The CCD sensor: A semiconductor circuit for capturing images. On the Nobel Prize in Physics awarded to Charles Kuen Kao, Willard S. Boyle, and George E. Smith (II)*

Atilà Herms

Department of Electronics, Faculty of Physics, University of Barcelona, Barcelona, Spain

Resum. El Premi Nobel de Física 2009 va atorgar el guardó compartit a dos científics (a més de Charles Kuen Kao): Willard S. Boyle i George E. Smith, ambdós dels laboratoris Bell de Murray Hill, a Nova Jersey, per haver inventat, el 1969, un dispositiu electrònic, el sensor CCD (*charge-coupled device* o 'dispositiu de càrrega acoblada'), que ha permès un desenvolupament important de les càmeres digitals, tant de fotografia com de vídeo. Només un any més tard, els laboratoris Bell ja tenien en funcionament la primera videocàmera basada en CCD. En càmeres comercials està essent substituït per sensors CMOS, però en certs sistemes de captura encara és el més important, especialment pel baix nivell de soroll. En aquest article s'analitza el funcionament del dispositiu i l'evolució que té.

Paraules clau: sensor CCD · fotografia i vídeo digitals · xarxes òptiques · laboratoris Bell · matrius d'estat sòlid · fotodetecció · transferència de càrrega

Abstract. The Nobel Prize in Physics 2009 was jointly awarded to two scientists (in addition to Charles Kuen Kao): Willard S. Boyle and George E. Smith, both of Bell Laboratories in Murray Hill, New Jersey, for having invented, in 1969, an imaging semiconductor circuit, the CCD (charge-coupled device), an electronic device which allowed an important development of digital cameras, both photography and video. A year later, Bell Labs already had the first CCD-based camcorder. In commercial cameras it is being replaced by CMOS sensors, but in some capture systems it is still the most important component, especially because of its low levels of noise. This article examines the functioning and evolution of the device.

Keywords: CCD sensor · digital photography and video · optical networks · Bell Labs · solid state arrays · photodetection · charge transfer

The Nobel Prize in Physics 2009 was jointly awarded to two scientists (in addition to Charles Kuen Kao): Willard S. Boyle and George E. Smith, both of Bell Laboratories in Murray Hill, New Jersey, for having invented, in 1969, an imaging semiconductor circuit, the charge-coupled device or CCD (Fig. 1.)

Applications of the CCD sensor

The most important aspect in the development of the CCD sensor is the device's applications in image capture. The CCD has fostered the advent of digital photography and video, with major advances for professional and amateur use, including the availability of image capture systems with features and at prices that would have been unimaginable in the not-too-distant past. Ever since its invention, the CCD has generated important new in-

Fig. 1. Left to right, Willard S. Boyle and George E. Smith. © The Nobel Foundation. Photos: Ulla Montan.

dustries and markets, ushering in a wide range of products, including digital cameras, camcorders, high-definition television, security monitoring, medical endoscopy, and equipment for use in modern astronomy. The underlying principles of the CCD have also proven to be crucial in the advent of optical networks, which comprise the transport technology underlying communication networks. Since 1983, telescopes have been equipped with solid state CCD cameras that enable astronomers to study objects that are thousands of times fainter than those studied with the most sensitive photography plates. Today, the majority of optical observatories, including the Hubble space telescope, are based upon CCD camera systems.

Fig. 2. (**A**) One of the earliest CCD devices from Bell Labs. (**B**) Kodak camera. (**C**) Sony Mavica camera.

Fig. 3. Evolution of the market for image capture devices [Source: Theuwissen A (2005)].

When Boyle and Smith envisioned the design of what they called "charge 'bubble' devices," the essence of these devices was the possibility of transferring a charge along the surface of a semiconductor, with the main application and goal being the capacity to store the charge, that is, the information. The first CCD was initially invented as a memory device, and it worked by injecting the charge through an input register. It soon became clear that a CCD could accumulate charges using the photoelectric effect, allowing electronic images to be created, stored, and then extracted for viewing or processing.

In 1971, M.F. Thompsett and other researchers at Bell Labs captured images using simple linear devices, and thus the CCD image capture device was born. Soon several companies, including Fairchild Semiconductor, RCA, and Texas Instruments, seized on the promising future of the device and began to develop research programs around it. For example, Fairchild, with G. Amelio (former Bell Labs researcher) at the helm, unveiled the first commercial image capture device, and in 1974 it commercially sold two-dimensional matrices with 100×100 pixels. In 1975, Kodak's Steve Sasson manufactured a camera with a CCD sensor from Fairchild. It weighed four kilos. The sensor was black and white and had 0.01 megapixels (Mp). The results were not totally convincing, but it was an encouraging first step (Fig. 2A,B).

