High Performance Silicon Imaging

Related titles:

Subsea optics and imaging
(ISBN 978-0-85709-341-7)

Handbook of terahertz technology for imaging, sensing and communications
(ISBN 978-0-85709-235-9)

Nitride semiconductor light-emitting diodes (LEDs)
(ISBN 978-0-85709-507-7)

High Performance Silicon Imaging

Fundamentals and Applications of CMOS and CCD Sensors

Edited by
Daniel Durini

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made.

For information on all Woodhead Publishing publications
visit our website at http://store.elsevier.com/

Typeset by Toppan Best-set Premedia Limited, Hong Kong

Printed and bound in the United Kingdom

(* = main contact)

Editor

D. Durini
Forschungszentrum Jülich GmbH
Central Institute of Engineering,
  Electronics and Analytics
ZEA-2 – Electronic Systems, S2425
  Jülich, Germany
(Formerly of the Fraunhofer
  Institute for Microelectronic
  Circuits and Systems, Duisburg,
  Germany)

E-mail: d.durini@fz-juelich.de

Chapters 1 and 2

D. Durini*
Forschungszentrum Jülich GmbH
Central Institute of Engineering,
  Electronics and Analytics
ZEA-2 – Electronic Systems, S2425
  Jülich, Germany
(Formerly of the Fraunhofer
  Institute for Microelectronic
  Circuits and Systems, Duisburg,
  Germany)

E-mail: d.durini@fz-juelich.de

D. Arutinov
Fraunhofer Institute for
  Microelectronic Circuits and
  Systems
Finkenstr. 61
D-47057 Duisburg, Germany

E-mail: david.arutinov@ims.
  fraunhofer.de

Chapter 3

M. Lesser
Steward Observatory
University of Arizona
UA Imaging Technology
    Laboratory
325 S. Euclid Ave Suite 117
Tucson, AZ 85721, USA

E-mail: lesser@itl.arizona.edu

Chapter 4

A. Lahav*
Tower Semiconductor Ltd
Ramat Gavriel Industrial Park
20 Shaul Amor Avenue
P.O. Box 619
Migdal Haemek 23105, Israel

E-mail: asafla@towersemi.com

A. Fenigstein and A. Strum
TowerJazz
4321 Jamboree Road
Newport Beach, CA 92660, USA

E-mail: avi.strum@towerjazz.com

Chapter 5

B. Choubey*
Department of Engineering
    Science
University of Oxford
Oxford OX1 3PJ, UK

E-mail: bhaskar.choubey@eng.
    ox.ac.uk

W. Mughal and L. Gouveia
School of Engineering
University of Glasgow
Room 519, Rankine Building
Oakfield Avenue
Glasgow G12 8LT, UK

Chapter 6

D. Ginhac
LE2I UMR 6306 – Université de
    Bourgogne
Aile des Sciences de l'Ingénieur
9, avenue A. Savary – BP 47870
21078 Dijon CEDEX, France

E-mail: dginhac@u-bourgogne.fr

Chapter 7

R. J. Gove
Aptina Imaging
3080 North First Street
San Jose, CA 95134, USA

E-mail: rjgove@icloud.com

Chapter 8

C. De Locht* and H. Van den
    Broeck
Melexis Technologies NV
Transportstraat 1
B-3980 Tessenderlo, Belgium

E-mail: cde@melexis.com; hav@
    melexis.com

Chapter 9

J. Bogaerts
CMOSIS nv, Belgium
Coveliersstraat 15
B-2600 Antwerpen, Belgium

E-mail: jan.bogaerts@cmosis.com

Chapter 10

R. Turchetta
Rutherford Appleton Laboratory
Science & Technology Facilities
    Council
Harwell Science and Innovation
    Campus
Chilton, Didcot, Oxfordshire OX11
    0QX, UK

E-mail: renato.turchetta@stfc.ac.uk

Chapter 11

R. K. Henderson*
University of Edinburgh
South Bridge
Edinburgh EH8 9YL, UK

E-mail: robert.henderson@ed.ac.uk

B. R. Rae
ST Microelectronics
33 Pinkhill
Edinburgh EH12 7BF, UK

D.-U. Li
Centre of Biophotonics
Strathclyde Institute of Pharmacy
    and Biomedical Sciences
University of Strathclyde
Glasgow G4 0RE, UK

Chapter 12

A. Strum* and A. Fenigstein
TowerJazz
4321 Jamboree Road
Newport Beach, CA 92660, USA

E-mail: avi.strum@towerjazz.com

Chapter 13

P. Centen
Grass Valley Nederland BV
Kapittelweg 10
4827 HG Breda, The Netherlands

E-mail: peter.centen@grassvalley.
   com

Chapter 14

S. Nikzad
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91125, USA

E-mail: Shouleh.Nikzad@jpl.nasa.
   gov

Woodhead Publishing Series in Electronic and Optical Materials

1
Fundamental principles of photosensing

D. DURINI, Forschungszentrum Jülich GmbH,
Germany and formerly Fraunhofer Institute for
Microelectronic Circuits and Systems, Germany and
D. ARUTINOV, Fraunhofer Institute for
Microelectronic Circuits and Systems, Germany

Abstract: This chapter discusses the fundamental theory behind photosensing and imaging. It starts with a brief overview of the human vision system, and continues with the basic ideas behind photometry and radiometry used as radiation measuring systems. Next, it presents a brief history of phototransduction explaining the origins of solid-state photosensing through the development of quantum mechanics and the different applications of the photoelectric effect, and mentions some important developments in the field of photodetectors prior to silicon based microelectronics.

Key words: human vision, photometry, radiometry, photoelectric effect, history of photodetectors.

> 'Nature and nature's laws lay hid in night;
> God said 'Let Newton be' and all was light.'
> *Alexander Pope*

> 'It did not last: the devil, shouting 'Ho.
> Let Einstein be', restored the status quo.'
> *Sir John Collins Squire*

1.1   Introduction

Since ancient times people have been trying to create images that could reflect their experiences, explain the world surrounding them and conserve their memories in a visual form. Since the very first mosaic paintings (also known as an 'abaciscus' or 'abaculus') (see Fig. 1.1), the same concept has been pursued: putting together hundreds or thousands of small coloured tiled stones or pieces of clay (named 'tesserae'), used as basic picture elements or 'picture cells' (pixels), a much bigger single final image can be created. The smaller these picture elements are and the more different intermediate values between complete darkness and complete illumination or different individual colours they might possess, the better is the resolution and the quality of the resulting image. The concept of mosaic painting has

1.1 Detail showing the Greek god Dionysus, part of the central image of the Dionysus Mosaic from around 220 – 230 AD currently situated in the Römisch-Germanisches Museum in Cologne in Germany (and still in its original place, where it was used as a kitchen floor). The Dionysus Mosaic has an area of 10.6 m x 7 m and was made out of approximately 1.5 million stone tiles (photograph taken by Daniel Durini in 2012).

been known for several thousand years: the earliest known mosaics made of different materials were found in the temple building in Abra in Mesopotamia, dated to the second half of the 3rd millennium B.C.

1.2   The human vision system

The concept of mosaic painting proved to be very successful mainly because it was developed based on the empiric knowledge of the functionality of our own human vision system and the ability of our brain to interpret the incoming information in a logical manner. Human vision is based on an
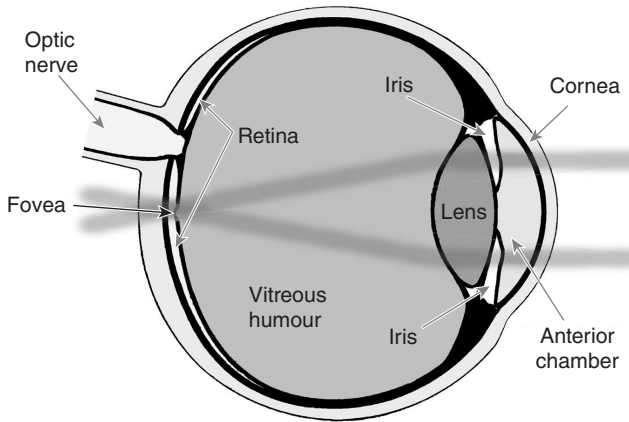
1.2 A schematic diagram of the human eye.

optical system projecting an image through a lens, the cornea, the vitreous fluid, and a layer of capillaries to focus it through several layers of neural membranes onto a system of passive cone photoreceptors located in the centre of the retina directly beneath a small cavity called the fovea (Davson, 1976), as it can be observed in Fig. 1.2. Currently accepted vision theories suggest that human beings use a tri-chromatic system to detect and separate colours, with photonic energy being measured using three different types of band pass cone-shaped absorbing photoreceptors (Curcio *et al.*, 1990; Rosenthal *et al.*, 2004), and the ability of the brain to combine separate fragments into one logical image entity.

According to Østerberg (1935, as cited by Ripps and Weale in Davson, 1976), the human retina contains approximately 110–125 million rods and 6.3–6.8 million cones, the two kinds of photoreceptors in the human eye. As concluded by Max Schultze (Schultze 1866, as cited by Ripps and Weale in Davson, 1976), they are associated with scotopic (nocturnal) and photopic (diurnal) visions, respectively, which formed the basis for the so called 'duplicity theory' defined to explain the visual ability of the human retina, with their properties as listed in Table 1.1.

As described in Østerberg's thorough survey (Davson, 1976), the high concentration of cones at the foveal centre (called 'fovea centralis') is of around 195,000 per mm^2, as can be observed in Fig. 1.3, abruptly falling to about 9,500 per mm^2 in the parafovea region (~2 mm from the foveal centre), and changes little from there to the retinal border – the 'ora serrata'. The cones in the region of the foveal centre, with an average spacing of approximately 1.9 μm between the neighbouring cells, are responsible for delivering a highly resolved image to the brain. Rods, on the other hand,

Table 1.1 The physiological basis of the duplicity theory in a human being (Ripps and Weale in Davson, 1976b)

Parameter	Rods	Cones
Operating conditions	Dim light	Daylight
Sensitivity	High	Low
Spatial resolution	Poor	Good
Temporal resolution	Poor	Good
Maximal sensitivity	Blue–green	Yellow–green
Directional sensitivity	Slight	Marked
Rate of dark adaptation	Slow	Fast
Colour vision	Absent	Present
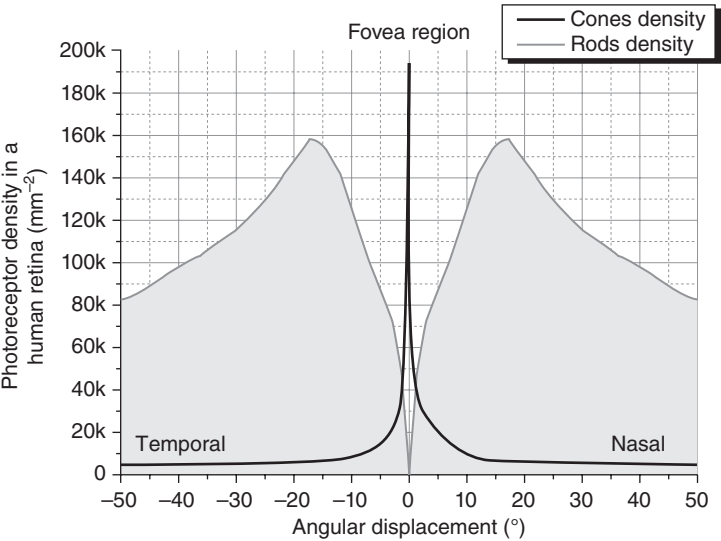
1.3 Average density of rods and cones in a human retina, plotted as a function of the angular separation from the fovea region (after data published by Ripps and Weale (1976a) in Davson, 1976).

are first encountered 130 µm from the foveal centre, their numbers reaching a maximum some 5–6 mm from the fovea centre, and then decreasing gradually toward the far periphery (Ripps and Weale in Davson, 1976). The mostly rod-populated regions of the human retina are reserved for the low-resolution essentially peripheral vision (Curcio *et al.*, 1990). This is why the ocular movements are performed for focused 'scanning' of the observed scenery, using for this task an impinging radiation angle of approximately 3° in contrast to the 120° or more (Curcio *et al.*, 1990) of an average human field of view. The signals sent to our brain can be pretty much defined as a

sort of a video stream, where a high-resolution highly focused vision signal is obtained through 'scanning' the observed scene using ocular movements and a variable pupil aperture to adjust to changing illumination for this task, projecting this high-quality information on top of a low-quality brighter scene image used as a background.

Since the diameter of the rod-free area (in the 'fovea centralis') is about 260 µm, corresponding to a visual angle of 53′ of arc, from the information presented above it can be concluded that in spite of the high receptor density of this area, it contains fewer than 8,000 cones, i.e. less than 0.2% of the total population of cones in the human retina, and less than 0.006% of the total number of photoreceptors, responsible for the 'high-resolution' visual information flowing into the brain. The density of rods and cones, plotted as a function of the angular separation from the fovea region, can be observed in Fig. 1.3 (after data from Ripps and Weale (1976b) in Davson, 1976).

Both cones and rods vary in size with retinal locus, becoming in general longer and more slender when going from the far periphery to the central retina. In the rod-free 'fovea centralis', typical cones are 40 µm long and about 0.9 µm in diameter. A peripheral cone outer segment, on the other hand, is approximately 10 µm in length and 2.5 µm across its base. The rods on the periphery of the retina are about 20 µm long and 2.5 µm in diameter; those in the region of the posterior pole measure 30 µm × 1 µm (all sizes taken from Ripps and Weale in Davson, 1976).

Human evolution determined the visual abilities we have today, which combined with the absorption, reflexion and scattering mechanisms taking place in front of the human retina yield the part of the spectra an average human being is able to experience. This part of the spectra is the one we normally call 'visible', which on average covers the wavelengths of radiation impinging the human eye in the region between approximately 420 nm (which we perceive as near ultra-violet and blue radiation) and 700 nm (which we perceive as red in colour). We are able to register only an extremely small part of the existing radiation, which makes our visual system although fascinating, rather limited in this sense.

To include studies about the human visual system into general sensory physiology, they have to be compatible with the biophysical principles which give a rational picture of the mechanisms involved. It is not enough to deliver a fixed average number of quanta of light (photons) in a specified narrow wavelength (λ) band lying between λ and ($\lambda + \Delta\lambda$) with a well defined temporal and spatial profile (Ripps and Weale in Davson, 1976). A significant fraction of the quanta of light may never reach the retina owing to losses due to reflection, scattering and absorption in the pre-retinal media; the exposure time may be meaningless if an eye being stimulated moves from its original position; and the geometry of the stimulating light

beam may be drastically changed if the observer's pupil diameter alters. Nevertheless, the first attempts to standardize the human visual perception were performed at the end of the 19th century, giving birth to the first definitions that form part of what we now know as photometry: measurement of electromagnetic radiation detectable by the human eye.

To quantify the human eye's ability to see fine detail, the term 'acuity' was defined (Cline *et al.*, 1997), which is a psychophysical metric for determining the minimal separation distinguishable between two point images or two dark bars (lines). In Curcio *et al.* (1990) the acuity, defined for the period of the highest spatial frequency grating being twice the angular subtense of line-to-line spacing, was of 77.1 cycles per degree or 0.78′ per arc on average for 37-year-old individuals. For the acuity measurements, resolution of gratings consisting of alternating light and dark bars (lines) requires that at least one row of not stimulated cones lie between rows of stimulated cones (Helmholtz, 1924, cited in Curcio *et al.*, 1990). Thus, it can be assumed that the minimal spatial resolution per single photoreceptor is half of the minimum distance between two distinguishable lines or two stimulated cones, i.e. 0.39′ of arc. If we try, despite the fact that the human visual system is a dynamic one, to express the pixel resolution an image sensor should have in order to distinguish the spatial detail a human brain may define based on the 'video stream' the eyes send to it, where 120° is the average angle for the human field of view in all directions and 0.39′ of arc being the acuity of a single photoreceptor, we would come to the conclusion that it could be delivered only by the astonishing amount of more than 340×10^6 pixels. Of course, having two eyes and using the incoming information from both of them helps the brain to increase the image resolution. It is nevertheless difficult and even pointless to compare the human vision system, understood as a video stream where a low-resolution field of view is additionally 'scanned' by a high-resolution focused one, with an image sensor aiming to reach the human eye image resolution in a single shot. New theories (Rosenthal *et al.*, 2004) suggest that the human retinal cones and rods might even be direct electromagnetic wave detectors rather than being simply photon counters, which in case they might prove true makes such comparisons even more unsuitable.

Finally, the luminance range of the human vision system exceeds 120 dB, a contrast ratio approaching 10^6 to 1 for a specific scene (Rosenthal *et al.*, 2004). The only minor problem if observed from the point of view of modern imaging systems or metrology applications, is the time required by the human vision system to adjust to changing illumination conditions and/ or reach full potential in what spatial resolution is concerned; in both cases, we are speaking about several seconds or even minutes by changing illumination conditions and at least several tens of milliseconds in

constant conditions. As stated in Ripps and Weale (1976b) in Davson (1976), if rod functionality is under test, then the period of dark-adaptation, considered in the dynamic range of the human vision system, should not be shorter than 30 minutes, a period that might be cut to no more than 10 minutes if it is the cones that are needed at maximum sensitivity.

Although millions of years of natural evolution might seem really hard to beat, solid-state photosensing is used nowadays in many applications in which the main goal is not only to create high-quality images to be perceived by human beings, but also to quantify this visual information beyond the visible range and be able to quantify the detected information in terms of irradiance, distance, velocity, etc.

1.3   Photometry and radiometry

The universal stimulus specified for the study of human visual perception was white light. To define it, in 1900 Max Planck described the electromagnetic radiation of the so called 'black body' (Planck, 1900), i.e., a radiation originated in a cavity in thermodynamic equilibrium with rigid opaque walls that might not be perfectly reflective at any wavelength, and that follows what is now called Planck's law as expressed by Eq. [1.1]. The radiation defined in Eq. [1.1] is assumed to be isotropic, homogeneous, unpolarized and incoherent.

$$L(T) = \frac{2hc^2}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \qquad [1.1]$$

In Eq. [1.1] $L(T)$ represents the temperature dependent radiance of the black body measured as the quantity of radiation that is emitted from its surface that falls within a given solid angle in a specified direction (expressed in watts per steradian per square metre or $W \cdot sr^{-1} \cdot m^{-2}$), T is the absolute temperature (in K), λ is the wavelength of the emitted radiation, k_B is the Boltzmann constant, h is the Planck constant, and c is the speed of light. For purposes of calibration of the stimulus used for the measurement of the human visual response, as defined by the General Conference on Weights and Measures (CGPM: http://bipm.org/en/convention/cgpm/) in 1946 and modified in 1967, this stimulus was defined as the light emitted by a standard black body of absolute temperature of 2042 K: the temperature of freezing platinum under a pressure of 101,325 $N \cdot m^{-2}$ (Kodak, 2008). In 1979, the CGPM established a new definition of the visual stimulus, defining it as the luminous intensity in a given direction of a source that emits monochromatic radiation of frequency 540×10^{12} Hz and that has a radiant intensity in that direction of 1/683 $W \cdot sr^{-1}$ (Kodak, 2008). One square centimetre of such a radiator defines the radiometric unit of candela (cd)
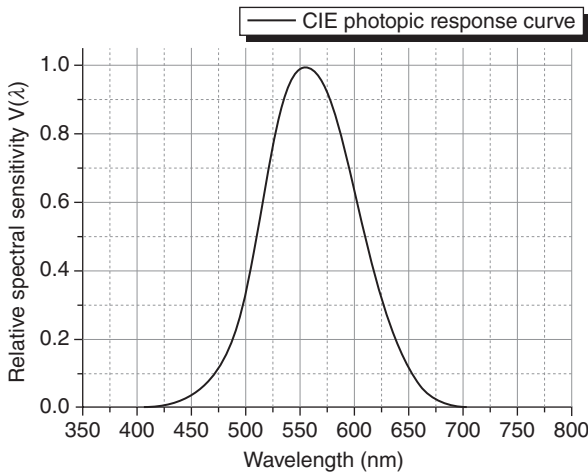
1.4 Photopic relative luminous sensitivity, as used by the CIE, based on data provided in Ripps and Weale (1976b) in Davson (1976).

which is equal to $\frac{1}{60}$ of the luminous intensity radiated by it perpendicularly to its orifice. A hypothetical point-source of 1 cd emits a luminous flux of 1 lumen (lm) per steradian. The flux density due to such a source at a distance of one metre is 1 lm/m^2 or 1 lux.

