is Planck's constant and c is the speed of light. If the incident optical radiation is a plane wave, then the minority carrier generation rate within the silicon substrate is given by

$$G_L(z, \lambda) = \sqrt{\frac{\epsilon_s}{\mu_s}} \frac{\lambda\alpha(\lambda)e^{-\alpha(\lambda)z}}{2hc} \left|\hat{E}_s^+\right|^2 \quad (1)$$

where z is the distance into the substrate, α is the absorption coefficient, ε is the permittivity, μ is the permeability, and E_s is the electric field intensity directed inward at the surface of the silicon substrate. The absorption coefficient of silicon decreases with increasing wavelength; therefore, the mean photon absorption depth ($1/\alpha$) is greater at longer wavelengths. The concept of photon absorption depth is of particular interest in regard to image sensors because it directly affects the quantum efficiency (*QE*)—that is, the ratio of captured photogenerated electrons to the number of photons incident per unit area and time—and the modulation transfer function (*MTF*), which is discussed later.

A typical photodiode or photocapacitor contains two distinct regions, a depletion region and a field-free region, as illustrated in Fig. 17. Electrons generated in the depletion region (region 1 in Fig. 17) will be confined within the photodiode by the built-in electric fields, and those generated in the field-free region (region 2 in Fig. 17) will randomly transverse the substrate until they either (1) recombine in the substrate, (2) diffuse toward the L_d boundary and are captured by the electric field, or (3) diffuse laterally and are captured in a neighboring photodiode or adjacent CCD structure. This last scenario is an important one, as it represents a

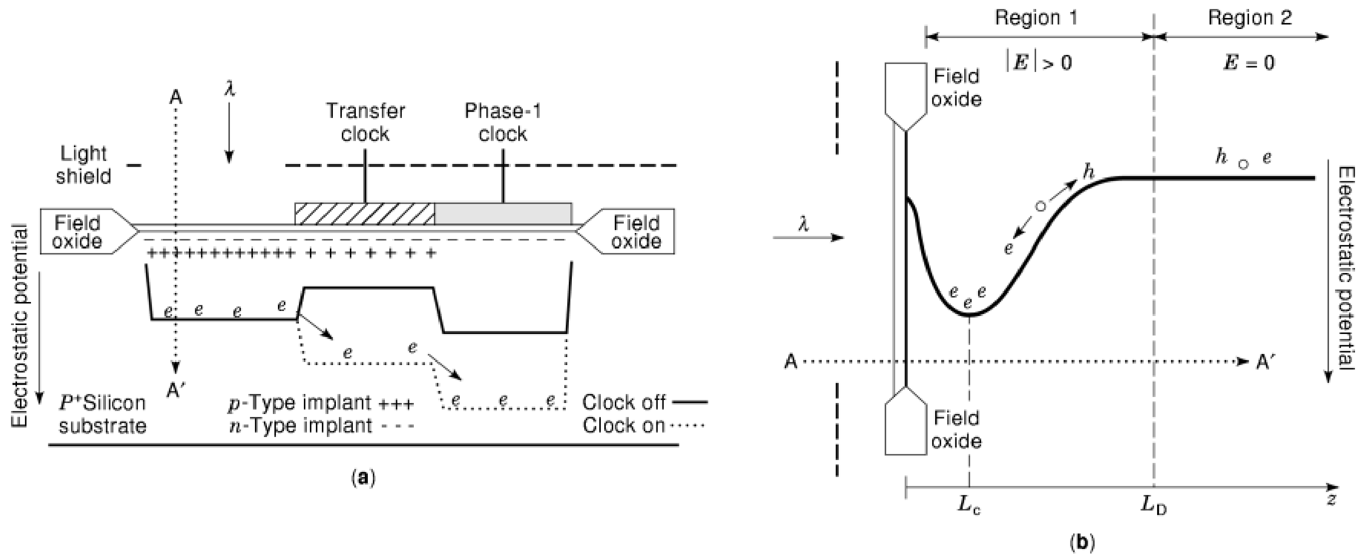

Fig. 17. Cross-sectional depiction of buried photodiode showing (a) electrostatic potential relative to adjacent transfer gate and CCD, and (b) electrostatic potential versus depth into silicon substrate.

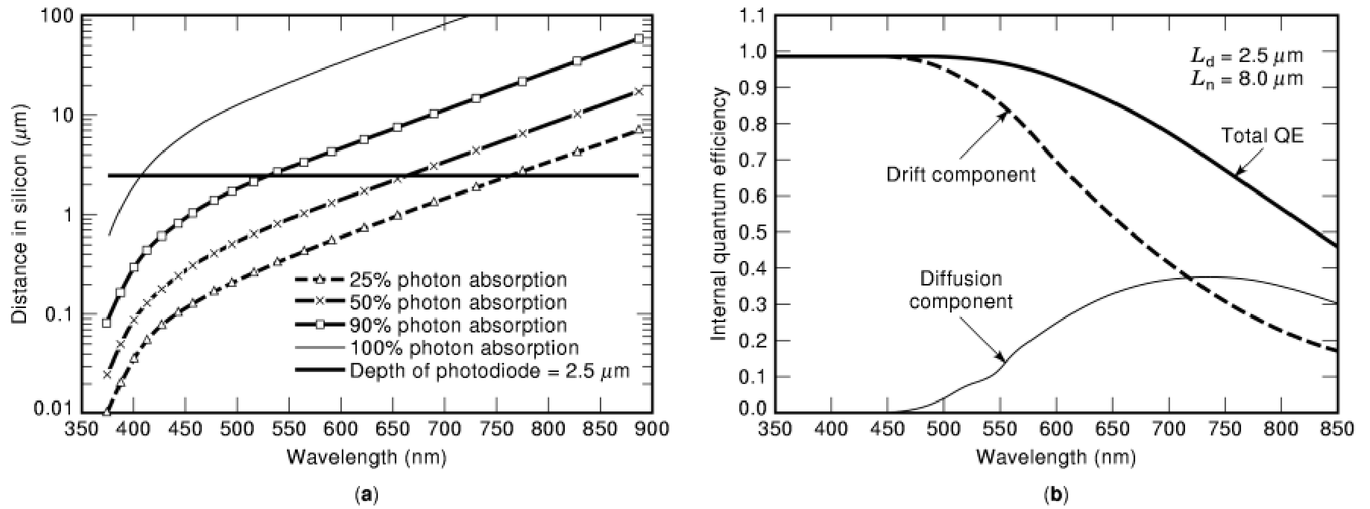

Fig. 18. (a) Percentage of photons absorbed versus depth into silicon and wavelength, and (b) components of internal quantum efficiency for a simple *pn* photodiode. By extending the depth of the photodiode deeper into the silicon, the diffusion component of QE can be reduced, thereby improving MTF.

mechanism that can degrade sensor performance, especially at longer wavelengths. Figure 18(a) shows several levels of photon absorption versus depth and wavelength, along with the depletion depth (2.5 μm) of a typical photodiode for reference.