Sony, led by Kazuo Iwama, also made major efforts and investments in developing CCDs for its cameras. When Iwama died, in 1982, a CCD sensor was placed in his grave in recognition of his role as a driving force behind the use of CCDs. In 1981, the Sony Mavica was brought to market, the first commercially sold camera with a 570×490 (0.280 Mp) CCD sensor. The Mavica captured black and white analogue images and saved them onto magnetic disks, while color images had to be viewed on a television monitor (Fig. 2C).

Since then CCDs have become the leading sensor technology for camcorders, and more recently for cameras. The first digital camcorder appeared in 1991, the Logitech Dycam 376×240, followed in 1992 by the first digital reflex camera, the Kodak DCS 200, with 1.5 megapixels. The first color digital camera, the Apple Quicktake, was launched in 1994, and the first mobile telephone with a built-in digital camera in 2000, by Sharp. Since then, the development of features and resolution has been spectacular. Today, household cameras use 12-megapixel sensors, and high-end professional cameras are equipped with a 50-megapixel CCD. A 196-megapixel CCD, for use in aerial photometry, was recently announced.

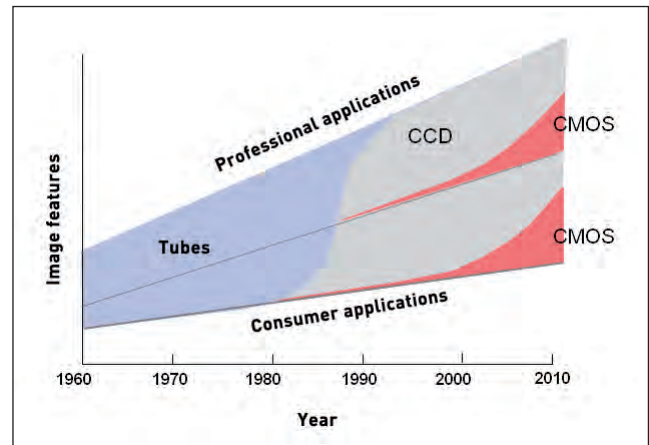In image sensors, we can primarily distinguish between two groups of applications. The first includes those that seek to imitate the workings of the human eye (cameras, photography, video, etc.), while the second includes applications that outstrip the human eye (high-speed, metrology, and extended wavelength, operating conditions, sizing, etc.). Analogous to the attempts to develop computers mimicking the human brain, solid-state image sensors have both advantages and limitations compared to the human eye. Lately, CCD sensors are being replaced in consumer applications by CMOS (complementary metal-oxide semiconductor) sensors, but the CCD is still the leader in the professional and scientific sectors. Figure 3 shows the approximate evolution in the use of CCD sensors and the recent growth in CMOS sensors.

In both fields of application, CCD sensors are the main device used to capture images. Let us examine the schema of a digital reflex camera as a paradigm for the image capture system (Fig. 4). The figure shows the sensor (CCD or CMOS) and the electronic processing circuits: (i) The first circuit includes the noise elimination methods and the adjustment of the gain from the amplifiers preceding the converter, simulating the behavior of the ISO/ASA sensitivity scale in the first film cameras, followed by the analogue/digital converters, which digitalize the signal of each pixel with a number of bits that depends on the pre-established precision of each camera. (ii) The second step consists of a circuit for white balance compensation, tone adjustments, contrast, color saturation, and other image processing features that depend on the characteristics of each camera, and the compression to JPEG format, with differing levels of quality. In CCD sensors, these circuits are distinct from the sensor in other chips; if the sensor uses CMOS technology, both the sensor and the processor may be integrated into a single chip, given that this technology is the same as integrated circuit manufacturing (one of the advantages of CMOS circuits). Generally speaking, in these cases the sensor and first processing circuit are usually integrated into the chip, but the second processing circuit is not, although, in the future, the development of 3-D integration technologies might enable the integration of all the processing electronics.