The most adopted standard and dimensionless spectral sensitivity of the human eye under diurnal conditions measured as its 'felt' response to the electromagnetic irradiation, described using the due photometric definitions was proposed by the International Commission on Illumination (CIE, 1983). It was described by the so called photopic spectral luminous efficiency function $V(\lambda)$ established in 1924, and shown in Fig. 1.4. The peak of this function is situated at 555 nm wavelength (green light) of impinging radiation. The CIE scotopic spectral luminous efficiency function, $V'(\lambda)$, established in 1951, has not been used in practical photometry and relates to human visual performance at very low levels of illumination and peaks at about 507 nm of irradiance wavelength (CIE, 1983). Nevertheless, even a function as firmly standardized as this one can vary systematically with a number of parameters: variations of the colour vision due to variations in the lenticular and macular pigmentation or the variation of the point of pupil entry of the stimulating beam during the measurements, to mention just a couple of them (Ripps and Weale in Davson, 1976), which makes the $V(\lambda)$ function lack any real physical meaning.

Eventually, the need arose to use the International System of Units (SI) to describe the electromagnetic radiation as a purely physical quantity over its entire possible spectra. The study of radiation under these terms became

the main task of radiometry, in contrast to the photometric quantities that are based on the response of the human eye. The defined units in radiometry are divided into two conceptual areas. The first one deals with the energy of radiation (measured in joules) or the temporal changes of that energy defined as power or radiant flux Φ (measured in joules per second or watt), related to the photometric unit of lumen – the measure of luminous flux Φ_v. In this sense, Eq. [1.2] (Kodak, 2008) defines a radiant flux Φ emitted by a source, proportional to the area enclosed by its spectral (wavelength dependent) distribution curve $\Phi(\lambda)$ which must be known.

$$\Phi = \int_0^\infty \Phi(\lambda)d\lambda \qquad\qquad [1.2]$$

The second conceptual area of radiometry is related to quantities that are geometric in nature, such as:

- irradiance or the photon flux density E_R measured in W/m^2, related to the photometric unit of lux – a measure of illuminance E_v
- the radiant intensity I_R, measured in W·sr^{-1}, related to the photometric unit of candela – a measure of luminous intensity I_v or
- radiance L_R defined as power per unit projected area per unit solid angle, measured in W·m^{-2}·sr^{-1} and related to the photometric entity of luminance L_v, measured in cd/m^2, among others.

To determine the capacity of the radiant flux Φ to create the sensation of light, the $\Phi(\lambda)$ curve from Eq. [1.2] must be transformed into a luminous flux (measured in lumen) by multiplying it by the photopic relative luminous efficiency function $V(\lambda)$ depicted in Fig. 1.4, and the maximum luminous efficacy factor K_m as expressed in Eq. [1.3] (Kodak, 2008).

$$\Phi_v = K_m \cdot \int_{\lambda_{min}}^{\lambda_{max}} \Phi(\lambda) \cdot V(\lambda)d\lambda \qquad\qquad [1.3]$$

For photopic vision, the maximum luminous efficacy of radiant flux as derived from its definition mentioned above, is $K_m = 683$ lm/W. In Eq. [1.3] λ_{min} and λ_{max} can be set to define the wavelength bandwidth in which the product of $\Phi(\lambda)$ and $V(\lambda)$ is non-zero. Practically, this involves using the visible spectral range.

1.4   History of photosensing

With time, once first radiation emitters were developed, where the specific wavelength of the emitting light could be defined and a specific time resolved active illumination became available, people started using photosensors not only to preserve visual memories or to communicate

with each other, but also for measurements of different natural phenomena and even for system controlling purposes. This is why modern applications of photosensors addressed in this book do not necessarily remain in the domain of imaging devices, a change that has proven to challenge the photosensing technology performances in unforeseen manners.

Radiation detection far beyond the visible part of the spectra in both direction of X-rays and high-energy γ particles (photons) on the one side and millimetre wavelengths on the other, single-photon counting ability for extremely low radiation, sub-micrometre spatial resolution or picosecond timing resolutions, are nowadays all not that extremely exceptional system requirements, mostly found in metrology, machine vision, scientific or medical applications. In the second decade of the 21st century, it can be stated that state-of-the-art photosensors are almost entirely fabricated using different semiconductor materials, the huge majority of which profit on maturity, yield, availability, and cost-performance ratio of the silicon based manufacturing technologies. Thus, this chapter and all the following ones are focused entirely on describing the different aspects of silicon based phototransduction.

The end of the 19th century was marked by the significant progress made in many areas of natural sciences and humanities. Physics was not an exception. Almost all phenomena known in those days were perfectly described by existing physical models with very high accuracy. The belief in the 'last days' of physics as a research discipline was so strong that many physicists speculated with the idea of there being nothing left to discover. The famous British physicist, Sir William Thomson (Lord Kelvin), said in his lecture given at the Royal Institution in London in April 1900, entitled 'Nineteenth century clouds over the dynamic theory of heat and light' (Thomson, 1901), that physics as a science is almost complete, and the only thing left is to provide more and more precise measurements of already known quantities and phenomena. The 'clouds' darkening the bright reality described in this lecture were just two phenomena left unexplained: the black body radiation and the earth motion through the light-bearing aluminiferous ether, a hypothetical mechanical medium used to explain propagation of light (replaced in modern physics by the theory of relativity).

The first 'unexplained phenomenon' is closely related to the paradox of classical physics, commonly known as the 'ultraviolet catastrophe', which states that the spectral density of the thermal radiation of an ideal black body (a light bulb or the sun could be considered as a black body to some extent) being at thermal equilibrium will tend to infinity at short wavelengths. In those times, two different rules were used to describe this type of radiation: Wien's displacement law (Mehra and Rechenberg, 1982) and Rayleigh–Jeans law (Rayleigh, 1900). The former is used to describe the

radiation at high frequencies but disagrees with the experiment in the low frequency domain. The latter, on the contrary, works well at low frequencies but fails at high ones. The problem was solved by German physicist Max Planck through the invention of the quantum theory of radiation. He postulated that electromagnetic radiation can only be emitted as discrete energy corpuscles – quantas – which are proportional to the frequency of the radiation as expressed in Eq. [1.4], where E_{ph} is the energy of the emitted radiation, v its frequency, and h the proportionality constant named after Max Planck. The frequency of radiation could also be expressed in terms of its wavelength λ, as also shown in Eq. [1.4].

$$E_{ph} = hv = \frac{hc}{\lambda} \qquad\qquad [1.4]$$

Based on Eq. [1.4], Planck developed a quantum theory of black body radiation which predicted finite energy emission and was also consistent with other known experimental results. The theory was inconsistent with experimental results and could continuously describe the whole spectrum of the black body radiation. Eventually, it became a fundamental part of modern quantum theory.

The second 'cloud' from the Lord Kelvin lecture was the inconsistency of the omnipresent moving ether with several experimental facts, mainly those obtained from the Michelson–Morley experiment, performed by the American physicists Albert Michelson and Edward Morley in 1887 (Michelson and Morley, 1887). This experiment aimed to define the dependence between the speed of light and its direction of travel. In case of success, it would prove the existence of moving ether. Negative results of the experiment shook the foundations of the broadly accepted concept at that time, turning it into a paradox. The inconsistency was solved after the 'special theory of relativity' (Einstein, 1905a) was developed by Albert Einstein in 1905. In one of his papers (Einstein, 1918), Einstein explained why the concept of relative motion with respect to the ether was unnecessary and even inconsistent from the physics point of view. The approach using this concept assumed different treatments of the same particular phenomena depending on the frame of reference chosen for its observation. The theories based on the concepts of quantum mechanics and the theory of relativity, which have formed the basis of the modern physics, encountered opposition at the beginning of the 20th century. Some wished to 'save' the ether concept; others refused the statistical nature of the quantum theory. At some point even Einstein doubted the statistical nature of physical phenomena. His famous quote 'I, at any rate, am convinced that He (God) does not throw dice' (Born, 1971) proves his apprehensions. Max Planck once said 'A new scientific truth does not

triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it' (Kuhn, 1970).

The development of the two theories started a new era in physics. The following years gave rise its many new discoveries. Although a detailed discussion of these discoveries remains beyond the scope of this chapter, they will be briefly discussed in the next paragraph with special focus on atomic physics and quantum mechanics.

After the discovery of the electron, the British physicist Sir Joseph John Thomson proposed a model for the matter's main building brick: an atom (Thomson, 1904). His atom consisted of a spherical volume filled by a positively charged medium populated by negatively charged objects – electrons as he named them some years before (Thomson, 1897), in the way schematically shown in Fig. 1.5. Some years after this first model, the New Zealand-born British physicist Ernest Rutherford showed through his experiments that the atom appeared to be a rather empty structure, consisting of a positively charged heavy core (the atomic nucleus) and point like negatively charged particles orbiting around it in a way similar to that of the planets orbiting the Sun in the solar system model (Rutherford, 1911). Being based on classical theory, the model could not explain the stability of such a construct. The charged electrons would be expected to radiate and lose energy while accelerating, and should eventually fall on the core destroying the atom. The model was then extended by the Danish physicist Niels Hendrik David Bohr who introduced his famous postulates
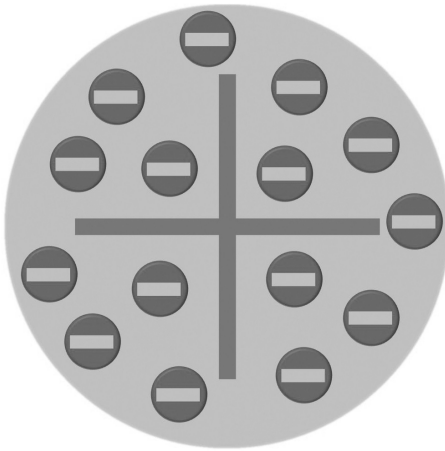
1.5 Graphical representation of Thomson's 'plum-pudding model'. Negatively charged localized areas distributed inside the positively charged medium.

in which he stated that electrons in the atom could occupy only certain well defined energy levels without emitting radiation (Bohr, 1913). The combination of Bohr's theory and Rutherford's experimental results will be later known as the Rutherford–Bohr model of the atom.

The next step towards better understanding of the atom and the matter surrounding us was made by Louis-César-Victor-Maurice, 6th Duke de Broglie, who extended the wave–particle duality concept of the photon to all elementary particles, considering not just the massless photons, but also all other particles possessing a mass (de Broglie, 1924). This helped explain quite a few observations made with different scattering experiments. All this was followed by the development of the matrix mechanics introduced by Werner Heisenberg, Max Born and Ernst Pascual Jordan in 1925 (Born *et al.*, 1925), and the wave mechanics introduced by Erwin Schrödinger in 1926 (Schrödinger, 1926), which resulted in the first self-sufficient quantum-mechanical formalism and generalization of the Rutherford–Bohr theories. Finally, in order to further understand the two clouds defined by Lord Kelvin, the relativistic and quantum theories were combined by the British physicist Paul Adrian Maurice Dirac (Dirac, 1928) and, accompanied by other discoveries of that time, gave birth to the concept of an 'antiparticle', a simultaneous co-product of each natural particle having the same mass but opposite charge.

All these theories helped explain a lot of already known phenomena and allowed us to discover many new ones. Looking beyond the clouds from Lord Kelvin's lecture made nuclear energy available, helped develop the first transistor, enabled the production of microchips, and basically shaped the world in the way we know it today. They led to a revolution in both physics and philosophy, and changed the lifestyle of humanity for ever.

Following these exciting developments in the fields of quantum mechanics and solid-state physics, an entire family of radiation detectors appeared based on the so called 'photoelectric effect'. The history of the photoelectric effect started in 1839 when French physicist Antoine-Henri Becquerel discovered the photovoltaic effect – the change of the electric properties of a material when exposed to impinging light (Becquerel, 1839). Almost 50 years after this discovery, a German physicist Heinrich Rudolf Hertz demonstrated that if a spark gap, an arrangement of two conducting electrodes separated by a gap usually filled with gas that is designed to allow an electric spark to pass between the conductors, is exposed to ultraviolet light, the spark will appear at lower voltages applied to the electrodes than in the absence of light. In other words, light causes negatively charged particles (electrons) to be 'knocked-out' from the orbitals of the atoms of the metallic electrode (cathode), effectively lowering the breakdown voltage of the spark gap. Throughout the following decade the effect was studied at first by the Russian physicist Aleksandr Stoletov (Stoletov, 1888) and

then, more intensively by the French physicist Alfred-Marie Lienard. Their work came up with the definition of the quantitative laws (although still lacking the explanation of the causes of such behaviour) followed by the photoelectric effect (Demtröder, 2010) as follows:

- The amount of electrons emitted from the surface of the metal per unit of time is directly proportional to the intensity of light impinging on it.
- Every material has so called long-wavelength cutoff of the photoelectric effect: a minimal frequency at which the photoelectric effect is still possible.
- The kinetic energy of the emitted electrons is proportional to the frequency of the incident light and is independent from the light intensity.
- The photoelectric effect is a non-inertial process: the photocurrent appears almost instantly after the cathode exposure to light.

Classical undulatory theory of light was unable to explain such properties of the studied phenomenon. According to the classical concepts, the kinetic energy of an electron had to depend on the amplitude of light impinging on it. Moreover, an electromagnetic (light) wave causing atom excitation would need a certain amount of time to 'pump up' the electrons circulating around the atom nucleus into a higher energy state where the separation of an electron from the atom's nucleus to which it was bound is possible. The experimental evidence was contradicting the accepted theories. The 'discrepancy' was finally solved by Einstein (Einstein, 1905b) who combined Planck's quantum theory with the corpuscular theory of light according to which the light represents a beam of particles – photons, a term coined by the American physical chemist Gilbert N. Lewis in 1926 (Lewis, 1926) – moving with the speed of light and yielding a certain amount of energy proportional to the frequency of radiation. Arthur Compton, a famous American physicist, confirmed in his experiments not only the existence of photons but also the corpuscular-wave dualism of the elementary particles (Compton, 1923). The electrons of an atom not only occupy discrete energy levels, but also absorb the impinging photon energy in a discrete manner. If the energy of an impinging photon results are sufficiently high, then an electron can be ejected from the atom leaving behind a vacant position in the atomic orbital, as can be observed in Fig. 1.6. The created vacancy can be later occupied by another electron in a process followed by a photon emission.

The mathematical description of the so defined external photoelectric effect is expressed by Eq. [1.5], where T is the kinetic energy of the ejected electron (i.e. $\dfrac{mv^2}{2}$), h the Planck constant, v is the frequency of the incoming
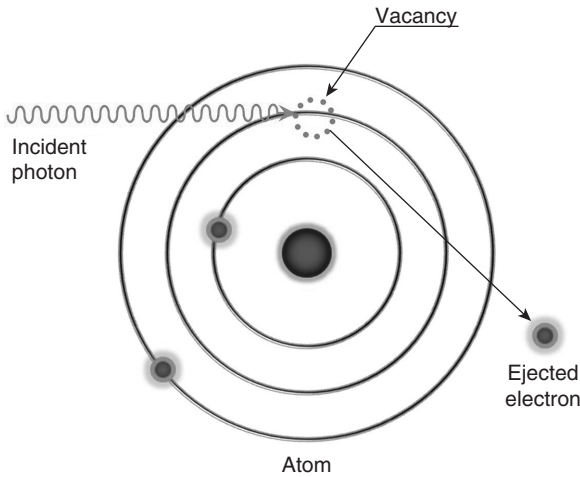
1.6 A graphic representation of the photoelectric effect based on the Rutherford–Bohr model of the atom.

Table 1.2 Values of the work function in eV for several semiconductor materials (based on data from Hölzl and Schulte (1979))

Material	Work function (eV)
Indium (In)	4.08
Gallium (Ga)	4.35
Silicon (Si)	4.95
Germanium (Ge)	5.15

light, and W, the so called work function of a certain material is the minimum energy required to liberate a single electron to a position outside the solid surface.

$$T = h\nu - W \qquad [1.5]$$

The work function W determines the threshold frequency of the external photoelectric effect as shown in Eq. [1.6], where ν_{min} is the minimal threshold frequency at which the photoelectric effect can still occur. This discovery earned Albert Einstein the Noble Prize in Physics in 1921.

$$W = h\nu_{min} \qquad [1.6]$$

The values of the work function for several commonly used semiconductor materials, namely indium, gallium, silicon and germanium are shown in Table 1.2. On the other hand, Fig. 1.7 depicts the dependence of the kinetic energy of an electron on the frequency of the impinging radiation interacting
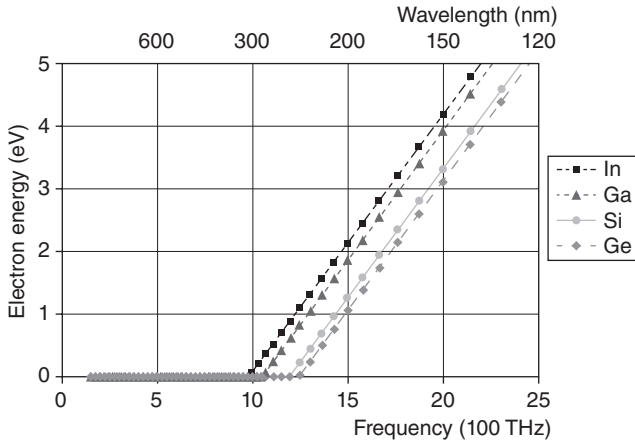
1.7 The energy required to excite an electron out of the solid material due to the photoelectric effect in indium, gallium, silicon and germanium vs. incoming radiation frequency.

with it. One can clearly see that for the materials chosen no photoelectric effect can take place for frequencies of impinging radiation below 1,000 THz (or wavelengths longer than 300 nm).

The electron energy depicted on the y axis in Fig. 1.7 is measured in electron-volts (eV), a unit defined as the amount of energy a fundamental charge (q) gains when passing through an electrical field induced through an electric potential difference of 1 volt. This unit is widely used in particle physics due to its small value compared to the unit of the joule defined in SI, where $1 \text{ eV} \approx 1.6 \times 10^{-19}$ J.

According to the already introduced Rutherford–Bohr model, the individual electrons bound to the nucleus of a certain atom, e.g. silicon, can only possess discrete energy levels separated by forbidden energy gaps. But, if they happen to acquire enough energy to 'escape' the forces that bind them to the atom nucleus, e.g. by absorbing incident radiation on them – which gives birth to the internal photoelectric effect – their behaviour will be defined by a new continuum of corresponding energy levels. According to the theory developed by Erwin Schrödinger, which describes how the wave function of a physical system evolves over time, these electrons will become quasi-free and will be able to move around in the silicon crystal within a complex electric field formed by the ions of the crystalline lattice and the valence electrons of the neighbouring atoms. The minimal required energy for the described internal photoelectric effect to take place, defined as the energy-gap energy E_g, is less than the work-function energy necessary for the electrons to completely leave the solid, required for the external photoelectric effect. An electron deficiency left in the covalent bond of the

silicon crystal may be filled by one of the neighbouring electrons creating a shift of the deficiency location. This movement of the deficiency locations, and the fact that it is much easier to describe it in these terms than to try to describe the movement of an entire system of quasi-free electrons, gave birth to a fictitious particle – a hole. It is thus reasonable to consider the birth of an electron-hole pair (EHP) every time an electron gets excited into the continuum of energy levels forming a conduction band. The flux of such quasi-free electrons across a crystal forms the electron current flow or electrical current I, defined as the temporal charge change, measured in coulombs per second, or amperes.