Solving the electron continuity equation for the drift (QE_E) and diffusion (QE_D) components of the quantum efficiency for the case of a simple *pn* photodiode on a silicon substrate yields (26)

$$\begin{aligned} QE_E &= (1 - e^{-\alpha(\lambda)L_d})T(\lambda) \quad \text{and} \\ QE_D &= \frac{\alpha(\lambda)}{\alpha(\lambda) + 1/L_n} e^{-\alpha(\lambda)L_d} T(\lambda) \end{aligned} \tag{2}$$

where L_n is the minority carrier diffusion length, L_d is the depletion depth, and $T(\lambda)$ is the photon transmission from the outermost surface, typically SiO_2 and/or polysilicon, to just inside the silicon substrate. Equation (2) is graphed in Fig. 18(b) for a photodiode of depth 2.5 μm and $T(\lambda) = 1$ to show the relative magnitude of each QE component.

Many imagers are fabricated using an epitaxial layer on a silicon substrate. This arrangement dramatically lowers the diffusion length at a predetermined depth. This, in turn, reduces the diffusion component of the QE and improves the imager's spatial frequency response. Solving the electron continuity equation for such a case yields a model that accurately predicts the QE over the wavelengths in the visible spectrum (27):

$$\begin{aligned} QE_k(\lambda) = T(\lambda)\Bigg[1 + &\frac{\eta e^{-\alpha d}}{(\eta^2 - \alpha^2)(\sigma_+ - \sigma_-)}[(\eta + \alpha)\sigma_- - (\eta - \alpha)\sigma_+] \\ &+ \frac{2\alpha\eta e^{-\alpha t}}{\sigma_+ - \sigma_-}\left(\frac{(u - \alpha)yv}{u^2 - \alpha^2} - \frac{uv - y\alpha}{\eta^2 - \alpha^2}\right)\Bigg] \end{aligned} \tag{3}$$

where

$$\begin{aligned} &v = \frac{D_s}{D_e}, \qquad y = \frac{N_s}{N_e}, \qquad u^2 = \frac{1}{L_s^2} + k^2, \\ &\eta^2 = \frac{1}{L_e^2} + k^2, \qquad \alpha = \frac{4\pi N''}{\lambda} \qquad \sigma_+ = (uv + \eta y)e^{+\eta(t-d)}, \\ &\sigma_- = (uv - \eta y)e^{-\eta(t-d)}, \qquad k = 2\pi f_g \end{aligned}$$

and D is the diffusion constants for electrons, N is the doping concentration, L is the diffusion length, d is the depletion depth of the photosite into the epitaxial layer, t is the thickness of the epitaxial layer, N'' is the extinction coefficient of the semiconductor, and f_g is the spatial frequency. The subscript s refers to the semiconductor substrate, and e refers to the epitaxial layer. At longer wavelengths, this model somewhat under estimates the QE, whereas a model developed by Stevens and Lavine (28) has been shown to be an improvement. The use of a gradient-doped epitaxial layer builds in a gradual electric field, effectively extending the photodiode depletion depth and decreasing the diffusion MTF losses, while still maintaining high QE.

The goal in creating a color imager is for each photosite to pass photons over a narrow wavelength range (e.g., 430 to 490 nm for blue, 520 to 580 nm for green, or 600 to 670 nm for red) and reject all others. This color filtering process, which affects the QE through the transmission term $T(\lambda)$ in Eq. (2), can be accomplished indirectly using a monochrome image sensor and external color filters or directly by the application of color filters on each photosite of the imager itself. Depending on the type and thickness of the materials, light incident at the outermost surface of a photosite will be either absorbed within the materials lying above the semiconductor substrate, reflected away, or transmitted through the materials and into the underlying substrate. The color filter absorption process is accomplished by depositing a red-, green-, or blue-dyed film over each photosite. Each film will transmit a large portion of the in-band photons while absorbing the out-of-band photons.

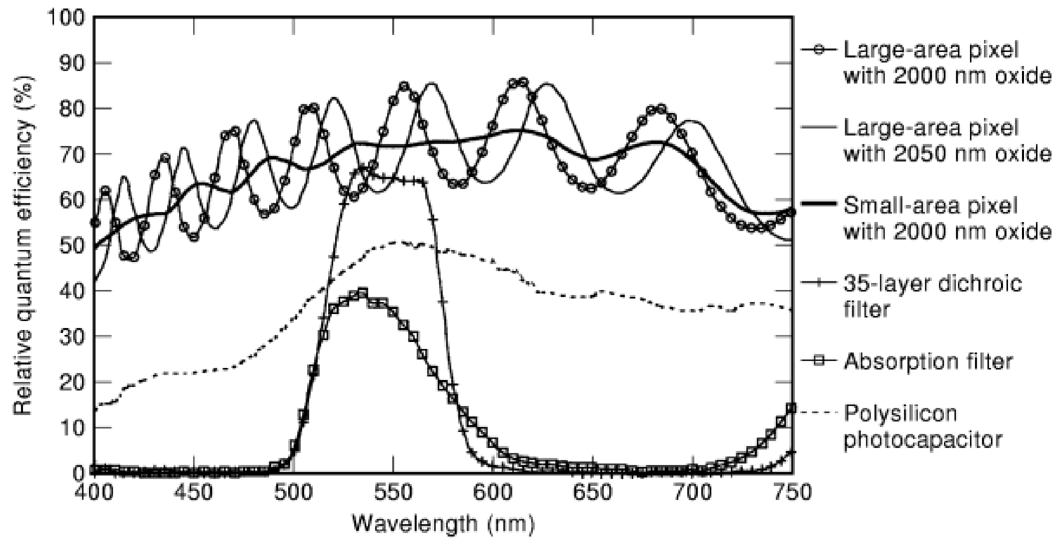

Fig. 19. A comparison of absorption and dichroic color filters, and the variation in quantum efficiency due to changes in oxide thickness. Note the phase shift with only 50 nm change in oxide thickness.

Dichroic filters are another technology used on some imagers, particularly linear imagers used in high-exposure applications. They are formed by depositing a series of thin films, with alternating high and low optical indices, over each photosite (29). The thickness of the films is made less than the temporal coherence length of naturally occurring light, so incident photons will undergo constructive and destructive interference within the dielectric film stack. A properly designed and manufactured dichroic filter will transmit photons over a narrow band with nearly ideal transmission and reflect all others. Figure 19 shows a comparison of absorption and dichroic color filters on a conventional imager.