Like the reflex camera (or SLR camera), CCD sensors and processors electronics have replaced the film used in the past.
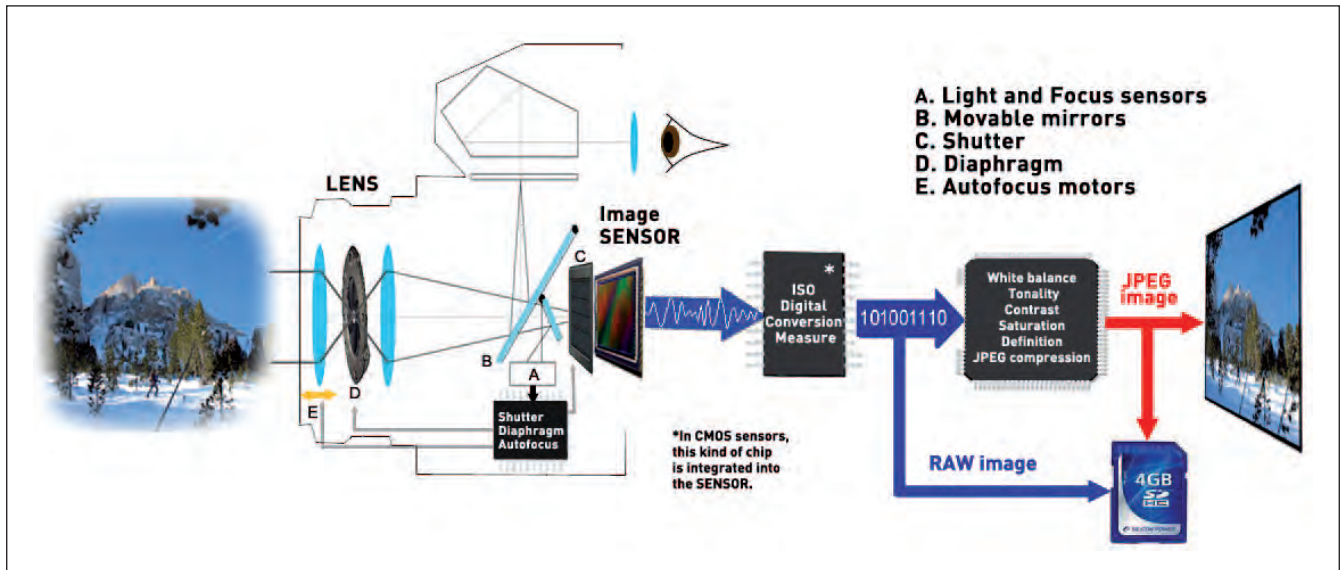
Fig. 4. Schema of an SLR (reflex) camera [Source: Bosch J (2009)].

As mentioned above, the earliest cameras contain sensors ranging from 3 Mp to more than 60 Mp and for all sorts of applications. The ideal number of pixels in a sensor compared to photographic film has yet to be defined. There are studies suggesting that 100 megapixels is needed to achieve the resolution of the human eye; yet according to some estimates a resolution equivalent to more fine-tuned photographic film can be reached with 18-megapixel sensors. In practice, results equal to film are achieved with 8- to 10-Mp sensors. However, other factors play a crucial role, such as the light conditions and sensitivity. The number of pixels is mainly important if the goal is to print large-sized photographs. The table below shows the maximum size of a printed photograph as a function of the number of megapixels of the original (Table 1). The size of the sensor is important in determining the resolution of the image captured. The larger the sensor, the higher the quality of light received per pixel, but the signal/noise ratio also increases, just as with photographic film.

Table 1. Relationship between the number of megapixels and the maximum size of a photograph

Megapixels	Without interpolation (cm)	With interpolation (cm)
3	12.7 × 17.8	28 × 36
4	15.2 × 20.3	33 × 48
5	15.2 × 22.9	33 × 48
6	17.8 × 25.4	41 × 61
8	20.3 × 30.5	46 × 71
11	22.9 × 35.6	51 × 76
14	25.4 × 38.1	61 × 91
16	27.9 × 43.2	76 × 101
22	33.0 × 48.3	101 × 152

Solid state matrices

The image capture sensor is made up of a solid state matrix. The acronym CCD (charge-coupled device) refers to an architecture of semiconductor photodetectors that generate a charge according to the incident light. This charge is then transferred through storage areas. CCD architecture has three basic purposes: (a) to generate and collect the charge, (b) to transfer the charge, and (c) to convert the charge into measureable resistance. The device that makes up each of the pixels in the matrix is a silicon MOS device. The light charge generated in each pixel is proportional to the amount of incident light. The aggregate effect of all the pixels produces a spatial representation of a scene or image.

Photodetection

To understand how a CCD sensor works, we must first examine the physical process of photodetection by the semiconductor material. The most appropriate semiconductor for capturing images, as explained below, is silicon. Silicon is a semiconductor with a gap, meaning the area where energy cannot exist, in this case measuring 1.12 eV in width (Fig. 5A). It is the energy separation between the valence and conduction bands; in other words, the energy that must be supplied to a valence electron such that it is able to free itself and move along the conduction band within an electrical field. In its place, the gap that the electron leaves in the valence band can be moved under the effects of the electrical field (actually, another bonded electron occupies the vacant position, but the effect is as if a positively charged particle, called a hole, has moved in the opposite direction) (Fig. 5B).