1.5   Early developments in photodetector technology

The internal photoelectric effect manifests itself as a change of the electrical properties of a metal or a semiconductor due to the increase in the amount of excited electrons (i.e. electrons that moved from the valence band into the conduction band of a semiconductor) caused by absorption of the impinging radiation. The measure of change of these electrical properties delivers indirectly the information regarding the energy of the impinging radiation itself. If a two-dimensional array of photosensitive elements is illuminated, and the information from every element is properly mapped on a two-dimensional surface, an image will appear that can be processed by our brain in the same way it happens with the mosaic image described in the beginning of this chapter. This is the basis of solid-state photosensing or imaging and is illustrated in Fig. 1.8.
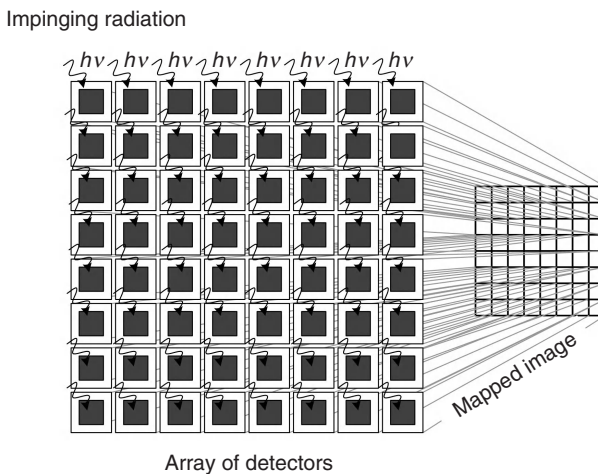
1.8 A graphic representation of the solid-state imaging principle using an array of photodetectors.

One of the first photodetectors invented was the thermal radiation receiver that reacted to the change in the temperature of sensitive elements. This principle was mainly conceived by Thomas Alva Edison in his so called 'tasimeter' (Edison, 1879), made possible after the discovery of the temperature dependence of the electrical resistance of certain materials such as the compressed carbon used by him. The invention was then further improved by S. P. Langley in 1888 (Langley, 1881) who invented the bolometer, an instrument that used a blackened platinum strip for the same purpose. This principle is still used in modern astronomy and other applications (e.g. night-vision), mainly based on semiconductor materials such as amorphous silicon in uncooled instruments or gallium doped germanium (Ge:Ga), as used by Frank J. Low (Low, 1961) in the 1960s, who is considered the father of astronomical bolometers.

After the development of the first vacuum tubes, physicists started to use them extensively in experiments. They helped Thomson to discover the electron, and invent the glow lamp. The first digital computers and TV receivers were based on vacuum tubes. In modern physics and engineering vacuum tubes are used as detectors for different types of radiation. The sketch of such a device used for light detection, called the photomultiplier tube (PMT), is shown in Fig. 1.9. This device was conceived aiming at detecting ultra low levels of radiation. In a PMT the light absorbed by the photocathode, for which normally alkaline (or multi alkaline) metal containing materials such as Ag-O-Cs, InGaAs:Cs, Na_2KSb or K_2CsSb (Hamamatsu, 2007) are used, results in the emission of electrons caused by the photoelectrical effect.
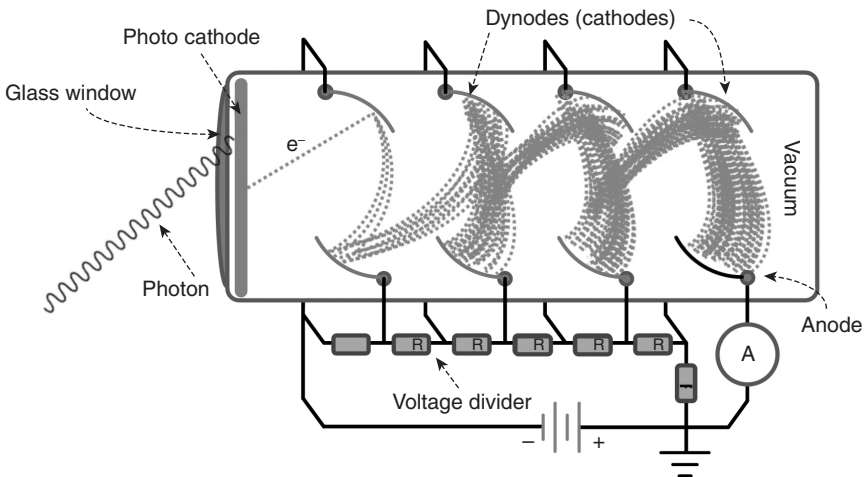
1.9 A schematic representation of the PMT.

Once generated, the electrons travel in vacuum from the photocathode in the direction of an anode placed on the opposite side of the vacuum tube due to an induced potential difference between the two electrodes. Within the vacuum tube, the electrons undergo avalanche processes at so called dynodes (see Fig. 1.9), biased at consecutively increasing high voltages, being consecutively multiplied in a process called secondary emission on their way to the anode. Properly chosen dynode materials such as BeO or MgO (Hamamatsu, 2007) usually allow for multiplication factors of more than one. Taking into account that modern PMTs consist of five or seven dynodes, each providing a multiplication factor of 10, they normally deliver a total of 10^5–10^7 electrons per single incident photon. Depending on the photocathode and window materials, the PMTs are sensitive to different wavelengths, i.e. different photon energies. Therefore a proper combination of the window and photocathode materials is a requirement for their optimal design. For example, for wavelengths below 200 nm, a Cs-I photocathode together with a sapphire window is normally used. For wavelengths above 300 nm, a more suitable GaAs photocathode combined with a borosilicate window should be used (Hamamatsu, 2007). These devices operate at several thousands of volts to achieve this performance, which makes the electromechanical complexity of the PMTs in addition to a somewhat poor spacial resolution of several mm^2 something hard to oversee. Nevertheless, the temporal resolution in the range of nanoseconds and the extremely low radiation sensitivity make them the instrument of choice in many applications such as, for example, optical emission spectroscopy or positron emission tomography (PET) among many others.

A step forward in the direction of minimization of PMT complexity and the further boost of their performance, especially where spatial resolution is concerned, was the development of the microchannel plate (MCP) detector. It consists of a lead glass carrier substrate populated by miniature periodically arranged photo-multiplying cells as shown in Fig. 1.10. A typical photo-multiplying cell within an MCP is a hollow cylinder with walls covered by a thin layer of Pb a few hundreds of nm thick, covered by an even thinner layer of silicon oxide, a few tens of nm thick. The first Pb resistive layer is used to provide electrical conductivity to the cell, whilst the second, dielectric layer, is used to cause the secondary electron emission. To the top and the bottom sides of the resistive layer a bias voltage is applied often reaching a few thousand volts (see Fig. 1.10).

Just like in PMTs, the MCPs are equipped with a photocathode where photon–electron conversion takes place. After an electron starts to move across the photo-multiplying cell, it hits the micro-channel wall. This is partially induced due to the small inclination of the MCP cell wall with respect to the photocathode, which is represented by the angle α in Fig. 1.10 (b). At the first hit the initial electron multiplication process takes place.
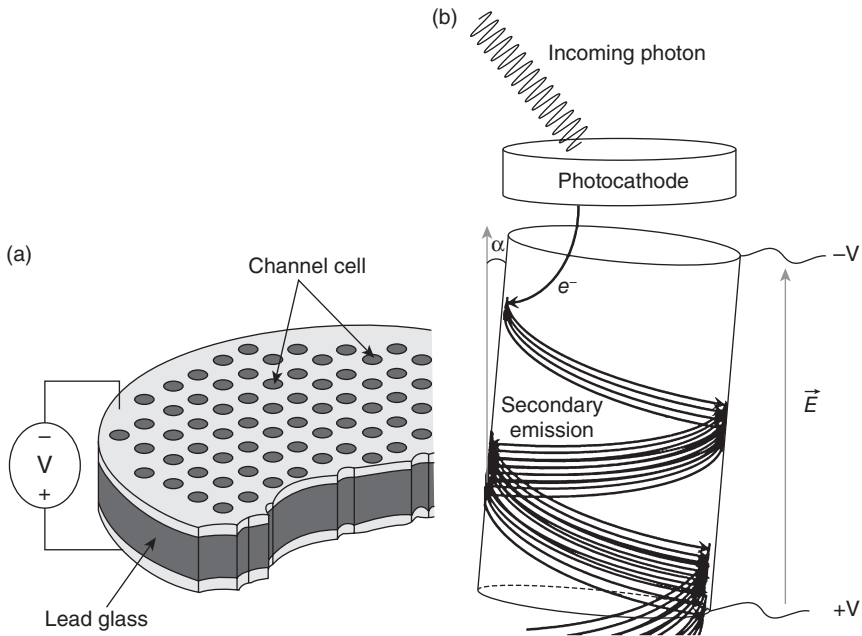
1.10 (a) A cutaway view of a MCP and (b) a cross section of the MPC cell causing the electron multiplication process due to secondary emission.

The newly created electrons continue to travel across the induced vertical electrical field (see Fig. 1.10(a)) and due to their lateral initial velocity they have a high probability of hitting the channel wall again on the opposite side of the cell, and creating more electrons in the following emission process. The walls of the micro channel cell act similarly to the PMT dynode system. At the end of this cascaded process, several thousands or hundreds of thousands of electrons emerging from the bottom side of the MCP enter the first chain of the readout electronics of the detector. Here, the MCP amplification factor depends on many parameters, e.g. the aspect ratio of the cell (the ratio of the depth of each micro channel related to its width) or the applied bias voltage. The MCPs provide similar electron amplification to the one provided by PMTs, but with a much higher spatial resolution. A typical micro-channel pitch of modern PMTs is around 15 µm. Moreover, they allow the building of portable systems capable of ultra low radiation level imaging, which enable their application in many different fields, most noticeably in those involved with night vision.

High biasing voltages and the expensive technology of MCPs nevertheless requires further optimizations. Very promising silicon based technology developed for integrated circuits since the 1960s proved to be the best

choice for this task with silicon as the material of choice. This new generation of photosensors is discussed in the next chapter.

1.6   References

Becquerel A. (1839) 'Mémoire sur les effets électriques produits sous l'influence des rayons solaires', *Comptes Rendus des Séances Hebdomadaires*, 9, 561–567.

Bohr N. (1913) 'I. On the constitution of atoms and molecules', *Philosophical Magazine Series 6*, 26(151), 1–25.

Born M. (1971) *The Born–Einstein Letters: Correspondence between Albert Einstein and Max and Hedwig Born from 1916 to 1955 with Commentaries by Max Born*, Walker & Company, New York.

Born M., Heisenberg W. and Jordan P. (1925) 'Zur Quantenmechanik II', *Zeitschrift für Physik*, 35, 557–615.

CIE (1983), *The Basis of Physical Photometry*, 2nd edn, Commission International de L'Eclairage Publ. No. 18.2, Central Bureau of the CIE, Vienna.

Cline D., Hofstetter H. W. and Griffin J. R. (1997) *Dictionary of Visual Science*, 4th edn. Butterworth-Heinemann, Boston.

Compton A. H. (1923) 'A quantum theory of the scattering of X-rays by light elements', *Phys. Rev*, 21(5), 483–502.

Curcio C. A., Sloan K. R., Kalina R. E. and Hendrickson A. E. (1990) 'Human photoreceptor topography', *The Journal of Comparative Neurology*, 292, 497–523.

Davson H. (1976) *The eye: Visual Function in Man*, Vol. 2A, 2nd edn, p. 616, Academic Press, New York.

de Broglie L. (1924) 'A tentative theory of light quanta', *Phil. Mag.* 47, 446.

Demtröder W. (2010) *Experimentalphysik 3: Atome, Moleküle und Festkörper*, Springer-Verlag, Berlin Heidelberg.

Dirac P. (1928) 'The quantum theory of the electron', *Proceedings of the Royal Society of London*, Series A, 117(778), 610–624.

Edison T. A. (1879) 'On the use of the tasimeter for measuring the heat of the stars and of the sun's corona', *Proceedings of the American Association for the Advancement of Science*, pp. 109–112.

Einstein A. (1905a) 'Zur Elektrodynamik bewegter Körper', *Annalen der Physik* 322(10), 891–921.

Einstein A. (1905b) 'Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt', *Annalen der Physik* 17(6), 132–148.

Einstein A. (1918) 'Dialog über Einwände gegen die Relativitätstheorie', *Naturwissenschaften*, 6, 697–702.

Hamamatsu (2007) *Photomultiplier Tubes. Basics and Applications*, 3rd edn, Hamamatsu Photonics, Hamamatsu (http://sales.hamamatsu.com/assets/pdf/catsandguides/PMT_handbook_v3aE.pdf).

Helmholtz H. (1924) *The Sensation of Vision: Treatise on Physiological Optics*, Vol. 2 (transl. J.P.C. Southall), Optical Society of America, USA.

Hölzl J. and Schulte F. K. (1979) 'Work function of metals' in *Solid Surface Physics, Springer Tracts in Modern Physics*, Springer-Verlag, Berlin Heidelberg, New York 85, 1–50.

Kodak (2008) 'Kodak fundamental radiometry and photometry. CCD image sensors', Revision 2.0 MTD/PS-0891, 24 November.

Kuhn T. (1970) *The Structure of Scientific Revolutions*, University of Chicago Press, p. 150.

Langley S. P. (1881) *The Bolometer*, American Metrological Society, New York.

Lewis G. N. (1926) 'The conservation of photons', *Nature* 118(2981), 874–875.

Low F. J. (1961) 'Low-temperature germanium bolometer', *J. Opt. Soc. Am.*, 51, 1300–1304.

Mehra J. and Rechenberg H. (1982) *The Historical Development of Quantum Theory*, Springer-Verlag, New York, Chapter 1.

Michelson A. and Morley E. (1887) 'On the relative motion of the Earth and the luminiferous ether', *American Journal of Science*, 34(203), 333–345.

Østerberg G. A. (1935) 'Topography of the layer of rods and cones in the human retina', *Acta Ophthalmol.*, 13(6), 1–97.

Planck M. (1900) 'Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum', *Verhandlungen der Deutschen Physikalischen Gesellschaft*, Leipzig, pp. 237–245.

Rayleigh L. (1900) 'Remarks upon the Law of Complete Radiation', *Phil. Mag.*, 49, 539–540.

Ripps H. and Weale R. A. (1976a) 'The visual photoreceptors', in Davson H. (ed.), *The Eye: Visual function in man*, Vol. 2A, 2nd edn, Academic Press Inc., New York, pp. 101–129.

Ripps H. and Weale R. A. (1976b) 'Visual adaptation', in Davson H. (Ed.), *The Eye: Visual Function in Man*, Vol. 2A, 2nd edn, Academic Press Inc., New York, pp. 5–41.

Rosenthal E., Solomon R. J. and Johnson, C. (2004) *Waves vs. Photons: A New View of How we Detect Light and Color*, Creative Technologies LLC, USA.

Rutherford, E. (1911) 'The scattering of α and β particles by matter and the structure of the atom', *Philosophical Magazine*, 21(6), 669–688.

Schrödinger E. (1926) 'An undulatory theory of the mechanics of atoms and molecules'. *Physical Review* 28(6), 1049–1070.

Schultze M. (1866) 'Zur Anatomie und Physiologie der Retina', *Arch. Mickrosc. Anat.*, 2165–286.

Stoletov M. A. (1888) 'On a kind of electrical current produced by ultra-violet rays', *Philosophical Magazine Series* 5, 26(160), 317.

Thomson J. J. (1897) 'Cathode rays', *Philosophical Magazine*, 44, 293.

Thomson W. (1901) 'Lord Kelvin, nineteenth-century clouds over the dynamical theory of heat and light', *Philosophical Magazine*, Sixth Series, 2, 1–40.

Thomson J. J. (1904) 'On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle with application of the results to the theory of atomic structure', *Philosophical Magazine*, Series 6, 7(39), 237.

2

Operational principles of silicon image sensors

D. DURINI, Forschungszentrum Jülich GmbH, Germany and formerly Fraunhofer Institute for Microelectronic Circuits and Systems, Germany and D. ARUTINOV, Fraunhofer Institute for Microelectronic Circuits and Systems, Germany

Abstract: This chapter discusses in detail the physics behind the different silicon based photodetector technologies. It analyses the challenges and perspectives of silicon based photosensors with a brief overview of hybrid and 3D sensor technology.

Key words: silicon based phototransduction, active pixel architectures, complementary metal-oxide-semiconductor (CMOS) photosensors, charge coupled device (CCD), silicon imaging, signal-to-noise ratio (SNR), hybrid and 3D detector technology.

'Technology feeds on itself. Technology makes more technology possible.'

Alvin Toffler

2.1 Introduction

Building on the foundation of the previous chapter, this chapter discusses the principles and technologies of silicon based photodetectors. It begins by discussing the properties of silicon and silicon based phototransduction. It goes on to review the MOS-C structure and p-n junction based photodetectors, and focuses afterwards on the development of charge coupled devices (CCD) and photosensors based on complementary metal-oxide-semiconductor (CMOS) technology. The chapter then discusses some of the technical challenges in areas such as high-performance pixel structures and their noise issues as well as miniaturization. It concludes by considering how hybrid and 3D technologies have been developed to address these challenges, and what technical hurdles still need to be addressed.

2.2 Silicon phototransduction

Silicon (Si^{14}) is a fourth group element of the periodic table that forms, if in solid-state, a diamond lattice crystal structure. In solid-state, it presents

25

covalent bonds of its four valence electrons oscillating around the nucleus in the $[Ne]3s^13p^2$ orbits (Sze, 2002). The wonderful thing about silicon is the fact that it features a high-quality interface to its own oxide, a high-quality dielectric, especially if it is oxidized under well controlled conditions, which made it the semiconductor of choice for the last 50 years of history of semiconductor based microelectronics.

In a silicon crystal formed by a huge amount of covalently bound atoms, the original identical electron energy levels start, as the once completely separated atoms start approaching each other, splitting into similar but nevertheless somewhat distinct energy levels forming energy bands. They must fulfil the 'Pauli Exclusion Principle' that states that no two identical 'fermions' (particles or quasiparticles obeying Fermi–Dirac statistics) can occupy the same quantum states simultaneously (McKelvey, 1982), so they adjust to the new circumstances. The most important energy bands that result from this process are the valence and the conduction bands, separated by an energy gap where no electron energy levels are allowed (McKelvey, 1982).

If crystalline silicon gets illuminated, in one of all scattering processes that take place between the impinging photons and the silicon crystal, there is an energy transfer between this impinging photon and the one scattered electron within the silicon crystal: the internal photoelectric effect takes place and one electron-hole pair (EHP) gets produced. Apart from the photoelectric effect, often being considered as the first in the line of photon-matter interaction phenomena, there exist at least two more important effects: the 'Compton scattering' and the 'pair production'. The probability, expressed as process cross-section in units of square centimetre per gram, of each of the three scattering effects to take place in silicon depending on the energy of the impinging radiation can be observed in Fig. 2.1 (based on the data extracted from Spieler, 1998).

The probability for the Compton scattering effect to take place starts increasing at around 500 eV or at approximately 2.25 nm wavelengths of impinging radiation (Spieler, 1998). It occupies a middle-energy region and represents a process of photon scattering (unlike absorption in photoelectrical effect) by a free electron. During this process the energy of the impinging photon is only partially transferred to the electron which results in change of photon energy and wavelength. This effect is widely used in medicine, e.g. in radiobiology and material sciences, used in gamma-ray spectroscopy, and is one of the most important proofs of the wave-particle duality. The phenomenon has been explained by Arthur Compton in his paper (Compton, 1923) which secured him a Nobel Prize in 1927. The inverse Compton scattering is another well-known process in which the energy gets transferred from the electron to the impinging photon.