Dichroic filters offer a number of advantages over standard absorption filters, in that they can be designed to give very precise control over the passband, they have better transmission, and they are resistant to change with time or exposure. Absorption color filters rely on special dyes, and therefore only a limited number of wavelength and bandwidth options exist. These dyes are also susceptible to fading. While having higher performance, dichroic filters are considerably more expensive due to the greater difficulty in manufacturing and limited number of manufacturers; therefore, their use is typically limited to high-end scanning systems and imaging satellites.

As an inherent part of silicon MOS processing, a thin silicon dioxide layer will exist at the photosite surface, and depending on the imager architecture, other layers such as polysilicon or tin oxide may also be present. Because the thickness of these layers is typically less than the temporal coherence length of the impinging light, optical interference can occur and modulate the spectral response in a similar manner to the dichroic filters described above. Large-area photosites (e.g., 15 μm $\times$ 15 μm and higher) appear more planar and are more susceptible to interference in the surface layers, as shown in Fig. 19, whereas the irregular surface topology inherent in smaller-pitch photosites has the effect of damping out the interference, yielding a flatter spectral response. The interference pattern is very sensitive to variations in layer thickness, hence, manufacturing process-induced nonuniformities can result in varying spectral response across a given image sensor, as shown in Fig. 19.

Predicting the response an image sensor will have to a specific optical input is often necessary when considering an imager for a particular application. If the absolute spectral energy distribution of the optical input is known, the exact output signal of the sensor can be calculated using the sensor responsivity spectral.

In general, the output of sensors from different optical systems cannot be compared unless the spectral energy distributions of the two light sources are identical, or at least known. Therefore, it is common practice among sensor manufacturers to specify the output of image sensors in response to a standard light source (e.g., CIE standard daylight fluorescent lamp).

The output of an image sensor is calculated by integrating the product of the imager spectral responsivity and the optical spectral irradiance over the entire wavelength spectrum, and then multiplying the result by the total image sensor integration period:

$$V_o = T_i \int_0^\infty R(\lambda) E_s(\lambda)\, d\lambda \qquad (\mathrm{V}) \qquad (4)$$

Here λ is the optical wavelength in nanometers, λ_{min} is the lowest optical wavelength of interest, λ_{max} is the highest optical wavelength of interest, $R(\lambda)$ is the sensor voltage responsivity in $\mathrm{V} \cdot \mathrm{m}^2/\mathrm{J}$, T_i is the optical integration period in seconds, and $E_s(\lambda)$ is the spectral irradiance in $\mathrm{W}/(\mathrm{m}^2 \cdot \mathrm{nm})$. The responsivity is related to the photosite quantum efficiency by

$$R(\lambda) = \frac{\eta\, \mathrm{QE}(\lambda) A_p \lambda}{hc} \qquad (5)$$

where η is the charge-to-voltage conversion factor, and A_p is the photosite's photoactive area. The total irradiance incident at the imager surface is calculated from the spectral irradiance as

$$E = \int_0^\infty E_s(\lambda)\, d\lambda \qquad (\mathrm{W}/\mathrm{m}^2) \qquad (6)$$

and therefore, the total radiant energy density incident upon the imager surface is

$$Q_R = E T_{int} \qquad (\mathrm{J}/\mathrm{m}^2) \qquad (7)$$

The absolute spectral irradiance $E_s(\lambda)$ can be rewritten as a scalar constant A_s times a relative spectral irradiance $S(\lambda)$:

$$E_s(\lambda) = A_s S(\lambda) \qquad [\mathrm{W}/\mathrm{m}^2 \cdot \mathrm{nm}] \qquad (8)$$

As mentioned above, the output response of different image sensors can be compared if both are illuminated with the same type of light source. The quantity used in this comparison is called the radiant responsivity R_R and is defined as the total sensor output voltage divided by the total optical energy density:

$$R_R = \frac{V_o}{Q_R} = \frac{\int_0^\infty R(\lambda) E_s(\lambda)\, d\lambda}{\int_0^\infty E_s(\lambda)\, d\lambda} \qquad [\mathrm{V}/(\mathrm{J}/\mathrm{m}^2)] \qquad (9)$$

If Eq. (8) is inserted in Eq. (9), R_R is seen to be dependent only on the relative optical energy distribution and imager responsivity, which enables direct comparisons of imager outputs using relative optical energy distributions for a given source.

When working in photometric units, the incident spectral irradiance is multiplied by the spectral luminous efficacy

$$K_\lambda = K_m V(\lambda) \qquad (\text{lm/W}) \qquad (10)$$

where $K_m = 683$ lm/W, and $V(\lambda)$ is the spectral luminous efficiency function, representing the relative sensitivity of the human eye (30). For light-adapted, or *photopic,* vision, $V(\lambda)$ peaks at 555 nm, and for dark-adapted, or *scotopic,* vision, $V'(\lambda)$ peaks at 530 nm. Multiplying Eq. (8) by Eq. (10) yields the value of illuminance at the image sensor plane, or

$$E_v(\lambda) = K_\lambda E_s = A_s S(\lambda) K_m V(\lambda) \qquad (\text{lx/nm}) \qquad (11)$$

The total illuminance at the image sensor plane is then

$$E_v = A_s K_m \int_0^\infty S(\lambda) V(\lambda)\, d\lambda \qquad (\text{lx}) \qquad (12)$$

where the total photopic energy density is defined as $Q_I = E_v T_{int}$, in lux-seconds. The quantity used to compare the response of different image sensors in photometric units is called the luminous responsivity R_I, and it is defined as the total sensor output voltage divided by the total photopic energy density, or

$$R_I = \frac{V_o}{Q_I} = \frac{\int_0^\infty R(\lambda) S(\lambda)\, d\lambda}{K_m \int_0^\infty V(\lambda) S(\lambda)\, d\lambda} \qquad [\text{V}/(\text{lx} \cdot \text{s})] \qquad (13)$$

Predicting the output voltage of an image sensor in a typical imaging system is more complicated than the plane wave case discussed above. As illustrated in Fig. 20, one now has to consider the light-gathering properties of the optics and the overall geometry of the imaging system. Referring to Fig. 20, the image sensor will normally reside in the image plane centered on the *x–y* origin. Assuming the distance to the object is greater than the lens diameter and the photosite area is less than the area of the object image, the imager output voltage can be found as a function of off axis angle using (31)

$$V(\theta) = \frac{\pi}{4} \frac{T_i N \cos^4 \theta}{F^2 (1+M)^2} \int_0^\infty R(\lambda) L(\lambda) T_o(\lambda) T_m(\lambda)\, d\lambda \qquad (14)$$

where F is the number of the lens, N is the number of TDI stages (unity for conventional imagers), M is the optical magnification, L is the spectral radiance in W/(cm^2 · μm · sr), T_o is the transmission of the lens, T_m is the transmission of the medium between the lens and object (typically air), T_i is the integration period, and θ is the angle off the center axis in the image plane.