The relative concentration of electrons and holes in each band depends on the temperature. A rise in the semiconductor's temperature leads to a vibration in the crystal structure that causes the release of some electrons, by providing them

with sufficient kinetic energy than the gap. We thus talk about thermally generated electron-hole pairs. The distribution function of these electrons is explained by the Fermi function:

$$F(E) = \frac{1}{\exp\left(\dfrac{E-E_f}{kT}\right)-1}$$

However, the energy needed to release an electron can be supplied by a photon. A photon with energy higher than the forbidden band causes an electron to leap from the valence to the conduction band and generates an electron-hole pair. Of all the semiconductor materials, silicon is the best for capturing light in the visible spectrum (Fig. 5A). The minimum energy needed for a photon to release an electron is E = 1.12 eV, and the corresponding wavelength is:

$$\lambda = \frac{ch}{E}$$

in which c is the speed of light and h the Planck constant. Therefore, the maximum wavelength that an electron can release is 1.110 nm, which corresponds to the infrared region, such that wavelengths in the visible spectrum produce electron-hole pairs. As the radiation begins to approach the ultraviolet spectrum, the light's penetration of the silicon also drops significantly, rendering it impossible for pairs to form. These factors
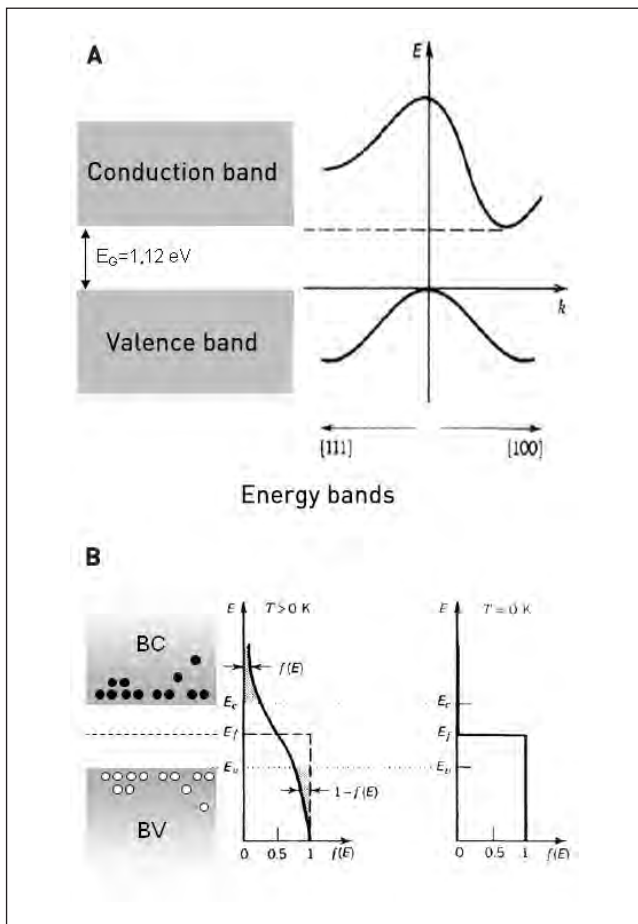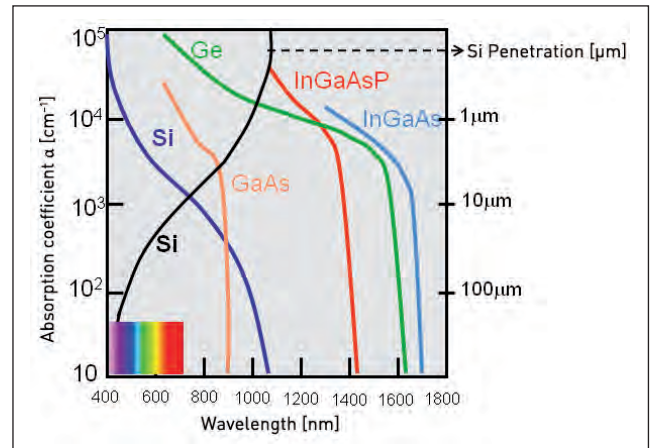
Fig. 6. Absorption and light penetration spectrum according to the wavelength of silicon and other semiconductors.

account for silicon being the most appropriate material for visible light detectors. Figure 6 illustrates why other semiconductor materials are better for detecting light in the infrared zone.