The so called 'pair production' is the third effect that takes place during photon-electron interaction and occupies much higher energies than the
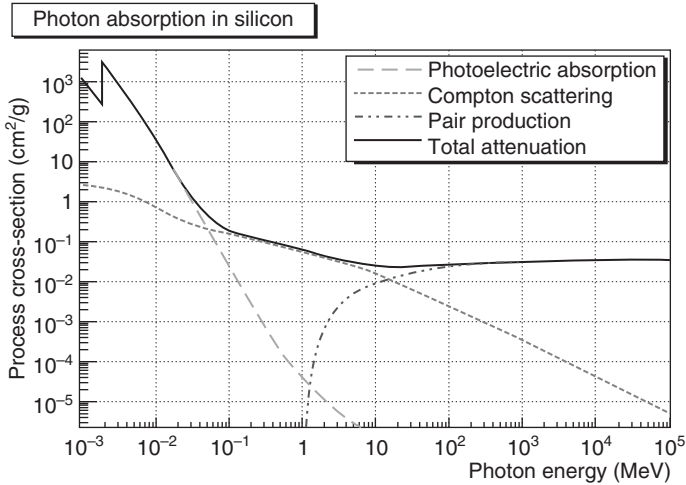
2.1 Process total cross sections as a function of energy of the impinging radiation in silicon, showing the contributions of the three most common processes: photoelectric absorption, Compton scattering and pair production (Spieler, 1998)

first two, starting at approximately 1 MeV of impinging radiation (or particle) energy, as can be observed in Fig. 2.1. If the impinging photon has an energy at least twice as high as the rest mass of an electron, during the interaction with the medium nucleus a 'particle-antiparticle' pair is generated, a process in which the whole energy of the impinging photon is spent. For example, if a photon with the energy of 1.022 MeV interacts with an atom, there is a certain probability that an electron-positron pair will be created. A positron, an antiparticle of the electron, has the same mass (0.511 MeV), the same spin, but the opposite signs for the electric charge and the lepton number. The first observation of the electron-positron pair production in a photon-nucleolus interaction has been made by the English physicist Patrick Blacket, for which he got the Nobel Prize in 1948. The probability of all three effects to take place, among others depends on the atomic number of the material where the interaction takes place, as well as on the energy of the impinging radiation. For the particular case of silicon being illuminated with the impinging radiation in the UV-NIR parts of the spectra (the wavelength range between 200 nm and 1127 nm), only photoelectric absorption is expected to occur.

If silicon is illuminated by a radiation with energies in the energy range between UV and NIR and the energy of the impinging photon E_{ph} (see Eq. [1.4]) is higher than the silicon energy gap $E_g = 1.11$ eV (measured at $T = 300$ K) between the maximum of the valence band and the minimum of the

conduction band (Kittel, 1996), band-to-band (or intrinsic) transition of an electron within the silicon grid takes place. This means that an impinging photon scatters an electron in the valence band forcing it to move into the conduction band and become free to move across the silicon crystal. If the photon energy results less than E_g, the photon energy will be absorbed only if there are available energy states in the forbidden bandgap of silicon due to chemical impurities or physical defects of the crystal grid, which are always present in any normally used extrinsic silicon (a doped silicon with altered electrical properties). The 'internal extrinsic transition' (Sze, 2002), as this process is known, represents the most important mechanism on which the entire silicon based phototransduction is based. Nevertheless, silicon is an indirect semiconductor, which means that the valence band energy maximum and the conduction band energy minimum do not share the same crystal momentum k. Actually, the E_g between these two bands in silicon at the same k, say the one present at the valence band maximum, results to be approximately 3.4 eV (Singh, 2003) – equivalent to a photon wavelength of 365 nm (see Fig. 2.2) in the UV part of the spectra – which would be the actual energy required for the intrinsic photoelectric effect to take place. If it was possible for this to occur only in this way, we would not be able to use silicon to detect visible radiation.

In order to understand why this is, nevertheless possible, we have to remember that for the extrinsic photoelectric effect to occur, in addition
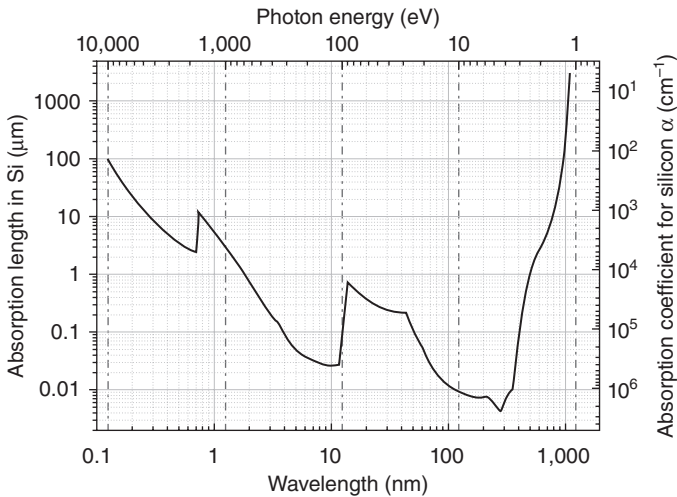
2.2 Wavelength (λ in nm) and photon energy (in eV) dependent impinging radiation absorption depth (in μm), obtained as the inverse of the absorption coefficient (α) in silicon, adapted from data taken from Theuwissen (1996) and Green and Keevers (1995).

to energy conservation, the momentum k of the electron-photon system must be conserved as well. The crystal momentum of the impinging photons k_{ph} is essentially zero compared to the electron k-values. Phonons, quantised units of lattice vibration within the silicon crystal in this case, have nevertheless a momentum similar to that of electrons. Thus, in silicon an indirect transition of the photon-scattered electron from the valence into the conduction band can take place for $1.1\,\text{eV} < E_{ph} < 3.4\,\text{eV}$ (the visible part of the spectra) only if a third particle, in this case a phonon, intervenes. This indirect optical transition occurring in silicon (where $E_{ph} < E_g$) occurs in two steps. At first, the electron absorbs the energy of an impinging photon and makes a transition to an intermediate state within the bandgap of an extrinsic silicon crystal. Simultaneously, an additional phonon is generated within the crystalline grid by the same impinging photon. Next, the photon scattered electron absorbs the additionally required momentum from a phonon and completes the transition into the conduction band. It can be inferred that the combined probability of a photon on one side and a phonon on the other to be absorbed is lower than the probability of either event to occur separately. However, due to the additional degree of freedom introduced to the system by the phonon momentum, electron transitions into many more states are possible (Wolfe *et al.*, 1989). The latter dramatically increases the probability of these events and allows for silicon based phototransduction.

If a flux of photons Φ_{ph} (quantized as a number of photons per area per second) with $E_{ph} > E_g$ is impinging on a silicon crystal, the wavelength dependent probability of absorption of each individual photon in silicon is proportional to the thickness of the silicon crystal it is impinging upon. Basically, the number of photons $n_{ph}(z)$ absorbed within an incremental distance Δz within the silicon crystal is given by Eq. [2.1] (Sze, 2002), where α is a proportionality constant defined as the absorption coefficient.

$$n_{ph}(z) = -\alpha \cdot \Phi_{ph}(z) \cdot \Delta z \qquad [2.1]$$

Considering the boundary condition $\Phi_{ph}(z) = \Phi_{ph}$ at $z = 0$, the relation between the impinging photon flux and the photon flux still existing at a certain depth z inside the silicon crystal can be expressed as shown in Eq. [2.2] (Sze, 2002).

$$\Phi_{ph}(z) = \Phi_{ph} \cdot e^{-\alpha z} \qquad [2.2]$$

Considering the photon energy (see Eq. [1.4]) and Φ_{ph}, the impinging photon flux can be expressed as impinging radiant flux $\Phi(z)$ in units of watt for a given silicon illuminated area. Based on α, the so called light absorption depth is defined as the depth inside a silicon crystal at which the impinging $\Phi(z)$ reduces its value due to different scattering mechanisms to below $1/e = 0.37\,\text{W}$. The absorption length for wavelength and/or photon energy

dependent radiation impinging on silicon at room temperature, based on the combined data obtained from Theuwissen (1996) and Green and Keevers (1995) can be seen in Fig. 2.2.

If we observe the graph shown in Fig. 2.2, we see certain dramatic drops of the absorption length in silicon between certain very close impinging radiation energies, e.g. around 100 eV or 1.8 keV. If the impinging photon energy is in the order of magnitude situated between the energy required for the ionization of an electron in the ith orbital of a silicon atom and the energy level of an electron situated in the next outer $(i + 1)$th orbital, then the probability of ionization becomes rather low: the photon energy is still not high enough to ionize the electron in the ith orbital (closer to the atom nucleus) and the binding energy of the electron orbiting in the $(i + 1)$th is too small for the photoelectric effect to take place in fulfilment of the momentum conservation law. If the impinging photon energy reaches the binding energy of the electron in the ith orbital, the probability (process cross-section) for the photoelectric effect to take place rises dramatically as well as its absorption coefficient.

For the case where $E_{ph} > E_g$, the same process will occur where the electron will transmit the 'extra' energy to a phonon and decrease its own energy to the level equivalent to the conduction band minimum. This will occur only up to energies of about 4 eV (Scholze et al., 2000) or 310 nm wavelengths (UV part of the spectra) of impinging photons. For photons above this energy, one electron is generated for every 4.4 eV up to 50 eV, and every 3.66 eV for photon energies above 50 eV (Scholze et al., 2000). Moreover, high energy X-rays ($E_{ph} > 10$ keV) have a low absorption cross-section (probability) in silicon, so that most photons will pass through the silicon lattice undetected (Wadsworth and McGrath, 1984). In order to assure more than 90% of probability of absorption of an impinging X-ray, e.g. for energies between $E_{ph} = 3$ keV and $E_{ph} = 15$ keV, the silicon depth should be 10 μm in the first case, and more than 12 mm in the second (Wadsworth and McGrath, 1984). Finally, for silicon, the wavelength value $\lambda = 1127$ nm, at which $E_{ph} < E_g$, represents the so called cut-off wavelength, at which no band-to-band transitions are possible and the silicon photosensitivity drops dramatically, can also be observed in Fig. 2.2.

Once the electron–hole pairs (EHP) are generated, the two charge carrier types have to be separated in order to avoid their almost immediate recombination. The length of time a minority carrier can 'survive' without recombining surrounded by majority carriers is the so called minority carriers' recombination time (or lifetime): τ_p for holes transiting an n-type region, or τ_n for electrons transition through a p-type region (Schroder, 1998). The distance these minority carriers are able to cross before their imminent recombination is called diffusion length: $L_n = \sqrt{D_n \tau_n}$ or

$L_p = \sqrt{D_p \tau_p}$, for D_n and D_p the charge carrier mobility (μ_n and μ_p) dependent diffusion lengths, respectively for electrons and holes (Schroder, 1998). Nevertheless, to avoid the recombination all together, a built-in or an externally induced electrical field is required where all the electrons will be attracted to the positive electrical pole and all the holes to the negative one. The straightforward solution for generating such an electrical field would be to use, citing two mostly used examples, a reverse biased *p-n* junction, first proposed by G. P. Weckler from Fairchild Semiconductor (Weckler, 1967) or a polysilicon-gate (metal-oxide-semiconductor capacitor (MOS-C)).

It is important to mention that 'excess' EHP are not only generated through the already explained internal extrinsic photoelectric effect, but also through $k_B T$ dependent thermal energy at temperatures above absolute zero that also creates band-to-band transitions: an additional electrical current normally called 'dark current'. These thermally generated carriers form part of the output current of the photodetector, reduce its capacity to store photogenerated charge carriers, and introduce an additional amount of noise (the so called 'dark shot noise'), due to the variance in the amount of thermally generated carriers under identical operating conditions, which follows, just as the variance in the amount of photons impinging the photoactive area does, the so called Poisson probability distribution. The rate of EHP thermal generation depends on the number of mechanical defects (crystalline grid dislocations) or chemical impurities (also known as Shockley–Read–Hall (SRH) generation-recombination centres) present in silicon, especially at the silicon surface or at the silicon-oxide interfaces. At the silicon surface, the potential periodicity of the crystalline grid no longer holds and exhibits an increased amount of so called 'dangling bonds', electrically active interface traps caused by missing silicon atoms and unpaired valence electrons.

The photodetector output current arises through photo and thermal generation of carriers according to the following main mechanisms that should be kept in mind when designing a photodetector in silicon:

- Thermal and photogeneration of EHP in the quasi-neutral areas in the silicon bulk within a diffusion length from the boundary of the region in which an electrical field has been induced (the space charge region – SCR). This mechanism is characterized in terms of a diffusion current subject to impurity concentration dependent recombination losses, which has a square function dependence with the distance.
- Thermal and photogeneration of EHP within the region in which the electrical field is induced (SCR), where carriers get swept out of this region constituting a drift current not subject by definition to any recombination losses.

- Primarily thermal generation of EHP out of fast surface states, which depends on silicon surface crystalline defects and is normally characterized in terms of the electron surface generation velocity s_{g-n}, given in seconds per centimetre (Schroder, 1998); the same mechanisms that can be found in the presence on any mechanically altered silicon region, e.g. the walls of shallow trench isolation (STI – a feature of the integrated circuit technology used for device isolation) structures or the silicon surface exposed to grinding or etching processes normally used for wafer thinning.
- Primarily thermal generation of EHP due to band-to-band and trap assisted tunnelling becomes very important in highly doped junctions due to very narrow regions in which normally high electrical fields are induced within the highly doped areas with relatively short tunnelling distances (Loukianova et al., 2003).
- Generation of EHP by impact ionization mechanisms (Brennan, 1999), where by increasing the reverse biasing voltage, electrons and holes acquire enough energy to, when colliding with the atoms of the crystal lattice, generate new electrons and holes in an avalanche process. The onset voltage for the reverse bias breakdown is called a 'breakdown voltage' V_{br}. An increase in reverse bias leads to a near exponential increase in the current until the multiplication rate increases to the point where the device becomes unstable and reaches the avalanche breakdown condition. This is still a fully reversible condition, but only for biasing voltages below the value required for the secondary breakdown which causes catastrophic failure.

One of the most important optical properties of silicon based photodetectors is their spectral responsivity, $\mathcal{R}$, which is a measure of the output electrical signal as a function of the impinging radiation energy, normally expressed in units of volts per joule per unit of area ($V/J/cm^2$) if the measured change in the electrical properties of the detector is to be expressed in volts and the impinging radiation energy is integrated over a defined time normally called 'charge collection time or photocurrent integration time T_{int}'. The spectral responsivity is normally characterized through the effective quantum efficiency η of the detector. η is the probability of an impinging photon to generate one EHP, added to the probability of the generated electrons to be separated from the holes and transported into a sense-node (SN) region without recombining, where the change in the electrical properties can be properly measured. For an optimal photodetector, the quantum efficiency should be as close as possible to 100%, and the amount of dark current and the statistical variation of the output signal (noise), should both be kept as low as possible. The signal noise floor defines the minimal radiation signal the detector can identify.

The mostly used silicon based photodetector structures and the way they deal with these issues will be described below.

2.3   Principles of charged coupled device (CCD) and complementary metal-oxide-semiconductor (CMOS) photosensing technologies

The entire silicon based imaging industry evolved around the concept of CCDs, first named charge 'bubble' devices, proposed by Willard Boyle and George E. Smith of the Bell Telephone Laboratories (Boyle and Smith, 1970) in October 1969 (Janesick, 2001) to be used as an analog shift register. The first publication reporting this invention appeared in *Bell Systems Technical Journal* in April 1970. In the same issue, a technical paper appeared in which Gilbert Amelio, Michael Tompsett and the same George E. Smith (Amelio *et al.*, 1970), experimentally verified the charge transfer principle between the CCD neighbouring cells and the feasibility of the entire concept. In May of the same year, Tompsett, Amelio and Smith proposed the first eight-bit CCD shift register consisting of a linear array of 26 closely spaced MOS-Cs with a *p-n* junction on either end, where a packet of charge was inserted into the first capacitor from one of the *p-n* junctions and then transferred down the array from one potential well into the next neighbouring one created by sequential pulsing of the electrode potentials (Tompsett *et al.*, 1970). Fabricated in a *p*-MOS technology, their best fabricated CCD array yielded a charge transfer efficiency of greater than 99.9% per electrode for transfer times of $2\,\mu$s. They also performed and reported the first attempt to create a visual image using their device, and generated the first silicon based image of a word 'CCD' (Tompsett *et al.*, 1970), which originated an entire revolution and proved that the CCD technology might be used for photodetection applications. The invention and its application to photodetection tasks brought Willard Boyle and George E. Smith the Nobel Prize in 2009. The CCD technology will be thoroughly described in Chapter 3 of this book.

Nevertheless, as described in Fossum (1997), before CCD technology was embraced by the industry, in the 1960s there were numerous groups working on solid-state image sensors. The problem with most of these developments was, nevertheless, that they all had an output proportional to the instantaneous local incident light intensity and did not perform any additional integration of the optical signal (Fossum, 1997) yielding low sensitivity. In 1964, the 'scanistor', an array of *n-p-n* junctions addressed through a resistive network to produce an output pulse proportional to the local incident light intensity, was reported by IBM (Horton *et al.*, 1964). In 1966, Schuster and Strull from Westinghouse Electric Corporation reported a 50×50 element monolithic array of phototransistors (Schuster and Strull,

1966). Self-scanned silicon image detector arrays had been proposed already in 1968 by Noble working for Plessey Company in the UK (Noble, 1968), where several configurations based on surface and buried reverse biased *p-n* junctions used as photodetectors were discussed, as well as a charge integration amplifier required for readout. Chamberlain reported further improvements of the image sensors in 1969 (Chamberlain, 1969), and Fry, Noble and Rycroft explored the issue of evident so called fixed-pattern noise (FPN) present in the array output signals in 1970 (Fry *et al.*, 1970). The FPN issue has accompanied all the image sensor technologies based on independent readout paths for each pixel element, and was considered for many years to be the major problem of MOS or CMOS based image sensors. The CCD technology proposed in 1970 minimized this issue by using one single readout path for each column of pixels leaving the FPN visible only on a column level. This was one of the main reasons for its adoption over many other forms of solid-state imaging (Fossum, 1997).

In parallel to the development of the CCD technology, the entire CMOS based microelectronic industry was making huge advances in what the processing technology is concerned. In the early 1990s, huge efforts were first started to take advantage of this highly developed process technology to try to create highly functional single-chip image sensors where low cost, high yield, and the possibility of inclusion of in-pixel intelligence and on-chip signal processing, integrated timing and control electronics, analog-to-digital converters, and the full digital interface – an electronic camera-on-a-chip, as named by Eric Fossum of the JPL (Fossum, 1995) – was the driving factor. In parallel, an independent effort grew from the USA NASA's 'need for highly miniaturized, low-power, instrument imaging systems for next generation deep-space exploration spacecraft' (Fossum, 1997) for which the developments were led by the US Jet Propulsion Laboratory (JPL) with a subsequent transfer of the technology to AT&T Bell Laboratories, Eastman Kodak, National Semiconductors and several other major US companies, and the start-up of the company Photobit, a spin-off of JPL (Fossum, 1997) that eventually formed part of one of the nowadays worldwide leading companies in mobile image sensors: Aptina Imaging (see Chapter 7).

The main problem with CMOS imaging, at least in its first years, was that the commercially available CMOS processes were not developed for imaging tasks at all, which meant that only the available *p-n* junctions depicted in Fig. 2.3, where a technology cross-section of a typical CMOS process is shown, could be used for photodetection if the overlying metal layers were removed. From Fig. 2.3 it can be concluded that photodiodes could be formed by using an *n*-well on the *p*-type substrate based junction, designed to isolate the *p*-type metal-oxide-semiconductor field-effect transistors (MOSFETs) from the *n*-type ones. Another possibility is to use
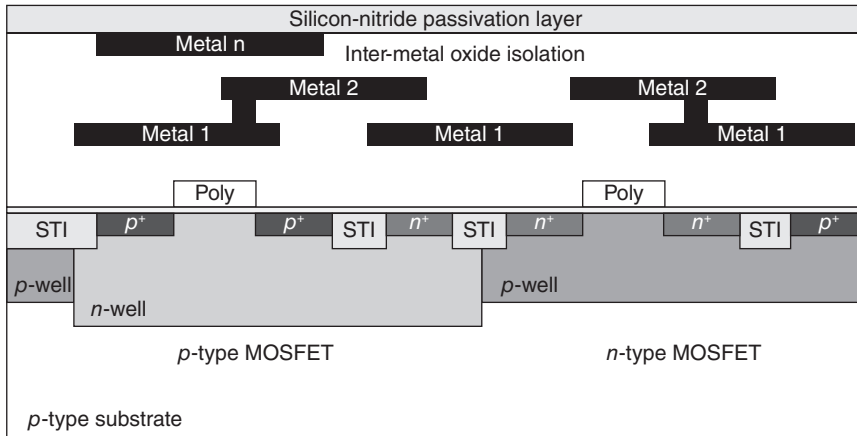
2.3 Technology cross-section of a typical CMOS process with *n* metal layers, one polysilicon layer, a *p*-type MOSFET (PMOS) and an *n*-type MOSFET (NMOS) with their respective *n*-well (n^+) and bulk (p^+) ohmic contacts.

the *Source* and *Drain p+* diffusions, used for the fabrication of *p*-type MOSFETs, fabricated on top of an *n*-well, or the equivalent *n+* diffusions on the *p*-substrate used in *n*-type MOSFETs. Considering photogate structures, an MOS-C gate on top of the *n*-well used for *p*-type MOSFETS could be used, or alternatively the equivalent MOS-C gate deposited on top of a *p*-type silicon used for *n*-type MOSFETs. Each of these basic photodetector structures has its advantages and disadvantages if compared with the other ones and none of them is really perfect.