Modulation Transfer Function. For an image sensor to perform well, it must be able to record accurately the spatial information contained in the incident optical image. The modulation transfer function (*MTF*) is the primary measure used to quantify the image sensor's spatial frequency response. For linear, space-invariant systems, the MTF is defined as the modulus of the optical transfer function (*OTF*), which in turn is defined as the Fourier transform of the point spread function (32). Discrete sampling image sensors are inherently space-variant. That is, the recorded image will vary with the position of the image relative to

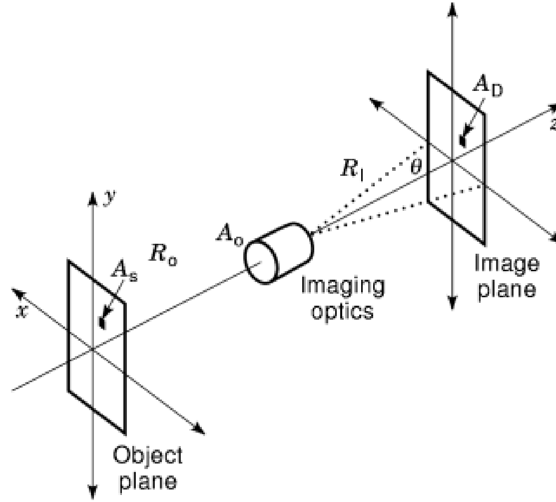

Fig. 20. Layout of typical imaging system. The image sensor would normally be located in the image plane to capture the image of object.

the image sensor, the variations being greatest at high spatial frequencies. This presents some problems when attempting to apply linear systems theory to the spatial frequency response of image sensors (33); however, for the following discussion this point is not considered.

If an image sensor is illuminated with a pure sinusoidal image, the MTF of the reconstructed image is calculated using the definition of normalized contrast as

$$\mathrm{MTF} = \frac{P_{\max} - P_{\min}}{P_{\max} + P_{\min}} \tag{15}$$

where $P_{\max}$ and $P_{\min}$ are the maximum and minimum signal levels taken along the reconstructed sinusoidal image. As the spatial frequency of the sinusoid image is increased, the MTF of the reconstructed image will decrease. There are three main components responsible for degrading the MTF: the discrete sampling effect caused by the photosite aperture, the photosite-to-photosite diffusion crosstalk, and, for CCD-based imagers, the charge transfer inefficiency (CTI) of the shift registers (34). All three components are multiplied together to produce the total imager MTF in either the horizontal or the vertical image space.

The aperture MTF arises from the sampling nature of the photosites. The sampling theorm (35) dictates that the photosite spacing, or pitch, must be less than or equal to $1/(2f_{\max})$ for all of the spatial information in the incident image to be captured and reproduced exactly (i.e., without aliasing), where $f_{\max}$ is the maximum frequency content in the image. As illustrated in Fig. 21(a), when an image sensor is illuminated with a sinusoidal image, each photosite will integrate a portion of the incident pattern $I(x)$. The quantity of charge captured in each photosite, $P(n)$, is proportional to the integral of $I(x)$ over the pixel aperture, or

$$P(n) \propto \int_{nX_p - 0.5X_a}^{nX_p + 0.5X_a} I(x)\,dx = 0.5 \left| x + \frac{\sin(2f_g\pi x + \theta)}{2f_g\pi} \right|_{nX_p - 0.5X_a}^{nX_p + 0.5X_a} \tag{16}$$

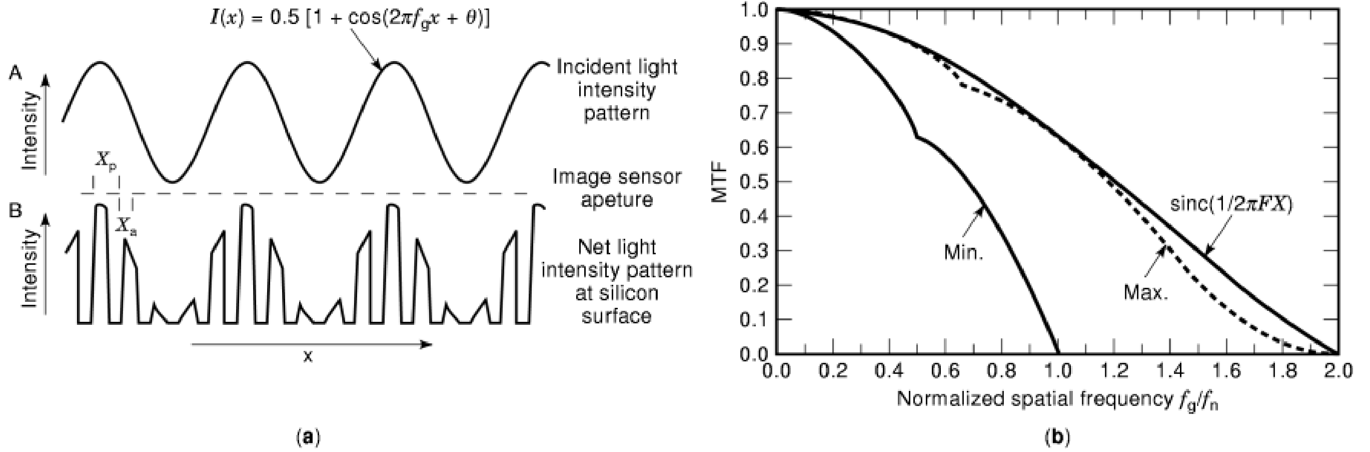

Fig. 21. (a) Effect of image sensor pixel aperture on incident sinusoidal image, and (b) plot of real and ideal aperture MTFs along with minimum possible aperture MTF. Note that the aperture MTF can be as low as zero at Nyquist.