Generation-collection of photogenerated charges

Once electron-hole pairs have been generated, they must be separated in order to prevent them from recombining again. This is achieved through the electrical field applied via a structure or device that enables the accumulation of the generated charge. The structure or device used in CCDs is the MOS condenser. When positive resistance is applied to the gate or metal, the electrons are attracted to the oxide, where they are confronted with the potential barrier made up of the oxide, and they accumulate there. The holes migrate to the substrate, as shown in the figure. The effect is like creating a potential well where the charge accumulates, namely, the electrons produced by light in this case, while some electrons are thermally generated, which correspond to noise. In low lighting conditions, the proportion of thermally generated electrons can be large, and the noise therefore more prominent.

Operating the CCD: Charge transfer

Once the charge has accumulated in each pixel, it is transferred to the output, which is the characteristic feature of the CCD. As mentioned above, the CCD was initially designed to be used as memory; however, it actually provides a record of displacement, in which the charge is transferred to a neighboring pixel at each clock signal, until the entire charge has been extracted. It should be borne in mind that when the CCD was developed, microelectronic processors were not as advanced as they are today, and it was impossible to establish connections for each pixel in order to extract the signal from every one. Therefore, designing a device that would continuously transfer the charge to its neighbor was a logical solution.

The CCD sensor is made up of a matrix of MOS condensers, in which the gate tensions are controlled appropriately and

Fig. 5. (**A**) Energy bands in a semiconductor (silicon). (**B**) Distribution of electrons according to the temperature of silicon.

sequentially by transferring the charges horizontally and vertically. Each gate or metal has its own control signal (clock). When the correct positive resistance is applied to the gate of a MOS condenser, a potential well is created in which the photo-generated charge is stored. If resistance is applied to two neighboring condensers, the two potential wells overlap in one of them and the charge is redistributed.

As shown in Fig. 7, if the resistance drops in one of the MOS condensers, the well becomes smaller and the charge moves on to the second well. When the resistance in the first condenser is 0, the entire charge has been shifted to the neighboring well. The process is similar to the transfer of water in communicating vessels by raising one container over the other. The movement of the electrons is caused by: (a) thermal diffusion, (b) self-induced electrical fields, and (c) lateral electrical fields caused by the resistance applied to the gate. In this case, the latter is what triggers the electrons' movement from one well to the other. In reality, a CCD pixel is made up of three MOS condensers. The one in the center is the condenser that actually receives the light, while the other two serve to avoid interferences between neighboring pixels and to transfer the charge. These two condensers must be properly screened to prevent them from continuing to generate a charge after the capture.
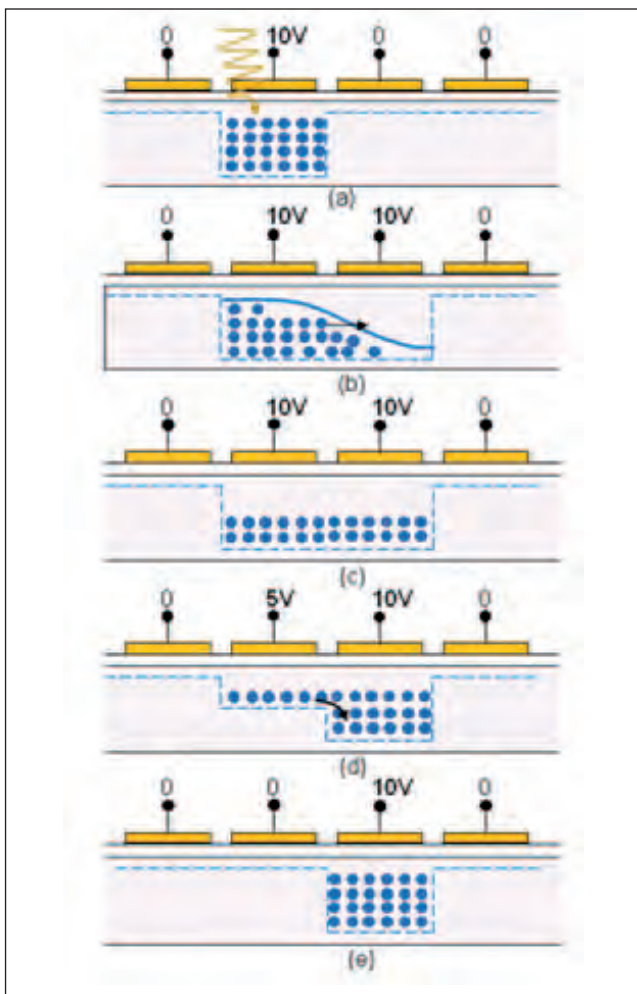

Fig. 7. Movement of the charge (electrons) between neighboring wells by applying a sequence of tensions at the gates.