The engineers designing the first CMOS imagers had to deal with a different imager development philosophy than the engineers working with CCD technology: in the world of CCD, the fabrication process was designed from the very beginning to enhance the CCD imaging performance, and at the end only NMOS or PMOS output amplifiers would get fabricated on the same substrate, whilst the readout circuitry is fabricated in a separate integrated circuit (IC); in the CMOS world, you have to use what is already there and try to make the most of it, putting the photodetectors and the entire readout circuitry on the same silicon substrate. The prediction made by Hon Sum Wong from IBM in 1996 (Wong, 1996), regarding CMOS technology and device scaling consideration, that CMOS imagers would benefit from further scaling after the 0.25 µm generation only in terms of increased fill-factor and/or increased signal processing functionality within a pixel, was made from this point of view. The increasing substrate doping concentration, always thinner gate-oxides, always lower maximal biasing voltages are all features that do not really benefit the photodetection

efficiency of CMOS image sensors. Nevertheless, at the end of the 1990s and during the 2000s several manufacturers changed this paradigm by further optimizing their processes to convert them into imaging enhanced ones. Instead of having CMOS processes that deliver highly functional logic circuitry with quite a bad front-end photosensing performance, nowadays we have photosensitivity enhanced processes still capable of delivering a quite acceptable CMOS functionality, and this made all the difference.

Just as used in the experiment carried out by Tompsett, Amelio and Smith, and also many others before them that proposed MOS photodetectors structures for their developments, transparent or semi-transparent gate metal-oxide-semiconductor capacitor based photodetectors have been successfully used in photometry and digital imaging ever since their early beginnings. Their basic principle of operation, just as proposed initially by the German physicist Julius Edgar Lilienfeld (Lilienfeld, 1928) in his attempt to create a voltage controlled current source, is that no DC current can flow under bias in this structure due to the presence of the insulating SiO_2 (gate-oxide) layer between the semi-transparent (and normally made of polycrystalline silicon, also called polysilicon) gate, as shown on the left-hand side part of Fig. 2.3. The operation of this device can be described as follows.

2.4   Metal-oxide-semiconductor-capacitor (MOS-C) structure-based photodetectors

In silicon based planar-technologies, n doped silicon regions are normally fabricated by ion implantation of phosphorus (P) or arsenic (As) atoms into a p-type – previously boron (B) doped – silicon substrate. Alternatively, p-regions are formed by ion implantation of boron atoms into an n-type – previously P or As doped – silicon. This is followed by a high-temperature impurity activation process in which the donor, e.g. P or As which form part of the 5th group of elements (and have five valence electrons), or acceptor impurities such as B which forms part of the 3rd group of elements (and has three valence electrons), that diffuse into the silicon substrate start forming part of the silicon crystalline grid by forming covalent bonds with the neighbouring silicon atoms. In order for this to occur, the boron atoms accept an 'extra' electron in order to form the required four covalent bonds with the neighbouring silicon atoms and get negatively ionized. They are called acceptors for this reason and form p-type semiconductor regions by increasing the total amount of positive holes in them. Phosphorus or arsenic atoms on the other hand lose an electron to form the four covalent bonds required by the surrounding silicon atoms getting positively ionized. For this they are called donors and form n-type semiconductor regions as they increase the total amount of negatively charged electrons in them. Both

donors and acceptors form the space charge due to their inability to move. They now form part of the extrinsic crystalline grid.

In addition to the process of properly doping the silicon substrate, thermally controlled high-quality SiO_2 based gate-oxide is grown on top of it trying to induce as few defects as possible at the silicon-oxide interface during this process. Once the thermally grown gate-oxide, with a typical thickness of less than 10 nm nowadays, has been fabricated, it gets covered by a polysilicon layer of some hundreds of nanometres in thickness, forming the MOS-C gate electrode.

Considering the energy band diagram of such a structure and taking into account the mostly positive mobile and fixed charge present in the gate-oxide, the flat band condition in which the conduction and valence bands have zero slope does not occur at zero bias applied to the polysilicon gate, but at a certain negative voltage U_{FB} if the MOS-C is fabricated on p-type silicon. Such a device can be observed in Fig. 2.4. The latter implies that the polysilicon to silicon work function difference $q(\phi_m - \phi_S)$, i.e. the difference between the Fermi level and the energy required for an electron to become free due to the external photoelectric effect and reach the so called 'vacuum' energy level, is not zero at zero applied bias, but is equivalent to a $U_{FB} < 0V$, and a certain amount of electrons is attracted through the positive charge to the silicon-oxide interface creating a somewhat weak inversion layer (see Fig. 2.4). Figure 2.4 shows a technology cross-section of a MOS-C structure fabricated on a p-type silicon substrate on the left-hand side, and on the right-hand side, the somewhat simplified energy-band diagrams for the equilibrium and flat-band condition of this structure, showing the Fermi levels, the metal work function ϕ_m, the silicon work function ϕ_s, the electron affinity χ, defined as the energy obtained by moving an electron from the vacuum energy level to the bottom of the conduction band, and the barrier potential $q\psi_B$, that indicates the potential difference between the Fermi level of intrinsic silicon E_i and actual doping concentration dependent silicon Fermi level E_{F-S}.

Now, if a positive voltage $U_{bias} > 0V$ is applied to the gate, as shown in Fig. 2.4, the holes are pushed away from the interface and a negative SCR gets formed by the ionised acceptors. The bands bend downwards and the concentration of holes decreases at the silicon-oxide interface. Due to the exponential dependence on distance between the Fermi level and the conduction-band edge, the majority carrier concentration (holes) drops off over a very short distance to a value that is negligible compared to the silicon p-type doping concentration. This allows for the assumption of an abrupt change from the SCR to undepleted silicon. With this approximation, the depletion layer (SCR) surface charge density Q_B, the silicon surface electric field E_s and the electrostatic potential ψ_s, whose dependence of the SCR represents the bending of the bands at the

U_{bias}

$+$

$-$

z

x (μm)

Gate-oxide

Polysilicon gate

STI

STI

Weak inversion layer
created under $U_{bias} = 0$ V

p-substrate

z (μm)

z'

Simplified energy-band diagrams
for a cut across the z-z' axis

Equilibrium
$U_{bias} = 0$V

Flat-band condition
$U_{bias} = U_{FB}$

Vacuum
level

$q\Phi_m$

E_{Fm}

qU_{bias}

E_{Fm}

qX

$q\Phi_S$

$q\psi_B$

E_c $E_i E_{F\text{-}S} E_V$

E_c $E_i E_{F\text{-}S} E_V$
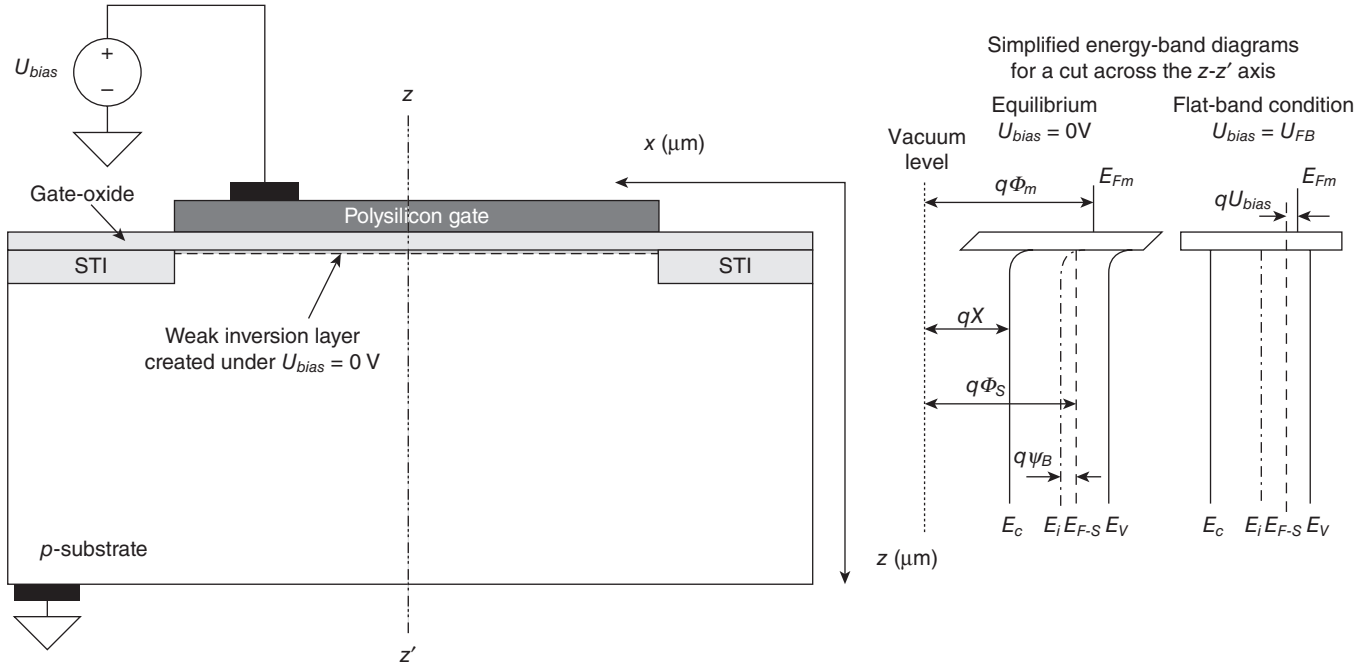
2.4 Technology cross-section and band diagram of a MOS-C structure fabricated on p-type silicon substrate for a non-ideal system in equilibrium, where $U_{bias} = 0$ V and $\Phi_m > \Phi_s$, and the flat-band condition where $U_{bias} = U_{FB}$ and $\Phi_m = \Phi_s$.

silicon-oxide interface, can be respectively, expressed through Eqs. [2.3], [2.4] and [2.5] (Sze, 2002), for q the fundamental charge, N_A the acceptor concentration density in silicon, W_{SCR} the SCR width, ε_0 the permittivity in vacuum and ε_{Si} the silicon dielectric constant.

$$Q_B = qN_AW_{SCR} \tag{2.3}$$

$$E_s = \frac{qN_AW_{SCR}}{\varepsilon_0\varepsilon_{Si}} \tag{2.4}$$

$$\psi_s = \frac{qN_AW_{SCR}^2}{2\varepsilon_0\varepsilon_{Si}} \tag{2.5}$$

The depth of the SCR can be expressed as given by Eq. [2.6], for electric field drop across the gate-oxide E_{ox}, scaled from E_{F-S} by the ratio of the silicon and oxide dielectric constants, where ε_{ox} is the oxide dielectric constant and d_{ox} the oxide thickness.

$$W_{SCR} = \sqrt{\frac{\varepsilon_0\varepsilon_{Si}}{qN_A}(U_{bias}+U_{FB})+\left(\frac{\varepsilon_{Si}}{\varepsilon_{ox}}d_{ox}\right)^2} - \frac{\varepsilon_{Si}}{\varepsilon_{ox}}d_{ox} \tag{2.6}$$

This condition is called 'depletion'. The structure is out of equilibrium at this point. The thermally and photogenerated EHPs are separated by the electric field created within the SCR. The positive charge (applied bias voltage) on the polysilicon gate at first gets neutralized only by the negative acceptor charge in the created SCR as all the majority holes are 'pushed' away from the silicon-oxide interface. With time, the minority electrons will diffuse into the SCR or get directly thermally or photogenerated within this area and transferred to the silicon-oxide interface decreasing the amount of acceptors required to achieve the overall electrical neutrality, i.e. decreasing W_{SCR}. In this way, the SCR and the equivalent electrostatic potential built within this region can be understood as the potential well in which the electrons can be collected once they are separated from the holes due to the existent electric field across the SCR, i.e. a collection bucket in which the minority electrons are collected that can store only as many generated electrons as required to establish the electrical neutrality and satisfy all the positive charge induced at the polysilicon gate through U_{bias}. In Fig. 2.5 minority carriers are photogenerated by incident photons that pass through the semi-transparent polysilicon gate electrode, and collected within the potential well (SCR) formed at the silicon-oxide interface. The amount of photogenerated minority carriers collected within an individual MOS-C is proportional to the input light intensity impinging on to the available (no metal covered) photoactive area of the MOS-C. Therefore, a two dimensional grid of pixels can record a complete image. Depending on whether the MOS-C was fabricated on n- or p-type substrate, it is defined as the p-type or the n-type MOS-C, respectively.
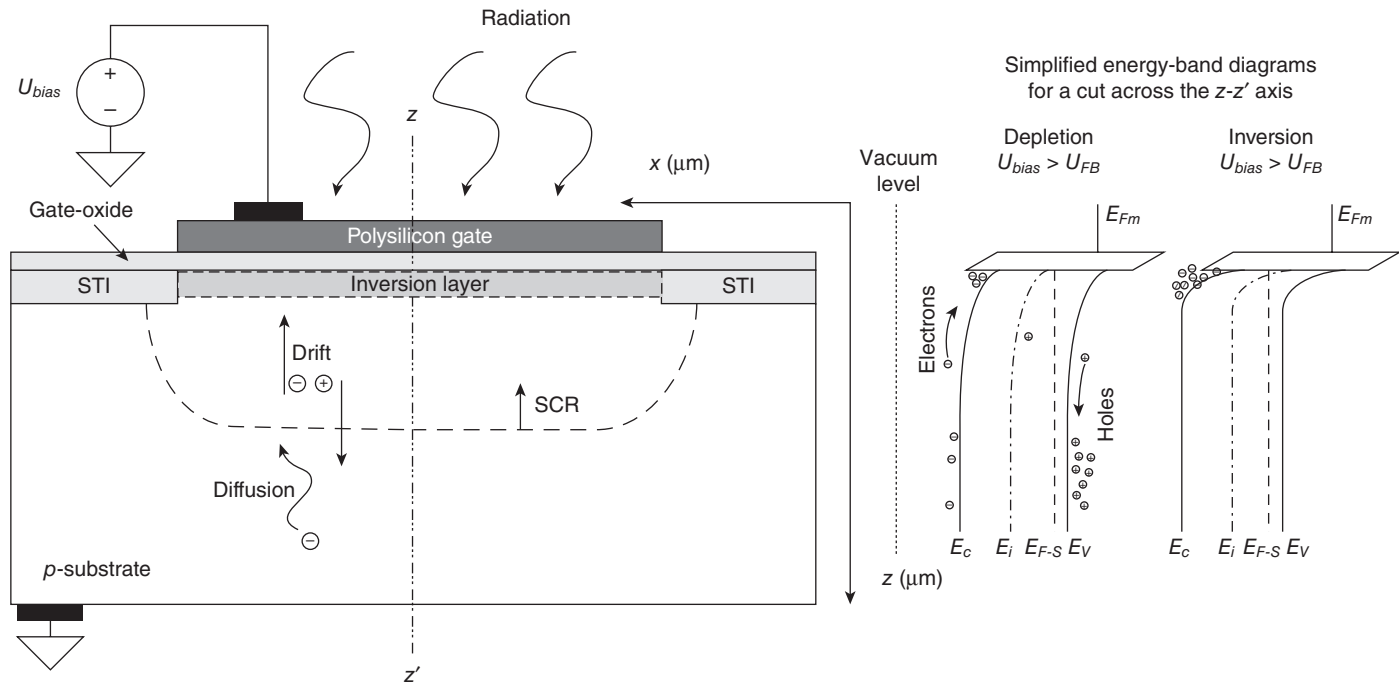
2.5 Technology cross-section and band diagram of a MOS-C structure fabricated on *p*-type silicon substrate for a non-ideal system under illumination in deep-depletion and inversion.

During this non-equilibrium process the intrinsic level at the interface reaches and eventually crosses through the non-equilibrium state created quasi Fermi level in silicon, i.e. $\psi_s \leq \psi_B$. Then the situation is that of the majority electrons at the interface, as shown in Fig. 2.4. This negative charge layer built there is named 'inversion' layer, and the status of the system is called 'inversion'. This so called strong inversion condition is accomplished when the surface potential has moved by approximately twice the distance of the Fermi level from the intrinsic level in the undepleted bulk, i.e. $\psi_s = 2\psi_B + 4k_B T/q$. The system has reached a new equilibrium condition, and under illumination, the amount of collected charge in this inversion layer corresponds to the amount of photogenerated and thermally generated charge during a certain defined charge collection or photocurrent integration time T_{int}. In inversion, the width of the inversion region W_{inv} is expressed by Eq. [2.7], and the resulting electric field at the surface of the semiconductor by Eq. [2.8] (Brennan, 1999). The minimum U_{bias} required for establishing the strong inversion condition is called the threshold voltage, defined by Eq. [2.9] (Brennan, 1999).

$$W_{inv} = \sqrt{\frac{4\varepsilon_0 \varepsilon_{Si} \psi_s}{q N_A}} \qquad [2.7]$$

$$E_s = \frac{q N_A}{\varepsilon_0 \varepsilon_{Si}} W_{inv} \qquad [2.8]$$

$$U_{th} = U_{FB} + 2\psi_B + E_{ox} d_{ox} = U_{FB} + 2\psi_B + \frac{d_{ox}}{\varepsilon_0 \varepsilon_{ox}} \sqrt{4 q N_A \varepsilon_0 \varepsilon_{Si} \psi_B} \qquad [2.9]$$

The MOS-C based photodetectors have several drawbacks. At first, the semi-transparent gate – usually made of polycrystalline silicon (polysilicon) – absorbs quite heavily in the UV and blue parts of the spectra (see Fig. 2.9(b) for wavelength dependent polysilicon gate transmittance curve) diminishing the amount of photons actually reaching the silicon substrate. Second, all the carriers generated through thermal or photon energies are collected at the silicon-oxide interface, full of SRH generation-recombination centres, which drastically affects the noise performance of the photodetector output signal. Although used in this way in the very beginning in the CCD technology, this kind of photodetector proved to be not very efficient, also having huge problems in what charge-transfer efficiency and noise performance are concerned. To solve this, 'buried' CCDs were introduced, where for example an additional p-type layer is fabricated on the n-type substrate, with a doping concentration maximum pushed away from the silicon-oxide interface, as proposed in Walden et al. (1970). The same idea could be applied to the structure shown in Fig. 2.5, by incorporating an n-type layer on top of the p-type silicon substrate. As explained in (Janesick, 2001), the first buried CCDs appeared commercially available in 1974,

produced by Fairchild, and tested by JPL. They yielded a noise floor of approximately $30\,e^-_{rms}$ and a charge-transfer efficiency (CTE) of 99.99% per pixel transfer (Janesick, 2001) in comparison to only 98% CTE and much higher noise floors for surface channel CCDs fabricated by the same manufacturer. The MOS-C based photodetectors found their rebirth in the 1990s with the appearance of CMOS based image sensors, named photogates due in the first place to the possibility to fabricate them in standard CMOS technology without any extra fabrication steps.

The *p-n* junction based photodetectors proved nevertheless much more efficient than their MOS-C based counterparts from the very beginning, and were incorporated into the so called interlined CCD technology very early and continued to be used in all existing applications until today, as will be explained further below.