where n is the photosite index, x_p is the photosite pitch, and x_a is the photosite active length. For a 100%-fill-factor photosite, $x_a = x_p$. Clearly, from Eq. (16) the value of $P(n)$ is dependent on the phase relationship, θ, between the incident sinusoidal pattern and the photosite array of the imager (36). Independently solving Eq. (16) for the minimum and maximum possible photosite responses and inserting these into the equation for normalized contrast gives

$$\mathrm{MTF_A} = \frac{P_{\max} - P_{\min}}{P_{\max} + P_{\min}} = \frac{\sin(\pi f_g x_a)}{\pi f_g x_a} = \mathrm{sinc}\left(\tfrac{1}{2}\pi F X\right) \qquad (17)$$

where F is the normalized spatial frequency (f_g/f_n), and X is the photosite aperture-to-pitch ratio (x_a/x_p). Equation (17) is plotted in Fig. 21(b) versus normalized frequency, along with the maximum possible aperture MTF. Using Eq. (17) to model the maximum aperture MTF is common practice and is accurate to within approximately 5% over the range of zero to the Nyquist frequency for photosites with $x_a = x_p$. Between the Nyquist frequency and twice the Nyquist frequency, the sinc function overestimates the MTF and therefore the amount of image aliasing. The minimum aperture MTF is also graphed in Fig. 21(b). The aperture MTF, for a given scene-to-imager phase, may reside anywhere between the minimum and maximum limits.

Photosite-to-photosite diffusion crosstalk originates from the diffusion component of the quantum efficiency, as discussed previously. The quantum efficiency model by Blouke and Robinson (27) can be used to estimate the diffusion component of the MTF by defining

$$\mathrm{MTF_D} = \frac{\mathrm{QE}_k(\lambda)}{\mathrm{QE}_0(\lambda)} \qquad (18)$$

Equation (18) is graphed in Fig. 22(a) for several wavelengths to demonstrate the influence that charge diffusion has on the MTF, and Fig. 22(b) shows the diffusion MTF at Nyquist frequency graphed versus pixel pitch and wavelength using Eq. (18) and a typical set of semiconductor parameters. This latter figure is useful when considering the choice of pixel pitch. For example, at the 650-nm wavelength the diffusion and aperture MTF are equal for a pixel size of approximately 5.2 μm. Therefore, attempts to increase spatial resolution by designing a smaller-pitch photosite would most likely yield disappointing results, as the diffusion MTF rapidly

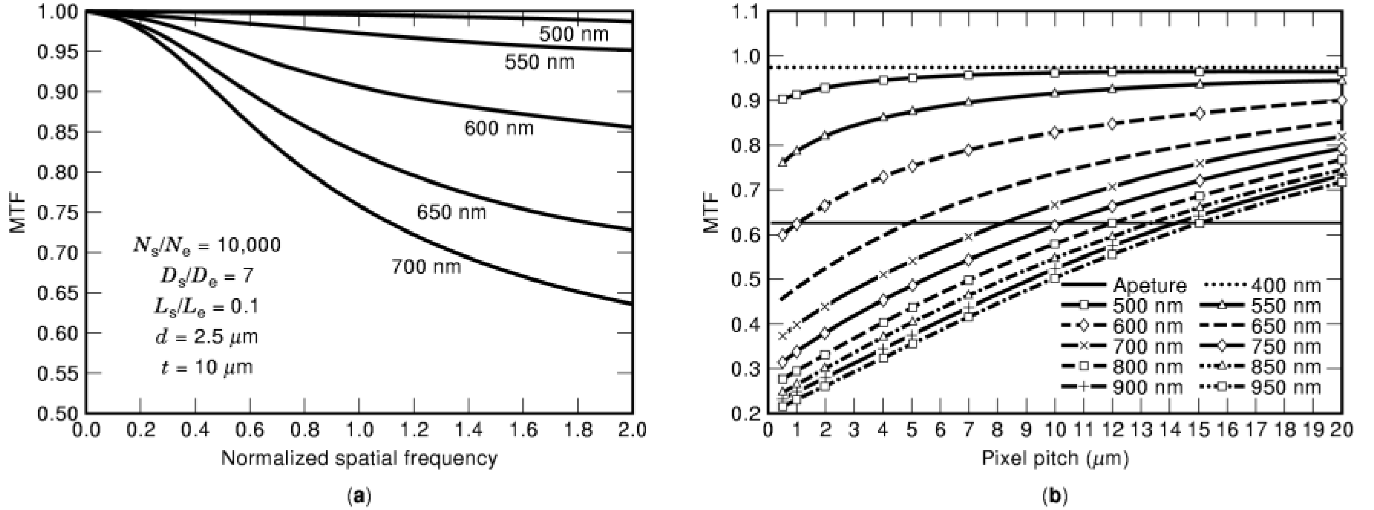

Fig. 22. Diffusion component of MTF versus (a) optical wavelength and (b) pixel pitch and wavelength. Degrading effects of increased mean absorption depth at longer wavelengths limit the minimum usable pixel pitch.

decays for pixels with smaller geometries. Significant changes to the parameters used in Eq. (18) would have to accompany the new photosite design to see improvements in overall imager MTF at 650 nm.

For higher-resolution CCD imagers, the number of charge transfers can be large and poor charge transfer efficiency (CTE) can play a role in reducing the MTF. The net effect of inadequate CTE is the displacement of signal carriers into trailing charge packets. This results in what looks like a tail emanating from brighter regions in captured images. This can be represented in an analytical fashion by considering a uniform series of k pulses, of size V_s, and injecting them into and transferring them through an N-stage shift register, with p phases per stage (37). The pulse series will be modulated by the inefficiency and give rise to a loss per transfer ε given by (38)

$$\epsilon = 1 - \exp\left(\frac{1}{pN}\ln\frac{V_S - V_L}{V_S}\right) \approx \frac{V_L}{pNV_S} \qquad (19)$$

where V_L is the total lost charge from the k injected packets. This process is actually a useful method for evaluating CCDs and gives not only a value of the change loss but also valuable information on the underlying physical mechanism(s) responsible for transfer inefficiency can be obtained by analyzing the leading- and trailing-edge distortion of the readout pulse train. For example, proportional losses, due to the fundamental limitations on charge transport (diffusion- and drift-aided), typically result in a mirrored response in the leading and trailing edges, whereas fixed loss mechanisms, such as surface-state trapping, result in an increased loss in the first pulse (or pulses, until the trapping states are filled within all stages of the shift register) and a nonsymmetric trailing edge. Nonlinear signal-dependent losses can result from barriers or wells in the channel potential profiles caused by deficiencies in processing or cell design, and can produce a pulse train with combinations of the above effects. The MTF due to charge transfer inefficiency can be modeled as (20)

$$\mathrm{MTF_X} = \exp\left[-N\epsilon\left(1 - \cos\frac{\pi f_g}{f_n}\right)\right] \qquad (20)$$