The control signals at the gate of MOS condensers are laid out in different phases that show a drop in the resistance on the descending slope of each pulse, which causes the well to drop and the charge to move to the neighboring condenser. This is an example of three-phase charge transfer. There are also other transfer schemas, such as two-phase, four-phase, and virtual phase. For equal pixel sizes, the four-phase transfer system is the one with the highest well capacity. In today's CCDs, the channel where the charge accumulates and which is used to transfer the charge to neighboring wells is buried so as to lower the noise from the charge capture by the surface states of the $Si-SiO_2$ interface. To bury the channel, there is an n-type layer just under the oxide which causes the well to be slightly displaced towards the interior of the substrate. As mentioned above, in order to prevent the image from deteriorating during the charge transfer to the outside, the registers where the charge is generated differ from the transfer registers in that the latter are protected from light.

Images are transferred vertically; when one row has moved to the serial read-out register the charges are serially transferred to the charge-resistance amplifier/converter, which is directly connected to the analogue/digital (A/D) converter and which in turn transfers the corresponding digital value to each pixel in the memory. The transfer speed depends directly on the bandwidth of the amplifier and the capacity of the A/D converter. In the case of a high number of pixels, the transfer speed is boosted by dividing the matrices into submatrices, each of which has its own amplifier and A/D convertor.

There are different CCD architectures in relation to the transfer process. The most common ones are the interline-transfer CCD and the frame-transfer CCD. Interline architecture tends to be used in consumer CCDs, while the frame architecture is more common in CCDs for scientific applications. In interline-transfer architecture, columns are interspersed with sensors, which are clad in a metal layer to transfer the image crosswise. After the light is captured, the photogenerated charge is transferred to the vertical registers in less than 1 μs; this transfer is so quick that no shutter is needed. Each row is then transferred to the horizontal register, which quickly transfers the row serially to the amplifier; the next row is then transferred, and the following one, until the last row. In this case, the fill-factor (percentage of the light detection area with respect to the total chip size) is not very high, but it can be slightly improved by using microlenses to concentrate the light on the detecting surface.

In frame-transfer architecture, the chip contains two virtually identical matrices, one for detecting light and the other for storage. The storage cells are protected with a metal layer. After the capture cycle or integration, the charge is quickly transferred to the storage cells; the transfer time to the covered matrix is approximately 500 μs. Once the image has been transferred to the storage matrix, it is moved row by row to the horizontal read-out register. The fill-factor in this case is very large.

The output amplifier converts the charge transferred to the well of the last MOS condenser into a resistance proportional to the charge. The resistance value is then digitalized by the A/D converter.
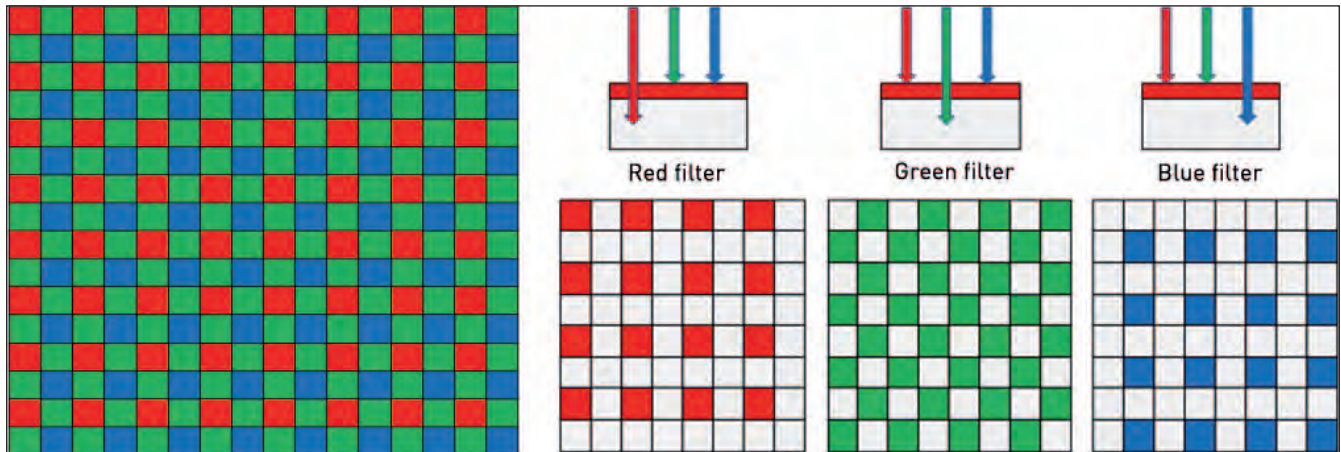
Fig. 8. Left, Bayer mask. Right, Generation of the color components in each pixel with a Bayer mask.