2.5   *p-n* junction-based photodetectors

The *p-n* junction is formed due to the diffusion and activation of impurity atoms during the annealing steps of the fabrication process. Once this fabrication process is over, large concentration gradients at the formed junction cause diffusion of charge carriers. Holes from the *p*-side diffuse into the *n*-side, and electrons from the *n*-side diffuse into the *p*-side. As mobile holes continue to leave the *p*-side, some of the acceptors near the junction are left uncompensated due to their inability to move. Similarly, some of the donors are left uncompensated once the mobile electrons start leaving the *n*-side of the junction. This causes a negative space charge (acceptors) to build up at the *p*-side of the junction and the positive space charge (donors) to build up on the opposite side. This SCR creates an electric field that is directed from the positive charge toward the negative charge. The maximum electrical field E_{max} will arise at the *p-n* junction (at z_0 along the *z*-axis in the upper left part of Fig. 2.6) and will decay in the direction of the *n* and *p* regions (where the SCR in the *n*-region, as shown in Fig. 2.6, ends at $z = z_n$ and in the *p*-region at $z = z_p$) according to Eq. [2.10] (Sze, 2002), where φ is the induced electrostatic potential.

$$E(z) = -\frac{d\varphi}{dz} = -\frac{qN_A(z+z_p)}{\varepsilon_0\varepsilon_{Si}}, \quad for \quad -z_p \le z \le 0$$

$$E(z) = \frac{d\varphi}{dz} = \frac{qN_D(z-z_n)}{\varepsilon_0\varepsilon_{Si}}, \quad for \quad 0 \le z \le z_n$$

[2.10]

At this point, an electric current starts to flow opposite to the direction of the original carrier diffusion currents, i.e. the generated electrical field causes a drift current of each carrier type to flow opposed to its diffusion
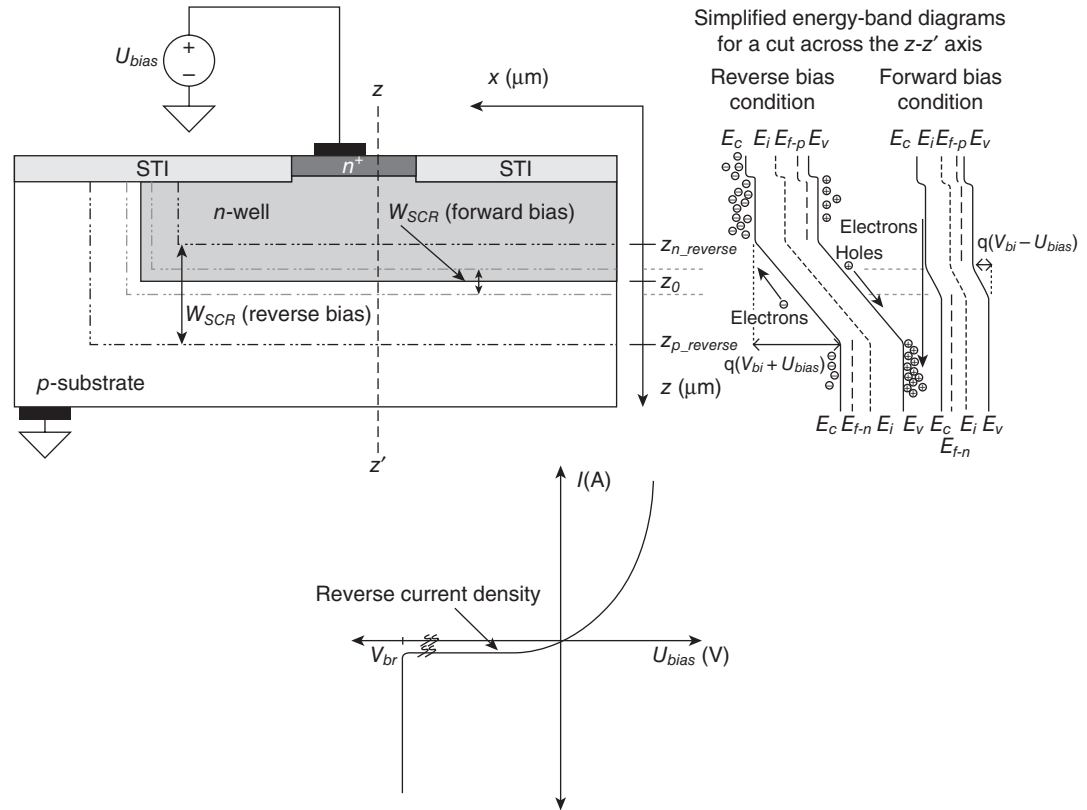
2.6 Typical p-n junction structure formed by an n-well and a p-type substrate as fabricated in CMOS planar technology based on STI: upper left – a 2D technology cross-section of the p-n junction; upper right – simplified energy band diagrams at the cut made across the z-z' axis on the right diagram, for the forward and reverse bias conditions, respectively; bottom – the ideal current-voltage characteristic of a p-n junction.

current. The overall electrical neutrality of the semiconductor requires that the total negative space charge in the *p*-side precisely equals the total positive space charge in the *n*-side. Derived from the first Maxwell equation or Gauss' law stating that the total electric flux exiting any volume is equal to the total charge inside, the unique space charge distribution and the electrostatic potential φ are given by Eq. [2.11], the so called Poisson equation for *p-n* junctions, where ρ_s is the space charge density given by the algebraic sum of charge carrier densities (*p* for holes and *n* for electrons) and the ionized impurity concentrations N_A and N_D (Sze, 2002). For all further considerations of the *p-n* junction in thermal equilibrium it will be assumed that all donors and acceptors are ionized (activated), and that in regions far away from the *p-n* junction charge neutrality is maintained and the total space charge density is zero. For the *p*-type neutral region it will be assumed that $N_D = 0$ and $p \gg n$, and that inside the *n*-well in Fig. 2.6, away from z_n, $N_A = 0$ and $n \gg p$.

$$\frac{d^2\varphi}{dz^2} \equiv -\frac{dE}{dz} = -\frac{\rho_s}{\varepsilon_0 \varepsilon_{Si}} = -\frac{q}{\varepsilon_0 \varepsilon_{Si}}(N_D - N_A + p - n) \qquad [2.11]$$

Under all these considerations, the doping concentrations of the *n*-type and *p*-type regions define the total width of the SCR W_{SCR} via Eq. [2.12], where $W_{SCR} = z_n + z_p$. For the case of thermal equilibrium, where $U_{bias} = 0\,\text{V}$, and the drift currents equal the diffusion ones.

$$W_{SCR} = z_n + z_p = \sqrt{\frac{2(V_{bi} + U_{bias}) \cdot \varepsilon_0 \cdot \varepsilon_{Si}}{q}\left(\frac{N_A + N_D}{N_A N_D}\right)} \qquad [2.12]$$

The total variation of the electrostatic potential across the junction, the so called built-in voltage V_{bi}, is also defined by the doping concentrations as expressed in Eq. [2.13] (Sze, 2002), where k_B is the Boltzmann constant, T temperature and n_i the intrinsic carrier concentration in silicon.

$$V_{bi} = \frac{k_B T}{q}\ln\left(\frac{N_A N_D}{n_i^2}\right) \qquad [2.13]$$

One way to alternate this thermal (dynamic) equilibrium of the *p-n* junction is to apply a bias voltage to it. The two dimensional technology cross-section of an *n*-well on *p*-substrate based *p-n* junction (upper left part) and its respective energy band diagrams under both operating conditions in thermal equilibrium (upper right part) are depicted in Fig. 2.6. In Fig. 2.6, E_c stands for the conduction band, E_v the valence band, E_i the intrinsic silicon Fermi level somewhere in the middle between E_c and E_v that for an intrinsic semiconductor governs via the Fermi–Dirac distribution the probability an electron has to occupy this energy level, and $E_{F\text{-}p}$ and $E_{F\text{-}n}$ the doping concentration dependent new Fermi levels for the *p*-type region

and *n*-type region, respectively. If the *p-n* junction is forward biased by an additional bias voltage U_{bias}, the electrostatic potential difference across the *p-n* junction (electrostatic potential barrier between the *n* and *p* regions) is reduced to $V_{bi} - U_{bias}$ enabling the electron flow across the junction (see the *I-V* diagram in the bottom part of Fig. 2.6). If the *p-n* junction is reverse biased by the same U_{bias}, the potential difference (electrostatic potential barrier) will increase by the same amount so that under thermal equilibrium only diffusion currents will flow across the junction.

As explained so far, if under illumination, a photocurrent and a dark current start flowing across the reverse biased *p-n* junction. If the structure depicted in Fig. 2.6 is taken into consideration, then the minority electrons in the *p*-type substrate will constitute the signal charge. They will be separated from the holes within the SCR and will eventually end up in the region of most positive electrostatic potential (the n^+ diffusion at the silicon surface in Fig. 2.7). All the assumptions made so far are no longer valid once the *p-n* junction is no longer in thermal equilibrium.

If the change in the electrical properties of the *p-n* junction (also called the photodiode PD) is to be measured, the junction must be initially set to a known electrostatic potential level, i.e. it has to be loaded to a certain defined initial (reset) voltage in such a way that the change of that initial voltage can be measured as generated signal and correlated to the properties of the radiation impinging on the photodetector. Once the PD is no longer electrically coupled to any fixed potentials, the change in the current flow as well as the change in the total amount of charge collected in it will cause the floating PD potential to change. The amount of additional charge that can be collected in the junction is defined by the SCR dependent initial junction capacitance C_{SN}, considered under U_{bias} bias and defined by Eq. [2.14] (Sze, 2002) – obtained by applying Eq. [2.12] – added to all the other parasitic capacitances present at the node due to metal layers, the reset MOSFET, etc. The potential variation at the PD which delivers the PD output voltage signal U_{out} is defined by the amount of thermally and photogenerated charge Q_{signal} added to the PD, as expressed in Eq. [2.15], defined having as the starting point U_{reset} and C_{SN}.

$$C_{SN} = \frac{\varepsilon_0 \varepsilon_{Si}}{W_{SCR}} = \sqrt{\frac{2(V_{bi} + U_{bias}) \cdot \varepsilon_0 \cdot \varepsilon_{Si}}{q} \left(\frac{N_A + N_D}{N_A N_D} \right)}$$

$$= \sqrt{\frac{q \cdot \varepsilon_0 \cdot \varepsilon_{Si}}{2(V_{bi} + U_{bias}) \left(\dfrac{N_A + N_D}{N_A N_D} \right)}}$$

[2.14]

$$U_{signal}(t) = U_{reset} - \frac{Q_{signal}(t)}{C_{SN}}$$

[2.15]
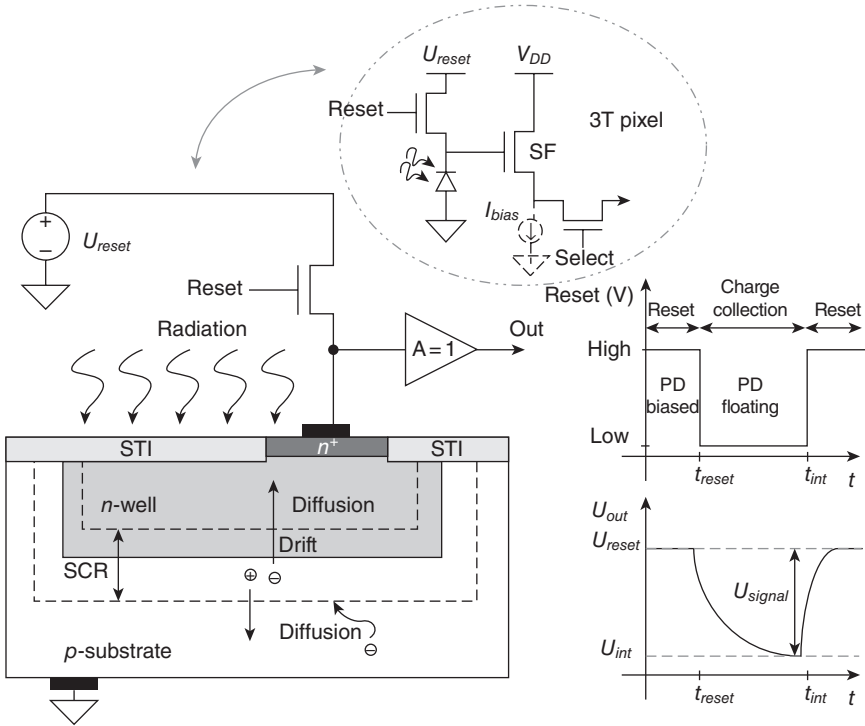
2.7 Schematic technology cross-section of a *p-n* junction based photodiode, showing the SCR and the drift and diffusion currents arising under illumination, and a reset transistor and the output signal buffer stage (bottom left); the reset signal time diagram and the photodetector output signal shapes (bottom right); in the upper part, the schematic representation of the equivalent typical 3T pixel used in CMOS imaging applications.

2.6 Noise considerations in pixel structures

Through cyclic reset operation of the PD, photosignals can be sensed in a periodic way enabling video streaming, just as proposed by Walter F. Kosonocky and James E. Carnes of RCA Laboratories (Kosonocky and Carnes, 1971), and inspired by the work performed by Noble (Noble, 1968), introducing what they called the 'digital signal regeneration stage' that they applied to a CCD based analog shift register to turn it into a digital circuit. As will be explained in more detail in Chapters 5 and 7, the most orthodox way to reset the *n*-well is by using a MOSFET. This allows for charging the PD to a defined U_{reset} during the reset operation and leaving the PD floating, i.e. electrically isolated from all other nodes during the

charge collection phase – which is accomplished by shutting the gate voltage of the reset MOSFET off – which originated the name 'floating diffusion' defined by Kosonocky and Carnes (1971). The technology cross-section of a *p-n* junction based photodiode under illumination to which a periodic reset signal can be applied through an *n*-type MOSFET, and the resulting output signal is sensed through a photodetector output buffer stage, normally used as an interface stage to readout electronics, is shown in Fig. 2.7. A source-follower in-pixel amplifier can be used in this configuration to buffer the pixel output signal and be able to transmit it over an entire column bus, just as first proposed by Noble (Noble, 1968), and an additional select switch (here also a MOSFET can be used) is incorporated into it to be able to select one particular pixel for readout if forming part of a matrix, just as proposed by Weckler in 1967 (Weckler, 1967) in his first proposed so called passive pixel structures (which at that time did not incorporate the in-pixel buffering stage). This particular architecture is known also as a three transistor active pixel (3T pixel: formed by the reset transistor, the source-follower transistor, and the select transistor in Fig. 2.7) and became one of the most-used pixels in CMOS based imaging applications (Theuwissen, 2008), demonstrated by JPL in 1995 in a 128 × 128 element array that had on-chip timing, control, correlated double-sampling and FPN suppression circuitry (Nixon *et al.*, 1995 as cited in Fossum, 1997).

One important issue existing in any photodetector circuit is the one related to its signal-to-noise ratio (SNR): a ratio between its desired output signal and the noise level existing in it, where noise represents all the unwanted signals that appear as a result of random processes that cannot be described by closed-form time-domain (i.e. stochastic) expressions. An acceptable way to represent any noise signal is to use its probability density function $p(x)$ which describes the probability density in terms of the input variable x (Leach, 1994). If the amount of available signal samples is assumed sufficiently large so that the result does not significantly change if this amount of samples is increased, and the measured entity is independent of time over the time interval of interest, the random processes existent are said to be stationary, and can be expressed as an ensemble average value $\bar{x}$ of the measured entity x.

Being this measured entity voltage or current, the variance σ^2 is the mean-square value of its AC component defined by Eq. [2.16] (Leach, 1994).

$$\sigma^2 = \overline{\left(x - \bar{x}\right)}^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 p(x)\,dx \qquad [2.16]$$

The square-root of the variance is called a 'standard deviation' of the signal and is often represented by the letter σ. For voltage or current signals, σ is

simply the square root of their mean squared (rms) AC components. Due to the absence of the DC component of the noise signals, the variance results equal to the rms values of the measured signals (Leach, 1994). Moreover, the spectral density of a random variable is defined as its mean-square value per unit bandwidth. For a noise voltage, it has the units of square volt per hertz (V^2/Hz). Equivalently for noise currents the respective units are A^2/Hz. For $v_n(t)$ being the zero-mean random voltage defined over a time interval $-T/2 \leq t \leq T/2$, assuming that $v_n(t) = 0$ outside this time interval, the mean-square value of $v_n(t)$ can be written as its average value in the defined time interval. On the other hand, if a frequency function $S_v(f)$ called 'spectral density' of $v_n(t)$ is considered and measured in units of V^2/Hz, then $Sv(f)\Delta f$ is interpreted as being the amount of noise voltage contained in the frequency band defined by f and $f + \Delta f$, where Δf is small. Both dependencies are expressed in Eq. [2.17] (Leach, 1994).

$$\overline{v_n^2} = \frac{1}{T} \int_{-T/2}^{T/2} v_n^2(t)\,dt$$

$$\overline{v_n^2} = \int_0^\infty S_v(f)\,df \tag{2.17}$$

The main types of fundamental noise mechanisms taking place in any electronic circuit are thermal noise, the already explained shot noise (following the Poisson distribution function), random telegraph signal (RTS) noise, and low-frequency ($1/f$) or flicker noise.

Thermal noise is caused by the random thermally excited vibration of the charge carriers in a conductor, observed first by John Bertrand Johnson (Johnson, 1928) of American Telephone and Telegraph Company (AT&T) in 1927 (since 1934 called Bell Telephone Laboratories), and theoretically analysed by Harry Nyquist (Nyquist, 1928) from the same company. Because of their work, thermal noise is also called Johnson or Nyquist noise. Based on their contribution, and regarding a complex impedance Z of a two terminal network, the mean-square thermal noise voltage $\overline{v_{n\_thermal}^2}$ generated by this impedance in the frequency band Δf is given by Eq. [2.16] (Nyquist, 1928). Because Z is a function of frequency, Δf must be small enough so that the impedance real part $\mathrm{Re}(Z)$ is approximately constant over the defined frequency band. Otherwise, the noise must be expressed by an integral. Following, the mean-square short-circuit thermal noise current $\overline{i_{n\_thermal}^2}$ in Δf is also defined in Eq. [2.18] (Leach, 1994). Thermal noise ultimately limits the resolution of any measurement system; even if the readout electronics could be built perfectly noise-free, the resistance of the signal source would still contribute noise to the signal output.

$$\overline{v_{n\_thermal}^2} = 4k_B T \, \mathrm{Re}(Z) \Delta f$$

$$\overline{i_{n\_thermal}^2} = \frac{4k_B T \Delta f}{\mathrm{Re}(Z)}$$

[2.18]

Flicker noise, on the other hand, is a noise with a spectral density proportional to $1/f^n$, where $n \approx 1$, which is the origin of its alternative name 'one-over-f-noise'. This type of noise seems to be a systematic effect inherent in electrical conduction, although no definite proof has been given about its origins. The $1/f$ spectral density of the flicker noise holds down to extremely low frequencies, and is normally modelled as a noise current flowing in parallel to the measuring device. In general, the mean-square flicker noise current in the frequency band Δf is defined in Eq. [2.19] (Leach, 1994), where $n = 1$, K_f is the flicker noise coefficient, and m is the flicker noise exponent. In the same equation the flicker noise spectral density $S_i(f)$ is also defined. Actually, if carrier trapping and detrapping mechanisms are modelled, especially in short-channel MOSFETs, then a Lorentzian spectrum (i.e. frequency dependence proportional to $1/f^2$) is obtained (Hosticka, 2007). If different trapping times are involved, then we obtain an envelope exhibiting $1/f$ behaviour.

$$\overline{i_{n\_flic\,ker}^2} = \frac{K_f I_{DC}^m \Delta f}{f^n}$$

$$S_i(f) = \frac{\overline{i_{n\_flic\,ker}^2}}{\Delta f} = \frac{K_f I_{DC}^m}{f^n}$$

[2.19]

The SNR is usually measured at the output of a circuit where the signal and noise voltages are larger (if an amplifier forms part of it, as depicted in Fig. 2.7) and easier to measure as given by Eq. [2.20] (Leach, 1994), specified in dB, where v_{so}^2 is the output signal voltage power and $\overline{v_{no}^2}$ is the mean-square noise output voltage.

$$SNR = 10 \log \left(\frac{v_{so}^2}{\overline{v_{no}^2}} \right)$$

[2.20]

Addressing the mentioned noise problems, the PD reset mechanisms give rise to two spurious noise components: the 'reset' and 'partition' noises, respectively, expressed through the mean-square reset noise voltage $\overline{v_{n\_reset}^2}$ in Fig. 2.8. Figure 2.8 shows the equivalent electrical model of all noise sources existing in the pixel configuration schematically depicted in Fig. 2.7. This kind of noise, closely related to 'thermal' noise, is important at all locations at which capacitors are charged and discharged through a circuit and appear due to the thermodynamic uncertainty of the charge on a capacitor.
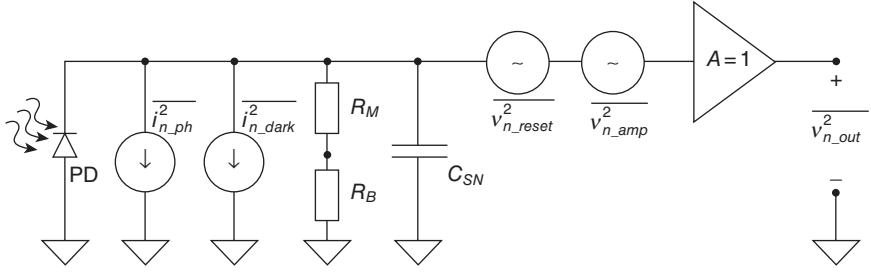
2.8 Schematic representation of the noise sources present in a PD 3T pixel depicted in Fig. 2.7.