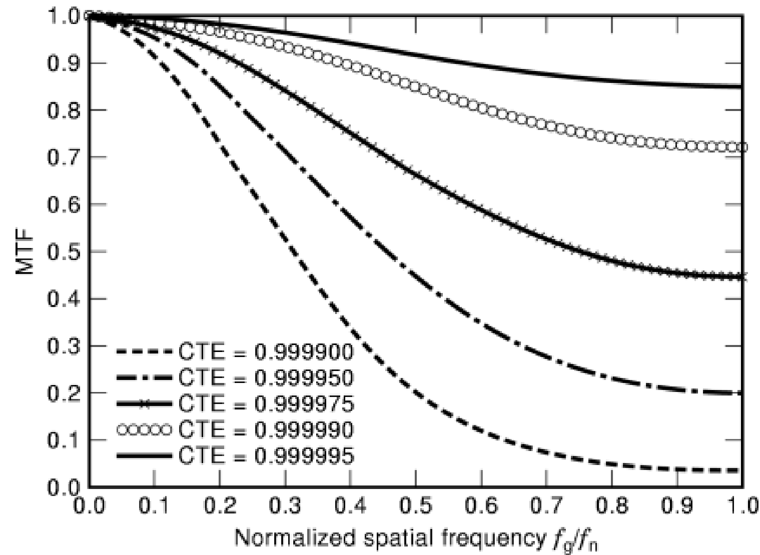

Fig. 23. CTE component of MTF for a two-phase 8000-pixel linear imager.

where N is the number of transfers per register, ε is the CTI per transfer ($\varepsilon = 1 - \text{CTE}$), f_g is the spatial frequency, and f_n is the spatial frequency at the Nyquist point. This is a periodic function with a minimum at the Nyquist frequency for a single-shift-register (monolinear) CCD design. Modern CCD processes can yield values as high as 0.999999 per transfer, and for most imagers, the CTE component is not a significant factor in determining the overall imager MTF. However, for imagers that operate at very low signal levels or temperatures, or high data rates, CTE must be considered. Figure 23 shows the effect on the MTF for efficiencies ranging from 0.999995 to 0.99990.

Area CCD image sensors have two CTE components contributing to the total MTF. One is due to the vertical (y) transfer structure and the other to the horizontal (x) structure. Both are calculated using Eq. (20). CID, APS, and linear imagers typically operate using very few (if any) vertical transfers; hence, CTE in the vertical direction, also referred to as lag, has minimal influence on the overall MTF for these types of imagers.

The total image sensor MTF in a given axis, or readout direction, can be written as

$$\mathbf{MTF_T = MTF_A \times MTF_D \times MTF_X} \qquad (21)$$

TDI image sensors are a special class of sensor, as discussed earlier. These sensors have several additional MTF terms due to the synchronized motion of the image scene over the imager, the possible misalignment of the scene as it travels over the imager, and the velocity tracking error, as described by Wong et al. (20).

Noise. The dynamic range of an image sensor is typically defined as the maximum output signal (i.e., signal saturation level) divided by the rms dark noise (or read noise) and is an important measure of imager performance. Because practical limitations on obtaining high saturation levels are set by photosite capacity, amplifier output range, and input light intensity, minimizing noise sources in both the imager (39,40) and the imaging system (41) is critical to obtaining high dynamic range.

Noise in image sensors is due to a number of factors, including variations in photon flux (n_p), CTI (n_x), input clocking (n_c), thermal dark current (n_d), charge trapping (n_t), kTC noise of the reset MOS transistor (n_k), and the charge-to-voltage factor of the output amplifier (n_a). Expressions for these sources are listed in Table 1

Table 1. Noise Sources in Image Sensors

Noise Source	Value (e^-)	Comments
Photon shot	$(\text{QE}\, E_p A T_i)^{1/2}$	Shot noise of incident photon flux
CTI	$(2\epsilon N)^{1/2}$	Signal fluctuations due to CTI
Clock	$\propto f_{clk}$	Signal variations caused by clock jitter and clocking luminescence
Dark current	$(J_d A_t T_i / q)^{1/2}$	Noise from thermally generated charge
Trapping	$(2nkA_tTN_s \ln 2)^{1/2}$	Surface channel CCD trapping
	$\langle 2nV_sN_t \exp(-t/\tau_e)\,[1 - \exp(-t/\tau_e)]\rangle^{1/2}$	Buried channel CCD trapping
Reset	$(kTC_f / q^2)^{1/2}$	kTC noise of charge-reset MOSFET
Amplifier	$\frac{q}{C_f}\sqrt{\frac{k_f}{WLC_{ox}f^{a}}\Delta f}$	Flicker component
	$\frac{q}{C_f}\sqrt{4kT\frac{2}{3}\frac{1}{g_m}\Delta f}$	Thermal component

for a typical image sensor, and the definitions of the parameters are listed in Table 2. The total rms noise level (n_T) is found by adding the individual components in quadrature:

$$n_T = \sqrt{n_p^2 + n_x^2 + n_c^2 + n_d^2 + n_t^2 + n_k^2 + n_a^2} \qquad (22)$$

Read noise is a term used to describe the noise level of an imager void of all optical input (i.e., $n_p{}^2 = 0$), and, as seen in Fig. 24, represents the effective imager noise floor. Read noise can be reduced by decreasing the operating temperature or through temporal averaging of the imager output signal.

Shot noise arises from the Poisson statistics governing the impinging photon flux and is equal to the square root of the number of captured photogenerated electrons. Temporal averaging can lessen the effect of shot noise, at the penalty of increasing the readout period. At moderate to high signal levels, shot noise is clearly the dominate noise component, as seen in Fig. 24.

CTI noise is a result of charge transfer inefficiency. The fluctuations in readout charge due to CTI are dependent on both the CTI value and the number of shift register transfers (42). The mechanisms responsible for clock noise include jitter in the input and signal processing clocks, clock crosstalk into the output signal as a result of capacitance coupling, and luminescence caused by the clocking of the shift register gates (43).

Dark current is a result of the inherent variations in the thermally generated minority carriers, and is very sensitive to changes in operating temperature, doubling in magnitude for every 5°C to 10°C increase in temperature. Noise attributed to the dark current is described by a shot-type mechanism and therefore is equal to the square root of the integrated dark signal.