Generating color

In the aforementioned process, a charge proportional to the light intensity is collected, but there is no chromatic information. To generate color, a system of basic color filters is applied: red (R), green (G), and blue (B). Each pixel has to have the basic color components. One solution is to use three sensors with one filter in each sensor corresponding to each component. This system is employed in professional camcorders, but in photography it is neither practical nor economical.

In household camcorders and cameras, a sun sensor is used, and the information in three color components from each pixel is obtained via RGB filters, in which there is a filter for one given component on each pixel. There are different ways to distribute the filters, including primary colors, complementary colors, color bands, and mosaics. The most widely used filter is called the Bayer mask (Kodak), in which the green is repeated to emulate the behavior of the human eye, which is more sensitive to this color (Fig. 8). When the filter is applied, each pixel only receives one color component; it obtains the other two via interpolation. Figure 8 provides a simple example of how the red components are obtained in the pixels that do not receive this color (Ra, Rb, Rc, Rd, and Re), i.e., through the interpolation of the red component of neighboring pixels (R1, R2, R3, and R4), which have a red filter on them.

The other two colors are generated in the same way, by interpolating information from neighboring pixels. Each camera manufacturer uses its own interpolation algorithms. Since the interpolation process can lead to poor image-quality optical effects, generally on the edges, each manufacturer tries to adjust its algorithms as much as possible. In many cases, depending on the transference architecture of the CDD, the fill-factor is less than 100%, primarily due to the restrictions on integration, especially in interline-transfer architecture, in which the fill-factor can be as low as 20%. Placing microlenses on top of each detector considerably boosts the fill-factor. The output of each pixel is directly proportional to the area of the detector. The use of microlenses raises the effective size of the detector and therefore the output resistance.

Improvements in CCD detection

One of the improvements applied to CCDs to boost their quantum efficiency is back-illuminating the device via the substrate. The front produces absorptions, interferences, and multiple reflections from the contacts, the polycrystalline-Si gates, and silicon oxide, and the transmittance of polycrystalline-Si drops starting at 600 nm and becomes opaque at 400 nm. All of these effects lead to a diminution in the detector's spectral response.

The capture of light from the back eliminates all of these effects and improves the efficiency for wavelengths under 600 nm. The problem in this case is the thickness of the substrate (>250 μm). The light penetration by silicon is drastically reduced at wavelengths close to 400 nm (blue tones), so the photons are absorbed and generate electron-holes far removed from the draining zone where the well is formed, and they are recombined without producing a response. Accordingly, the substrate must be made thinner, reduced to a thickness of 10 μm. Thus, thin devices illuminated via the substrate avoid the effects of the interferences triggered by the upper layers of the MOS condensers. Moreover, quantum efficiency levels of around 85% can be achieved with an appropriate anti-reflection layer. The substrates are thinned using a chemical
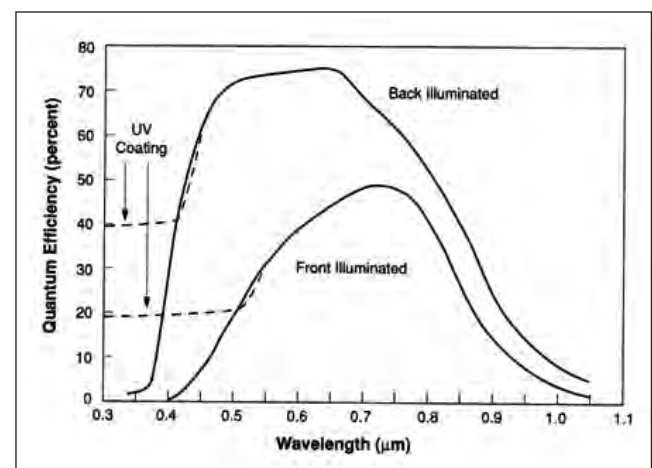

Fig. 9. Quantum efficiency of a front-illuminated and back-illuminated detector [7].

process after the device is manufactured, which makes them considerably more expensive. Consequently, back-illuminated CCD's are primarily used in scientific applications or high-end cameras (Fig. 9).

One of the main problems that must be eliminated in light detectors is noise. There are different sources of noise in CCDs, such as photon shot noise, dark current noise, fixed-pattern noise, reset noise, amplifier noise, and quantification noise. A variety of techniques can be used in processor electronics to lower the noise, such as correlated double sampling; alternatively, the device can be cooled, since dark current noise mainly arises from the thermal generation of electrons, which are coupled with photogenerated electrons that in turn depend directly on the temperature. Table 2 shows the drop in dark current according to the temperature.