In Fig. 2.8, the complex impedance Z, expressed by [2.21], consists of the pixel sense-node capacitance C_{SN}, formed by the *p-n* junction capacitance added to all other parasitic capacitances present in the pixel structure depicted in Fig. 2.7 (mainly the capacitances related to all metal connections added to the SF-gate and reset transistor capacitances), connected in parallel to R_S, the sum of the undepleted silicon bulk resistance R_B added to the resistance of all the joint metal connection layers R_M.

$$Z = R_S \left\| \left(\frac{1}{j\omega C_{SN}} \right) \right. = \frac{R_S}{1 + j\omega C_{SN} R_S} = \frac{R_S}{1 + (\omega R_S C_{SN})^2} - j \frac{\omega R_S^2 C_{SN}}{1 + (\omega R_S C_{SN})^2}$$

[2.21]

Considering Eqs [2.17] and [2.18] for the mean-square thermal noise voltage, both in time-domain, the first component, $\overline{v_{n\_kTC}^2}$, of the mean-square reset noise voltage $\overline{v_{n\_reset}^2}$ – considering $\omega = 2\pi f$ – is expressed through Eq. [2.22], where ξ is the so called 'reset constant' that takes the value of 0.5 for the case of 'soft' reset and 1 for 'hard' reset operations (Pain *et al.*, 2000). The precise expression in time domain should incorporate the term $1 - e^{-\frac{t}{R_B C_{SN}}}$ multiplying the result of Eq. [2.22].

$$\overline{v_{n\_kTC}^2} = \xi \int_0^\infty 4 k_B T \, \mathrm{Re}(Z) \, df = \xi \int_0^\infty \frac{4 k_B T R_M}{1 + (2\pi f R_M C_{SN})^2} \, df$$

$$= \xi \frac{2 k_B T}{\pi C_{SN}} \int_0^\infty \frac{dx}{1 + x^2} = \xi \cdot \frac{k_B T}{C_{SN}}$$

[2.22]

As it can be inferred from Eq. [2.22], $\overline{v_{n\_kTC}^2}$ is often referred as '$k_B T/C$ (*k-T-* over-*C*) noise'. In silicon based imaging it can be expressed also as the so called 'equivalent noise charge' (ENC): considering the $k_B T C_{SN}$

mean-square noise voltage $\overline{v_{n\_kTC}^2} = Q_n^2/C_{SN}^2$, then the ENC due to $k_B T/C$ can be expressed as $ENC_{k_B T/C} = \dfrac{\sqrt{\xi \cdot k_B T C_{SN}}}{q}$.

The 'partition noise', closely related to the $k_B TC$ noise used to be embedded in the reset noise model with a scaling factor α_{part} applied to the conventional thermal noise source (Lai and Nathan, 2005). The α_{part} is found to be inversely related to the fall time of applied gate pulse (in this case, the pulse applied to the 'reset' transistor in Fig. 2.7). It is believed that a fast fall time will trap residual charge in the channel after the transistor pinches off. The leftover charge is the primary source of partition noise. For standard used clocking signals provided normally by 'Flip-Flop' devices, rise and fall times present in the signal are of around 2 ns, for which value α_{part} = 1..3, and the mean-square partition noise voltage can be expressed as Eq. [2.23], where C_{G_RST} is the gate capacitance of the reset transistor in Fig. 2.7. Following the same development pursued to obtain Eq. [2.22], the 'partition' ENC is defined by Eq. [2.23].

$$\overline{v_{n\_par}^2} = \cdot \frac{2\alpha_{part}(k_B T C_{G\_RST})}{\pi^2 C_{SN}^2}$$

$$ENC_{part} = \frac{\sqrt{2\alpha_{part} \cdot k_B T C_{G\_RST}}}{q\pi}$$

[2.23]

Regarding Fig. 2.8 and its relation to the 3T pixel schematic diagram shown in Fig. 2.7, the second (source-follower–SF) transistor shown in the electrical diagram forms part, together with the bias current I_{bias} (normally substituted by a 'current-mirror'), of the already mentioned SF pixel buffer amplifier. This allows the PD output voltage U_{out}, related through a proportionality factor C_{SN} to the amount of integrated charge Q_{signal} collected in the n-well during T_{int}, to be observed without removing Q_{signal}, i.e. in a non-destructive manner. The variance of the pixel output signal caused by the SF buffer amplifier, defined here as the SF noise, is defined by Eq. [2.24] as the variance σ_{SF}^2 of the SF output signal or its ENC, where $\overline{v_{n\_SF}^2}$ is the input-referred mean-square source-follower noise voltage.

$$\sigma_{SF}^2 = \frac{C_{SN}^2 \cdot \overline{v_{n\_SF}^2}}{q^2}$$

$$ENC_{SF} = C_{SN} \cdot \frac{\sqrt{\overline{v_{n\_SF}^2}}}{q}$$

[2.24]

Taking into account the already mentioned dark current shot noise and its Poisson distribution, it can be stated that $\overline{n_{dark}}$, defined by Eq. [2.25] in terms of the PD dark current flowing during T_{int}, is not only the mean number of thermally generated carriers, but also its variance σ_{dark}^2.

Thus, the number of thermally generated electrons fluctuates about $\overline{n_{dark}}$ with a standard deviation of $\sigma_{dark} = \sqrt{\overline{n_{dark}}}$.

$$\sigma_{dark}^2 = \overline{n_{dark}} = \frac{I_{dark} \cdot T_{int}}{q}$$ [2.25]

Another already mentioned noise component of importance is the photon 'shot noise,' which originates from the nature of the impinging radiation itself: the number of photons striking on an object also follows the Poisson probability distribution function. Just as in the case of dark current noise, the variance of the mean value of carriers generated by impinging photon flux equals its mean value, i.e. $\sigma_{ph}^2 = \overline{n_{ph}}$. This is a fundamental problem, so the aim for 100% fill factor (the ratio between the photoactive area and the entire pixel area) and quantum efficiency is at the same time the aim for reduced photon shot noise. The ENC due to photon noise, represented by the variance of the number of carriers generated by impinging photon flux occurring during T_{int}, σ_{ph}^2 , is defined by Eq. [2.26].

$$\sigma_{ph}^2 = \overline{n_{ph}} = \frac{I_{ph} \cdot T_{int}}{q}$$ [2.26]

Another type of phenomena also related to the amount of photons actually impinging the silicon photoactive regions are the reflection and absorption characteristics of the layers deposited on top of silicon in every photodetector array. If any electrical signals are to be carried to the photodetector (e.g. digital signals required for the reset transistor, U_{bias} or V_{DD}) or out of it (U_{out}, for example), metal conduction lines must be used; the more the better, as that allows for a more efficient circuit design and spares silicon area. These metal lines must be, nevertheless, electrically isolated from each other and also from the underlying silicon. For this, silicon oxide is normally used. Moreover, to avoid any undesired diffusion of hydrogen or other types of atoms into the silicon or out of it during operation, an additional silicon nitride based passivation layer is normally deposited on top of every IC. The entire system of layers lying on top of the silicon surface resembles the so called Fabry–Perot interferometer (Demtröder, 1996), where the varying transmission function of each material layer is caused by interference between multiple reflections of light caused by each material surface. Constructive interference occurs if the transmitted beams are in phase, and this corresponds to a transmission peak of the material. If the transmission beams are out of phase, destructive interference occurs and this corresponds to a transmission minimum. Whether the multiple reflected beams are in phase or not depends on the wavelength of the incoming radiation, the angle at which the light travels through the material layer, the thickness of

this layer, and the refractive index change between each pair of overlying materials (Hecht, 1989).

A typical impinging radiation wavelength dependent transmission curve of silicon nitride passivation layer, measured for a 0.35 μm CMOS technology, obtained as a ratio of the measured quantum efficiency curves obtained from two similar PDs, one fabricated with the silicon-nitride based passivation layer and the second one fabricated without it, can be observed in Fig. 2.9(a) (Durini and Hosticka, 2007). The wavelength dependent transmission variations due to the reflection interferences of the oxide and nitride overlying layers, as well as a strong absorption for wavelengths shorter than 450 nm, can be observed in the same graph. Additionally, the
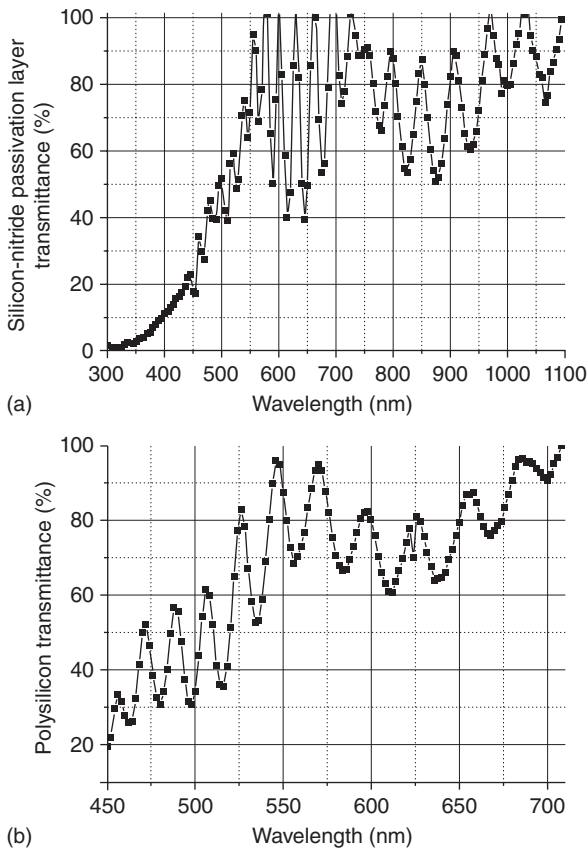
2.9 Wavelength dependent transmittance curves obtained from two similar photodetectors fabricated with and without the overlying: (a) silicon-nitride based passivation layer and (b) approximately 250 nm thick polysilicon (gate) layer.

entire CCD concept, also imported to CMOS based imaging technology in form of so called 'photogates', is based on semi-transparent gate electrodes, normally fabricated out of polysilicon. Polysilicon layers reflect or absorb light and significantly alter the amount of photons actually reaching the photoactive regions in silicon, as can be observed in Fig. 2.9(b), where the wavelength dependent transmission curve of a 250 nm thick polysilicon layer can be observed (Durini and Hosticka, 2007). By carefully observing the graphs in Fig. 2.9, it can be easily concluded why front-side illuminated arrays of photodetectors fabricated in this way are absolutely inappropriate for detection of light in the UV part of the spectra with wavelengths between 100 nm and 450 nm.

Thinning the silicon substrates to below 50 μm to achieve back-side illumination of CCD (Janesick, 2001) or CMOS based imagers for this range of wavelengths has been a very effective solution for the kind of issues that brought a lot of additional technical challenges with it and will be thoroughly discussed in Chapter 4 of this book.

Finally, taking all the above information in mind, the SNR of the pixel configuration shown in Fig. 2.7 can be expressed for the total amount of photogenerated carriers in the structure n_{ph} at the end of T_{int}, as shown in Eq. [2.27] (Durini, 2009).

$$SNR = 10 \cdot \log \left(\frac{\overline{n_{ph}}^2}{\sigma_{dark}^2 + \sigma_{ph}^2 + \sigma_{kTC}^2 + \sigma_{part}^2 + \sigma_{amp}^2} \right) =$$

$$= 10 \cdot \log \left(\frac{\overline{n_{ph}}^2}{\overline{n_{dark}} + \overline{n_{ph}} + \dfrac{k_B T C_{SN}}{q^2} + \dfrac{2\alpha_{part} \cdot k_B T C_{G\_RST}}{\pi^2 q^2} + \dfrac{C_{SN}^2}{q^2} \cdot \overline{v_{n\_amp}^2}} \right)$$

[2.27]

Observing Eq. [2.27] it can be concluded that in order to increase the SNR and lower the signal noise contributions, only a somewhat limited amount of measures can be undertaken:

- Increase the amount of photogenerated carriers n_{ph}. Increasing the signal charge will additionally reduce the influence of the photon shot noise on the overall pixel performance. To achieve this, the amount of photons impinging on the photoactive area of the pixel within T_{int} should be increased. All possible optical losses should be additionally reduced, e.g. those due to reflection or absorption of layers covering the actual silicon. The latter can be achieved by increasing the pixel fill-factor, by eliminating all the reflection or absorption losses of the layers overlying the pixel array, which might be partially accomplished by using anti-reflection coatings (ARC) or even completely removing the inter-metal

isolation stack on top of the pixels, or by increasing the radiation impinging angle for every pixel. All possible electrical losses of photogenerated carriers, mostly due to recombination processes in silicon, must additionally be reduced. This can be achieved by completely depleting the entire volume where absorption of impinging radiation might occur, which is not always an easy task. For higher energy particles, the entire silicon wafer with thicknesses of several hundreds of μm should be completely depleted of majority charge carriers.

- Decrease the dark current and the $\overline{n_{dark}}$. The 'dangling' bonds on the silicon surface or on the walls of neighbouring STI structures should be at all times satisfied to reduce their generation rates and at best 'screen' their noise contributions. This can be achieved by using highly doped 'burying layers' diffused on these regions, and keeping the stored signal charge away from these regions rich in SRH generation-recombination centres, and also by avoiding the SCR reaching these regions, or by increasing the recombination rates in the regions not relevant for phototransduction such as channel-stop layers, etc. Another solution for reducing the dark currents would be to cool down the device, as for every approximate additional 8 °C, the amount of dark current doubles its value.

- Reduce C_{SN} and reset noise components. For the pixel configuration shown in Fig. 2.7, reducing C_{SN} would mean reducing the area of the pixel covered by the n-well (the photodetector itself). This would imply relying mostly on the diffusion currents of photogenerated carriers from the regions outside the n-well for charge collection, bound to inevitable losses through recombination processes on the path. In some cases, the reduction of C_{SN} even represents quite an attractive option, where the performance of the charge-collection should be examined in detail. Nevertheless, the capacity contribution of the PD itself to the C_{SN} is mostly marginal, the SF-gate capacitance being one of the major contributors to it. Hence, the source-follower gate area should be minimized taking into account the trade-off related to its buffering performance directly related to the readout speed and the load impedance of one entire pixel column in CMOS technology or the CCD output signal path. Other effects related to reducing the C_{SN} are the reduction of the pixel full-well capacity, and also an increase in the charge-to-voltage conversion factor (also called pixel gain constant) or the pixel spectral responsivity, as for a reduced C_{SN} and a fixed voltage output swing, more volts per collected electron at the pixel output are obtained.

The commonly used low-frequency noise reduction method in discrete-time digital circuits is the so called 'correlated double-sampling' or CDS

(Janesick, 2001). This method consists of subtracting the actual signal output voltage from a previously sampled SN reset voltage value that already contains all the noise components related to the SN reset operation. In this way, all the low-frequency noise signal components will be eliminated from the resulting signal output voltage. The problem with the pixel configuration shown in Fig. 2.7 is the absence of a memory cell where the reset signal value could be stored during the charge collection stage. For this kind of 3T active pixels, only not-correlated double-sampling can be performed, where the output signal related to one T_{int} is subtracted from the reset signal from the following T_{int}. The assumption here is that the reset function related noise components remain similar between two reset operations under similar operating conditions, which is unfortunately not always the case.

The reduction of the white noise (in the first place, dark and photon shot noise, as well as the flicker noise components) is, however, much more involved. There are several techniques which can be followed for low-noise analog circuit design, as it will be discussed more thoroughly in Chapter 5 of this book. For example, in amplifiers the gain and bandwidth management is absolutely essential. The front gain stages should provide enough gain so that the noise contribution of the following stages results negligible. On the other hand, the transconductances of active loads and current sources should be sufficiently low so their noise is not greatly amplified. In addition, the bandwidth limitation, particularly important in the case of white noise must be also carefully controlled (Hosticka, 2007). Also, buried source-follower metal-oxide-semiconductor field-effect transistor (SF-MOSFET) structures have been proposed for lowering the amount of RTS and flicker noise in pixels, mainly generated due to the in-pixel amplifier operation (Wang, 2008).

2.7   High-performance pixel structures

Statistical values of $0.7 \, e^-_{rms}$ of noise per pixel (Chen *et al.*, 2012), measured by applying the photon transfer method (PTM) for indirect characterization of silicon based imagers (Janesick, 2007) and taken as a basis for the development of the first internationally accepted standard for semiconductor based imagers and camera systems, EMVA 1288, defined by the European Machine Vision Association (EMVA, 2013), have been reported in literature with a tendency of further development in the direction of near single-photon counting capability of silicon based photosensors. This would not have been possible without the breakthrough achieved in 1982, when Nobukazu Teranishi and his team of the Microelectronics Research Laboratories of the Nippon Electric Co. Ltd, proposed what they called a 'no image lag photodiode structure' for the interline CCD image sensors

(Teranishi *et al.*, 1982). Interline CCD (Janesick, 2001) means that each charge-coupled cell consists of two main parts: the photoactive part with the photodetector structure of choice (e.g. a photodiode), and a charge storage cell from where the signal charge will be transported along one entire column of pixels until its serial readout at the CCD output buffer stage. A technology cross-section of the last three cells of such an interline CCD (where U_{CCD_m} is the biasing voltage of the last CCD cell m), followed by the floating-diffusion SN, the reset gate biased through U_{reset}, and the drain diffusion (forming all together the reset transistor or the 'digital signal regeneration stage' as proposed by Kosonocky in Kosonocky and Carnes (1971), is schematically shown in Fig. 2.10. Image lag depicts the residual charge that remains in the photodetector or in the photodetector SN after the reset operation has been performed adding this residual charge to the new collected one causing false output information, diminishing the full-well capacity of the photodetector and increasing the noise floor. To eliminate the image lag, the authors lowered the donor concentration of the sandwiched *n*-type region of the PD in order to lower the potential required to completely deplete the PD from the majority carriers, lowering
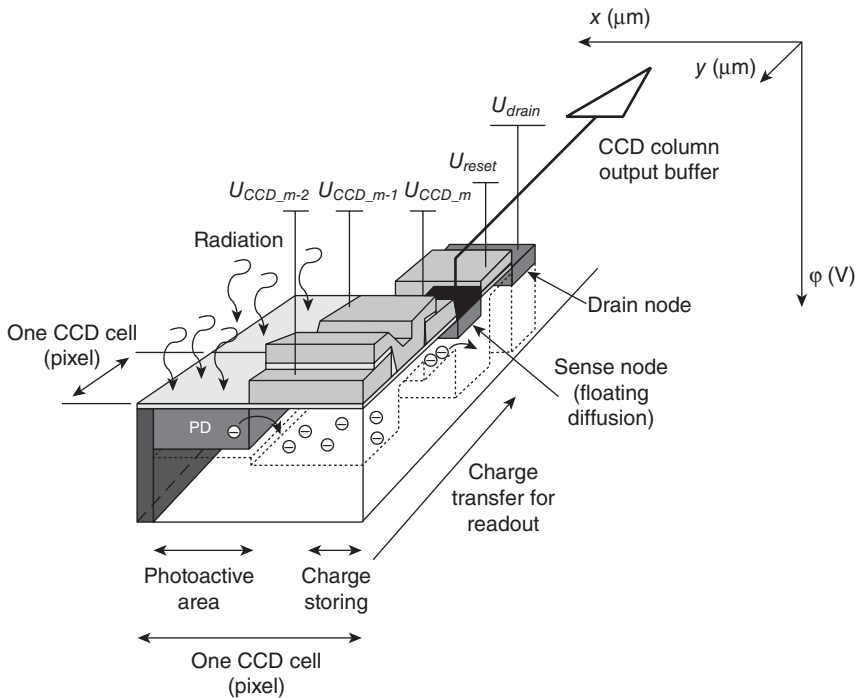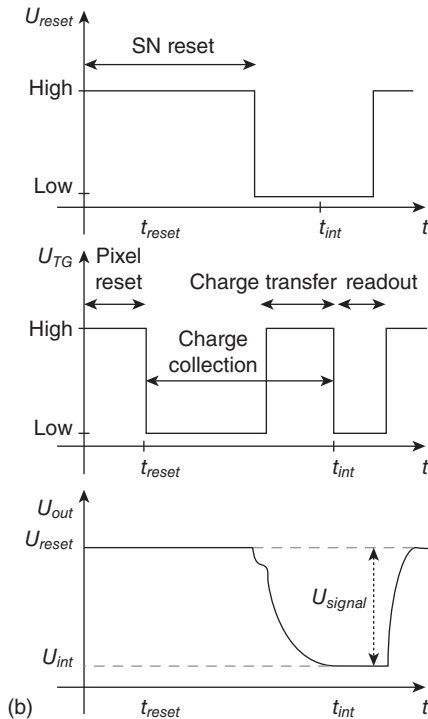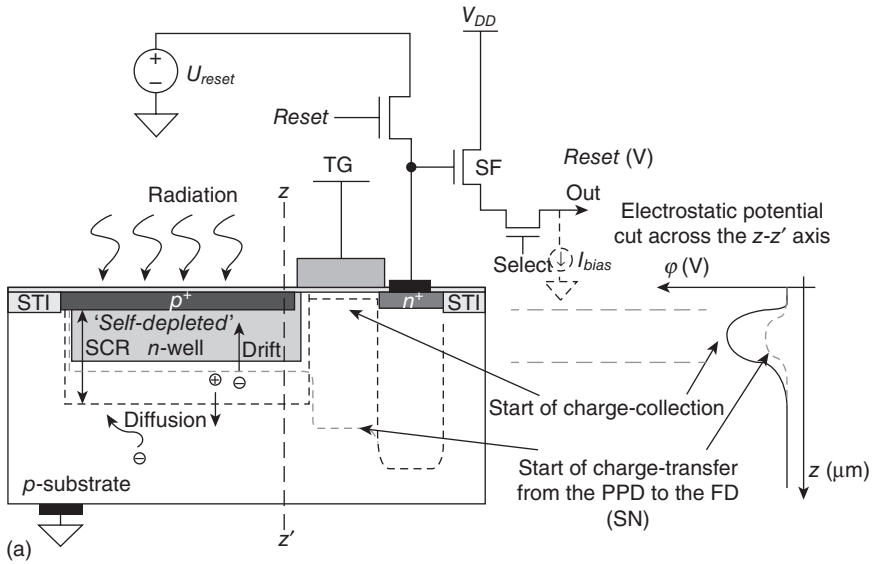

2.10 Schematic technology cross-section of an interline CCD cell (pixel).

the charge transfer time significantly (Teranishi *et al.*, 1982), and introduced the p^+ layer on the silicon-oxide interface of the p-MOS CCD technology to prevent the signal electrons from reaching this interface rich in SRH generation-recombination sources.