Trapping noise is produced by capture and emission of charge carriers from surface and bulk interface states, and can be a significant component when the trap capture and emission time constants are comparable

Table 2. Parameters Used to Calculate Noise Components

Parameter	Description
E_p	Incident photon flux density
A	Photosensitive area
A_t	Total active imager area
T_i	Integration period
QE	Photosite quantum efficiency
e	Charge transfer inefficiency
N	Number of charge carriers
n	Number of transfers
J_d	Dark current density
N_s	Density of surface states
N_t	Density of bulk states
V_s	Volume the charge carriers occupy
T	Absolute temperature
C_f	Capacitance of the charge-to-voltage node
f_{clk}	Shift register clocking frequency
t	Length of time of a shift register transfer
t_e	Bulk trap emission time constant
W	Width of MOSFET gate
L	Length of MOSFET gate
k_f	Device specific constant
C_{ox}	Oxide capacitance per unit area
f	Frequency
Df	Frequency bandwidth
g_m	MOSFET transconductance
k	Boltzmann's constant
a	Transistor dependent constant

to the CCD clocking rate (44). Surface channel CCDs have interface trap densities orders of magnitude greater than buried channel CCDs, and therefore trapping noise can be a dominant noise component in these devices. Many of the traps can be permanently occupied by adding a small fixed background charge, or *fat zero,* to each CCD stage. There is an additional shot noise component associated with the fat zero charge, however, which must be weighed against the potential reduction in trapping noise.

kTC noise is a consequence of the uncertainty in the charge detection node reset voltage due to thermal variations in the reset transistor transconductance (45). It is proportional to the capacitance of the charge detection node. For MOS and CID imagers, this capacitance includes the address line capacitance; therefore this term is large in comparison with CCD and APS imagers. *kTC* noise can be removed using a signal detection scheme termed correlated double sampling (*CDS*). Two common circuit models for CDS are shown in Fig. 25, along with the associated input clocking and output waveforms.

Noise associated with the on-chip amplifier is dependent on the imager architecture and output amplifier configuration (46,47). Typically, the amplifier will have both a thermal (white) noise and a $1/f$ noise component as depicted in Fig. 24(a). The $1/f$ component is significantly reduced by CDS signal processing, at the expense of doubling the thermal noise power. A graph of imager sensor noise versus output signal level, an example of which is in Fig. 24(b), proves extremely useful in evaluating image sensor noise characteristics and performance, and is referred to as the photon transfer function (48).

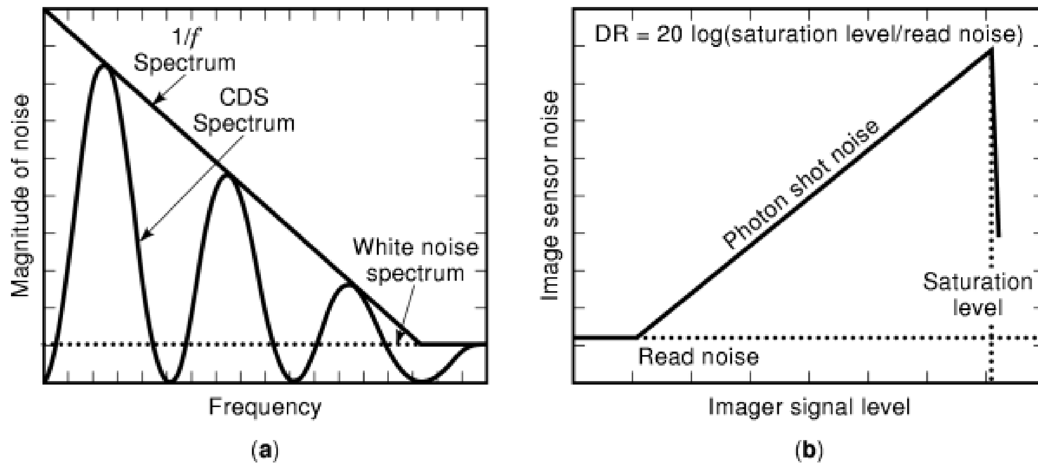

Fig. 24. (a) Typical noise spectrum in CCD image sensor with and without CDS; (b) photon transfer response.

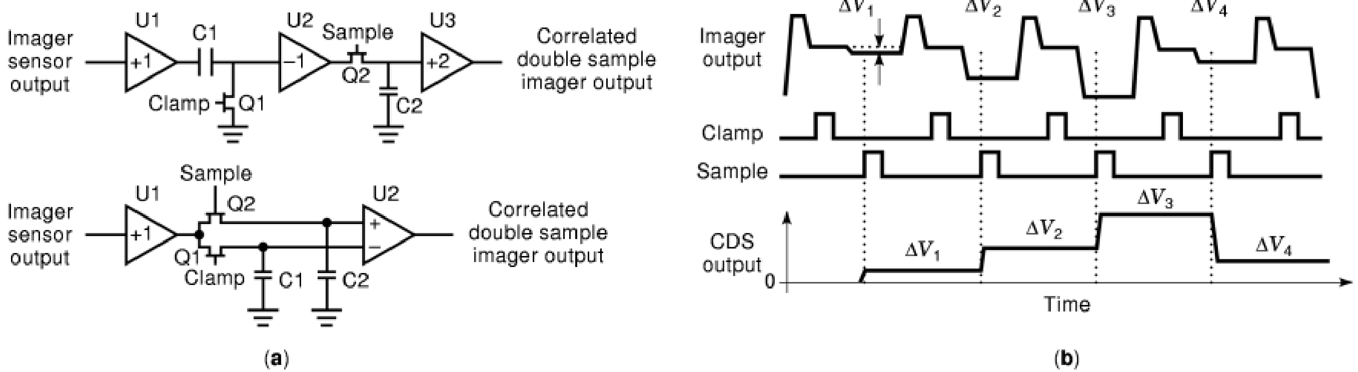

Fig. 25. (a) Electrical schematic of CDS circuits; (b) associated clocking signals. In addition to removing *kTC* noise, the conversion time available for the analog-to-digital converter is now doubled.

BIBLIOGRAPHY

1. P. Wendland, A charge-storage diode vidicon camera tube, *IEEE Trans. Electron Devices*, **ED-14**: 285–291, 1967.
2. R. K. Willardson, A. C. Bear, Semiconductors and semimetals, In *Infrared Detectors II*, San Diego, CA: Academic Press, 1977.
3. R. W. G. Hunt, *The Reproduction of Color*, 5th ed., Fountain Press, 1995.
4. W. Schreiber, *Fundamentals of Electronic Imaging Systems*, 2nd ed., New York: Springer-Verlag, 1991.
5. W. S. Boyle, G. E. Smith, Charge-coupled semiconductor devices, *Bell Syst. Tech. J. Briefs*, **49**: 587–593, 1970.
6. R. H. Dyck, G. P. Weckler, Integrated arrays for silicon photodetectors for image sensing, *IEEE Trans. Electron Devices*, **ED-15** (4): 1968.
7. G. F. Amelio, M. F. Tompsett, G. E. Smith, Experimental verification of the charge-coupled device concept, *Bell Syst. Tech. J.*, **49**: 1970.
8. G. Michon, Method and apparatus for sensing radiation and providing electrical readout, U. S. Patent No. 3,786,263, 1974.
9. C. Greve, *Digital Photogrammetry: An Addendum to the Manual of Photogrammetry*, Am. Soc. Photogrammetry Remote Sensing, 1996.