Table 2. Dependence of dark current on temperature [7]

Temperature (°C)	Dark current (pA/cm2)
60	1136
40	276
20	55.8
0	9.02
−20	1.11
−40	$9.58 \times 10-2$
−60	$5.53 \times 10-3$
−80	$1.66 \times 10-4$

The cooling technique is based on a thermoelectric cooler, i.e., a Peltier device that extracts heat from the CCD. Cooling is also used in scientific applications (mainly in astronomy) and in high-end cameras. As shown above, the spectrum detected by silicon detectors is limited to between 400 and 1100 nm; therefore, X-rays are completely out of their range. A layer of shiny material over the CCD absorbs the radiation of X-rays at energies of around 10 KeV and results in the production of luminescence at a peak of 550 nm, which can easily be detected by a silicon CCD.

Brief biography of the Nobel Prize winners

Willard S. Boyle (Amherst, Nova Scotia, Canada, 1924) earned his Bachelor's degree in 1948 and his PhD in 1950 from McGill University. He spent two years as a Physics Professor at the Royal Military College of Canada. In 1953, he joined Bell Labs, where together with Don Nelson he invented the first continuously operating ruby laser (1962). That same year, he was appointed Director of the Space Science and Exploratory Studies, which supported the Apollo space program and contributed to choosing lunar landing sites. He returned to Bell Labs in 1964, where he worked on developing integrated circuits. In 1969, he and George E. Smith invented the CCD, for which they were awarded numerous prizes. In 1975, he was appointed Executive Director of Research at Bell Labs, where he remained until he retired in 1979. Since then he has moved to Wallace, Nova Scotia (Canada). He is married and has four children, ten grandchildren and one great-grandchild.

George E. Smith (White Plains, New York, USA, 1930) earned his Bachelor's degree from the University of Pennsylvania in 1955 and his PhD from the University of Chicago in 1959 (with a three-page dissertation). From 1959 until his retirement, he worked at Bell Labs, where he coordinated research on new lasers and semiconductor devices. He also directed Bell Labs' VLSI Department. He and Willard S. Boyle invented the CCD in 1969. Upon retiring, he and his wife Janet spent 5 years sailing around the world. He now lives in Waretown, New Jersey.

To learn more

www.dalsa.com/sensors/Products/ccd.aspx

www.dpreview.com [Digital Photography Review]

www.harvestimaging.com

www.advancedimagingpro.com/publication/pub.jsp [Advanced Imaging Magazine]

www.outbackphoto.com/dp_essentials [Photography using digital SLR]

www.bhphotovideo.com/c/browse/Photography/ci/989/N/4294538916

www.ieeeghn.org/wiki/index.php/Oral-History:George_E._Smith [Interview with G.E. Smith]

http://inventors.about.com/od/cstartinventions/a/CCD.htm

http://nobelprize.org/nobel_prizes/physics/laureates/2009/press.html

Bibliography

1. Amelio GF, Tompsett M, Smith GE (1970) Experimental Verifications of the Charge Couple Concept. Bell Systems Technical Journal 49:593-600

2. Bockaert V (2007) The 123 of Digital Imaging. Interactive e-book [www.123di.com] Bosch J (2009) Curs de fotografia digital. UB, Barcelona

3. Bossiers JT et al. (2006) Technical challenges and recent progress in CCD imagers. Nuclear Instruments and Methods in Physics Research A565:148-156

4. Bota S (2009) Curs 'Ulls de silici'. Universitat de les Illes Balears, Mallorca

5. Boyle WS, Smith GE (1970) Charge Coupled Semiconductor Devices. Bell Systems Technical Journal 49:587-593

6. Boyle WS, Smith GE (1971) Charge-coupled devices-A Bs new approach to MIS device structures. IEEE Spectrum 71:18-27 (IC09)

7. Holst GC (1998) CCD Arrays Cameras and Displays, 2nd ed. JCD Publishing, Winter Park, FL

8. Janesick J (2000) Scientific Charge-Coupled Devices. SPIE Press Monograph, vol. PM83

9. Theuwissen A (1995) Solid-State Imaging with Charge-Coupled Devices. Series: Solid-State Science and Tech-

nology Library, vol. 1. Kluber Academic Publishers, Dordrecht, The Netherlands

10. Theuwissen A (2005) CMOS Image Devices. FSRM Course on Training on Microsystems

11. Theuwissen A (2009) The 2009 Nobel Prize in Physics: W. Boyle and G. Smith for the CCD. Europhysics News 40(6)

12. Tompsett MF, Amelio GF, Bertram Jr WJ, Buckley RR, McNamara WJ, Mikkelsen Jr JC, Sealer DA (1971) Charge-coupled imaging devices: Experimental results. IEEE Transactions on Electron Devices 18(11):992-996