10. F. Sabins, *Remote Sensing: Principles and Interpretation*, 2nd ed., New York: W. H. Freeman, 1987.
11. E. G. Stassinopoulos, J. P. Raymond, The space radiation environment for electronics, *Proc. IEEE*, **76**: 1423–1442, 1988.
12. G. Hopkinson, Cobalt 60 and proton radiation effects on large format, 2-D, CCD arrays for an earth imaging application, *IEEE Trans. Nucl. Sci.*, **NS-36**: 2018–2025, 1992.
13. G. R. Hopkinson, C. J. Dale, P. W. Marshall, Proton effects in charge-coupled devices, *IEEE Trans. Nucl. Sci.*, **43**: 614–627, 1996.
14. R. E. Hummel, *Electronic Properties of Materials*, New York: Springer-Verlag, 1985.
15. S. Sze, *Physics of Semiconductor Devices*, New York: Wiley-Interscience, 1981.
16. W. E. Engeler, J. J. Tiemann, R. D. Baertsch, Surface charge transport in silicon, *Appl. Phys. Lett.*, **17** (11): 1970.
17. D. F. Barbe, Imaging devices using the charge-coupled concept, *Proc. IEEE*, **ED-63** (1): 1975.
18. M. F. Tompsett, The quantitative effects of interface states on the performance of charge-coupled devices, *IEEE Trans. Electron Devices*, **ED-20** (1): 1973.
19. R. H. Walden *et al.*, The buried channel charge-coupled device, *Bell Syst. Tech. J.*, **51**: 1972.
20. H. Wong, Y. Yao, E. Schlig, TDI charge-coupled devices: Design and applications, *IBM J. Res. Dev.*, **36** (1): 1992.
21. J. Beynon, D. Lamb, *Charge-Coupled Devices and Their Applications*, London: McGraw-Hill, 1980.
22. P. Jespers, F. Van De Wiele, M. White, *Solid State Imaging*, Leyden: Noordhoff, 1976, pp. 447–481.
23. H. Tseng, J. Ambrose, M. Fattahl, The evolution of the solid-state image sensor, *J. Imaging Sci.*, **29** (1): 1985.
24. E. Fossum, Active pixel sensors: Are CCD's dinosaurs? *Proc. SPIE*, **1900**: 1993.
25. E. Hecht, A. Zajac, *Optics*, Menlo Park, CA: Addison-Wesley, 1974, pp. 444–447.
26. D. Seib, Carrier diffusion degradation of modulation transfer function in charge-coupled imagers, *IEEE Trans. Electron Devices*, **ED-21** (3): 1974.
27. M. Blouke, D. Robinson, A method for improving the spatial resolution of frontside-illuminated CCD's, *IEEE Trans. Electron Devices*, **ED-28** (3): 1981.
28. E. G. Stevens, J. P. Lavine, An analytical, aperture, and two-layer carrier diffusion MTF and quantum efficiency model for solid state image sensors, *IEEE Trans. Electron Devices*, **41**: 1753–1760, 1994.
29. F. Flory, *Thin Films for Optical Systems*, New York: Marcel Dekker, 1995.
30. R. Boyd, *Radiometry and the Detection of Optical Radiation*, New York: Wiley, 1983.
31. G. Holst, *CCD Arrays, Cameras, and Displays*, Winter Park, FL: JCD Publishing, 1996.
32. S. Park, R. Schowengerdt, M. Kaczynski, Modulation-transfer-function analysis for sampled image systems, *Appl. Opt.*, **23** (15): 1984.
33. W. Wittenstein *et al.*, The definition of the OTF and the measurement of aliasing for sampled imaging systems, *Opt. Acta*, **29** (1): 41–50, 1982.
34. S. Chamberlain, MTF simulation including transmittance effects and experimental results of charge-coupled imagers, *IEEE Trans. Electron Devices*, **ED-25** (2): 1978.
35. A. Jerri, The Shannon sampling theorem—its various extensions and applications: A tutorial review, *Proc. IEEE*, **65**: 1565–1596, 1977.
36. J. Feltz, M. Karim, Modulation transfer function of charge-coupled devices, *Appl. Opt.*, **29** (5): 1990.
37. R. W. Brodersen, D. D. Buss, A. F. Tasch, Experimental characterization of transfer efficiency in charge-coupled devices, *IEEE Trans. Electron Devices*, **ED-22** (2): 1975.
38. M. Kimata *et al.*, Low-temperature characteristics of buried channel charge-coupled devices, *Jpn. J. Appl. Phys.*, **22** (6): 975–980, 1983.
39. M. Gupta, *Electrical Noise: Fundamentals & Sources*, New York: IEEE Press, 1977.
40. E. Dereniak, D. Crowe, *Optical Radiation Detectors*, New York: Wiley, 1984.
41. H. W. Ott, *Noise Reduction Techniques in Electronic Systems*, 2nd ed., New York: Wiley, 1988.
42. A. Mohsen, M. Tompsett, C. Sequin, Noise measurements in charge-coupled devices, *IEEE Trans. Electron Devices*, **ED-22** (5): 1975.
43. J. Janesick *et al.*, The future scientific CCD. In *State-of-the-Art Imaging Arrays and Their Applications, Proc. SPIE*, **501**: 1984.
44. M. F. Tompsett, Quantitative effects of interface states on the performance of charge-coupled devices, *IEEE Trans. Electron Devices*, **ED-20**: 45–55, 1973.
45. R. W. Broderson, S. P. Emmons, Noise in buried channel charge-coupled devices, *IEEE Trans. Electron Devices*, **ED-23** (2): 1976.

46. P. Centen, CCD on-chip amplifiers: Noise performance versus MOS transistor dimensions, *IEEE Trans. Electron Devices*, **38**: 1206–1216, 1991.
47. P. R. Gray, R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed., New York: Wiley, 1977, pp. 664–667.
48. J. Janesick, K. Klaasen, T. Elliott, CCD charge collection efficiency and the photon transfer technique. In *Solid State Imaging Arrays, Proc. SPIE* **570**: 1985.

ROBERT H. PHILBRICK
HERBERT J. ERHARDT
Eastman Kodak Company