It was the birth of the nowadays well-known pinned photodiode (PPD), extensively used in interline CCDs by Eastman Kodak Co. (Burkey *et al.* 1984) and other CCD producers from the very beginning of this technology, and successfully imported into the world of CMOS based image sensors giving birth to a 4T pixel structure (because of the four transistors in the pixel structure: the reset transistor, the select transistor, the SF and the transfer-gate) as reported in a joint work of Eastman Kodak Co. and Jet Propulsion Laboratory - JPL (Lee *et al.*, 1995), and further optimized by Eastman Kodak Co. and Motorola (Guidash *et al.*, 1997), depicted in Fig. 2.11 for a CMOS technology based on a p-type substrate, and thoroughly described across this entire book.

As can be observed in Fig. 2.11, the PPD consists of an n-well, the doping concentration of which is tailored to make it fully depleted if sandwiched between a highly doped grounded 'pinning' p^+ layer on the silicon surface, and the equally grounded p-type silicon substrate. The p^+ layer satisfies all the dangling bonds at the silicon surface, drastically diminishing the thermal EHP generation rate and noise contributions through electron trapping and de-trapping mechanisms normally originating there, and pins the surface potential to the ground level. In this way the self-depleted n-well remains isolated from most of the SRH generation-recombination sources. To ensure this, additionally a certain minimal distance should be incorporated between the n-well and the STI walls, which should at best be covered with p^+ layers. The polysilicon transfer-gate (TG) is used to electrically decouple the PPD from the SN floating-diffusion.

As can be observed in the time-diagrams depicted in Fig. 2.11(b), the PPD is at first emptied of all remaining electrons collected there by shutting the reset transistor and the TG simultaneously ON (applying U_{reset} to the floating-diffusion (FD) and $U_{TG\text{-}high}$ to the TG). By doing so, the electric field induced in the channel region of the TG will affect the fringe region between the TG and the PPD (some hundreds of nm) creating a drift-field there that will cause the accumulated electrons to drift from this TG-PPD interface region through the TG channel into the FD. For this to happen, the electrostatic potential in the FD, the TG and the PPD must fulfil the condition $\phi_{FD} > \phi_{TG} > \phi_{PPD}$ throughout the entire process of charge transfer from the PPD to U_{reset}. The charge now 'missing' at the TG-PPD fringing region will cause a charge gradient across the PPD, and the electrons will start diffusing from the regions with higher electron concentration, away from the TG, into this lower electron concentration region. Once they get there, they will be equally drifted and transferred to U_{reset} sustaining the

2.11 (a) Schematic technology cross-section of a p^+np junction based pinned photodiode 4T pixel structure, showing the SCR and the drift and diffusion currents arising under illumination, and the electrostatic potential profile across the z-z' axis; (b) reset (RST) and transfer-gate (TG) signal time diagrams together with the sense-node (SN) output signal.

concentration gradient as long as there are still remaining electrons left in the PPD. For the latter, it must be considered that the charge leaving the PPD will cause the U_{PPD} to increase until the so called pinch-off or pinning voltage $U_{pinch-off}$ is reached within the PPD, where the SCRs caused by the p^+-n-well and the n-well-p-substrate p-n junctions overlap making impossible any further potential increase within the PPD. This is the pixel reset operation, as defined in the time-diagram in Fig. 2.11(b).

For the charge-collection operation, the FD remains in reset mode whilst the photoactive region of the pixel gets electrically decoupled from the SN by shutting the transfer-gate OFF. The thermally and photo-generated electrons start diffusing into the PPD if generated within the diffusion distance from the PPD SCR, or get directly drifted to the region of maximum electrostatic potential if generated within the SCR. This region of maximum electrostatic potential within the PPD is located away from the silicon surface thanks to the pinning layer introduced there, as can be observed in the one dimensional cut across the z-z' axis made for the electrostatic potential and shown on the right side of Fig. 2.11(a). Due to the gradient in the electrostatic potential induced at the silicon surface, all the EHP generated by short wavelength impinging radiation (in the UV and blue part of the spectra) will get fast drifted into the potential maximum, reducing their recombination rate within the highly doped pinning layer. The spectral responsivity of the PPD is thus increased for these parts of the spectra if compared with other p-n junction photodetectors. By further accumulation of electrons within the PPD, its electrostatic potential will eventually sink below $U_{pinch-off}$, the PPD will no longer be fully depleted, and two parallel connected capacitances will appear originated by the two p-n junctions (see Fig. 2.11(a)). This will increase the full-well capacity of the PPD.

At a certain point defined by the image sensor frame rate, the reset transistor is shut OFF, and the FD is left floating. Once enough time has passed for the reset transistor to get settled and if the correlated double sampling operation is to be performed, at this point the reset pixel output signal is to be read out through the SF stage and saved somewhere at the input of the CDS stage. Next, the TG is shut ON and the collected charge is transferred from the PPD into the floating SN, making its electrostatic potential sink proportional to the amount of charge transferred into it. For a properly designed PPD-TG interface where neither potential barriers nor additional potential wells appear on the charge transfer path, the 100% charge transfer is complete after a couple of hundreds of ns. To compensate the output voltage shift caused by the capacitive coupling of the FD to the transfer-gate (the TG is shut OFF at first) followed this action by the readout of the actual pixel output signal, proportional to the optical power of the impinging radiation integrated over T_{int}. At the CDS stage, this output value is subtracted from its now directly correlated reset value (some

authors call this process a true CDS) minimizing all low-frequency noise contributions to the signal in this process. The pixel output signal is now ready for digitalization.

Just to summarize, being the pixel photoactive area electrically decoupled from the pixel sense-node enables charge-collection and simultaneous pixel non-destructive readout i.e. (true) CDS not existent in CCD technology, enhancing high-speed imaging applications and yielding low-noise performance. The capacitance of the SN can be tailored for an increased charge-to-voltage conversion factor (pixel gain), reduced k_BTC-noise, and desired pixel full-well capacity (FWC) (or saturation capacity, as defined in the EMVA 1288) mostly defined by it. The FD should be designed as small as possible to minimize its dark current and other leakage current contributions related to the silicon substrate effects, defining the SN capacitance mainly through proper design of the SF-gate and other parasitic capacitance contributions. For global shutter applications (thoroughly discussed in various chapters of this book), where the FD must retain the signal information over several milliseconds, its so called 'global shutter efficiency', defined as the difference between the original signal and the signal change rate over time, should be as close to the original signal as possible. All these attributions make the PPD based 4T pixel quite attractive.

A similar approach was pursued by JPL in the beginning of the CMOS based image sensors (Mendis *et al.*, 1993), where they proposed a Photogate 4T active pixel structure where, instead of the PPD, a MOS-C photodetector was used, using most of the just mentioned advantages but still suffering under MOS-C structure related drawbacks explained earlier in the text.

2.8   Miniaturization and other development strategies followed in image sensor technologies

Scaling of the CMOS technology was a driving force of the microelectronics industry and market for several decades. The increase of IC performance and capabilities, and perfection of IC manufacturing technologies were constantly causing growth of the market. A well-known trend described by Gordon Moore (the co-founder of Intel) in 1965 (Moore, 1965) states that the number of transistors per unit area of an IC roughly doubles every two years, which indirectly translates into a major increment of the IC functionality through the reduction of the process technology node, as can be observed in Fig. 2.12.

Nevertheless, following Moore's law is becoming more and more challenging. Every next step in the reduction of the feature size means miniaturization of IC building blocks like transistors and interconnection lines, decreasing the distance between them, and lowering bias voltages
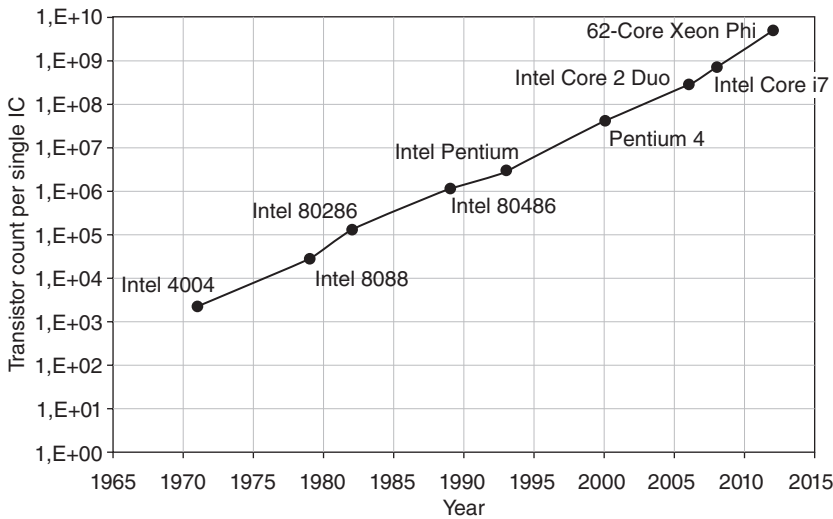
2.12 Transistor count of the Intel processors versus production year. Note the logarithmic scale of the vertical axis. Data from Intel Corporation, 2013.

accompanied by the decrease in the thickness of the transistor gate dielectrics. One of the fundamental problems of semiconductor technology scaling on the transistor level is a physical phenomenon called tunneling, where an electron (or other elementary particle) having a specific energy is able to travel through a potential barrier of a higher energy level. In modern CMOS technologies, e.g. those with minimum feature sizes below 90 nm, the gate-oxide thickness is in the order of one nanometer, a thickness at which the tunneling effect is likely to occur. Further reduction of the gate oxide thickness only increases this probability. Such uncontrolled electron transition through the gate oxide is a source of leakage currents and causes unstable operation of the transistor. The degradation in the performance of analog transistors occurs in transistors with even thicker gate dielectrics. Unlike in digital designs, transistors of analog circuits are not working as simple switches spending most of the time in one of the two basic states (ON or OFF), but acting as amplifiers. For the interconnection lines the technology node reduction translates into higher wire densities, larger average distances the traces have to cover, higher resistances, and higher parasitic capacitances, all of which cause signal propagation delays preventing further increase of the circuit response speed. Finally, increase in the production costs is unavoidable as more sophisticated photolithography techniques are required for further developments.

To be able to fight against performance deterioration and nevertheless continue with developments pursuing Moore's law, several strategies have

been developed. They could be divided into two major approaches: the one following Moore's law called 'More Moore', normally pursued in pure digital developments, and the one pursuing the increase of analogue or transduction functionalities in ICs is called 'More than Moore' approach.

Following the 'More than Moore' approach, state-of-the-art CMOS integrated image sensor (CIS) technology is reaching a point where its overall performance is becoming comparable to the one normally expected from CCD based image sensors. The CIS performance is getting closer nowadays to – although still on a research and development level despite very promising results – the one expected from broadband and image intensifier devices such as those based on micro-channel plates (MCP), photo-multiplier tubes (PMT) or similar technologies. The latter is mainly due to the fact that nowadays CMOS based fabrication technologies are evolving in the direction of enhanced imaging performance while maintaining their CMOS manufacturability characteristics: high yields and other maturity advantages of CMOS manufacturing technology at very competitive prices. This is definitely a new trend, as it has not been that long since the fabrication of CMOS imagers meant high logical complexity on the pixel and the imager levels, following the 'camera-on-a-chip' concept (Fossum, 1995), although tolerating quite bad front-end optical performance when compared with CCDs (Magnan, 2003). Improving the sub-optimum photodetection performance means undertaking dedicated changes on the CMOS manufacturing level. A considerable effort, both CMOS process related and also economically speaking, is to be made in order to develop such technologies. The result is that not every CMOS foundry is currently able to produce state-of-the-art CMOS imagers that can compete on the global market. In fact, there is only a handful of CMOS-foundries nowadays (Aptina Imaging, Taiwan Semiconductor Manufacturing Company Ltd (TSMC), Sony, etc.) that can still afford the necessary investment in order to remain competitive in the commercial market of standardized CIS normally found in cell-phones, PCs or tablets. These CIS are always aiming at (mostly dictated by marketing policies) smaller pixel sizes and higher pixel count in imagers. The challenges are enormous in this field of application and normally technology driven (thoroughly described in Chapter 7) as we are already at the physical limit of a relatively acceptable front-end photo-transduction performance of commercial imagers with pixel sizes getting below $1.1\,\mu m$. These pixel sizes always produce more noise and, basically, increasingly worse images, basing its success on effective on-chip and out-chip image processing tools.

On the other hand, there are numerous high-performance niche markets exhibiting strong need for imagers with increasingly ambitious performance specifications: sensitivity nearing single-photon detection (with standard active pixels reaching noise floors below $1\,e^-_{rms}$, as described above), spectral

responsivity beyond the silicon capabilities (by incorporating other detector material to CMOS processed image sensors), high-speed, huge dynamic ranges, on-chip intelligence, etc. The developments aimed at applications in machine-vision (see Chapters 5 and 6), automotive (as described in Chapter 8) or space applications (as described in Chapter 9), high energy particle detection (as described in Chapter 10), biology and medicine (as described in Chapters 11, 12, and 14), as well as many other industrial and scientific applications, or even broadcast and HDTV as explained in Chapter 13, are performance driven and base their goal specifications on the specific application requirements. Attaining these high-performance specifications often implies having CMOS technologies specialized to extreme degrees. The measures undertaken are normally also accompanied by CMOS post-processing steps such as wafer thinning, flip-wafer or flip-die bonding, wafer surface passivation, and so on, thoroughly described in Chapters 4 and 14 that deal with back-side illumination image sensors. On the other hand, broadband and extended spectral imaging has been enabled through new hybridization and micro-structuring techniques, where other photodetection materials have been used together with CMOS readout integrated circuits (ROIC) in integrated solutions.

Being able to substitute the currently used highly expensive technologies (such as PMT or MCP), which need cooling, thousands of volts for biasing and quite expensive infrastructure by compact, integrated, affordable and reliable semiconductor (e.g. CMOS) solutions is a goal worth working for. 'Camera-on-a-chip' CMOS solutions with front-end performances of the quality normally expected from CCDs but having them at more accessible prices and increased logic complexity, is also an attractive goal. Nevertheless, the over 30-year-old, and mature, CCD technology has its own advantages over CMOS technology which points out an unavoidable coexistence of both technologies, in the first place due to the degrees of freedom CCD technology has for manufacturing process changes, if compared with any CMOS technology, with the performance and not necessarily the cost as the driving element.

Near single photon counting with picosecond time resolution has been one of the main breakthroughs of the last couple of years. This was achieved in the form of single-photon counting avalanche diodes (SPADs) as proposed by Sergio Cova and his team (Cova et $al.$, 1981), based on the pioneering work of Roland H. Haitz (Haitz, 1965), with biasing voltages well above the p-n junction breakdown voltage V_{br} and special quenching circuits. They are realized nowadays in advanced CMOS technologies (Karami et $al.$, 2012), which as explained in Chapter 11, yields quite acceptable SPAD fill-factors, achieving the performance normally expected from special dedicated processes for SPADs (Zappa et $al.$, 2007). The CMOS SPADs yield nowadays less than 100 dark counts per second (cps) at room

temperature for a 30 μm diameter circular SPAD photoactive area and less than 100 ps time-jitter, fabricated in a 0.35 μm CMOS technology as reported in Bronzi *et al.*, (2013). Based on the same concept, but joining several avalanche photodiodes working in the so called Geiger mode (just as used by SPADs), the concept of so called silicon photomultipliers was proposed as an alternative to the PMT technology (Dautet *et al.*, 1993). The results obtained from these technologies are very promising, but still have many drawbacks that have to be solved before they enter the market in the form of products: pixel sizes, dark count-rates, dead-times and fill-factors in case of SPADs; and fabrication technology related issues, as well as sometimes insufficient spatial resolution, in the case of silicon photomultiplier (SiPMs).

The quite attractive and more developed CCD based counterpart of SPADs and SiPMs using avalanche processes for near single-photon counting performance was proposed in 1992 by Jaroslav (Jerry) Hynecek working at Texas Instruments (Hynecek, 1992); named low-noise charge-carrier multiplier (CCM) and located in a CCD channel. The main idea of the invention is to originate quite low and controllable avalanche process based multiplication of the signal charge in each CCD cell which adds, keeping the multiplication noise low, from one cell to the other across a CCD register until multiplication values of several hundreds are occasionally reached. In 2001 Paul Jerram and his team (Jerram *et al.*, 2001) of Marconi Applied Technologies (afterwards e2V) applied the same concept to their low-light-level CCDs (LLLCCDs) which caused commercial production of cameras based on these concepts under the name of EMCCD, offered since then by several manufacturers.

As an alternative for dealing with all the mentioned problems accompanying the 'More Moore' developments, the American physicist Richard Feynman proposed in his lectures given in Tokyo in 1985 an IC stack to increase processing power and functionality of the processors. He stated that 'another direction of improvement is to make physical machines three dimensional instead of all on the surface of a chip. That can be done in stages instead of all at once: you can have several layers and then many more layers as time goes on' (Nishina, 2008). The implications of this new approach and the developments undertaken ever since in this direction will be briefly addressed below.

2.9   Hybrid and 3D detector technologies

Three dimensional IC stacking has been known since the 1970s as 3D packaging: the ICs are stacked vertically, on top of each other and interconnected using wire-bonds as shown in Fig. 2.13(a). Each IC of the stack has special bonding pads where wire bonds are attached. The choice of material for the pads and wires depends on the application (Al, Au or
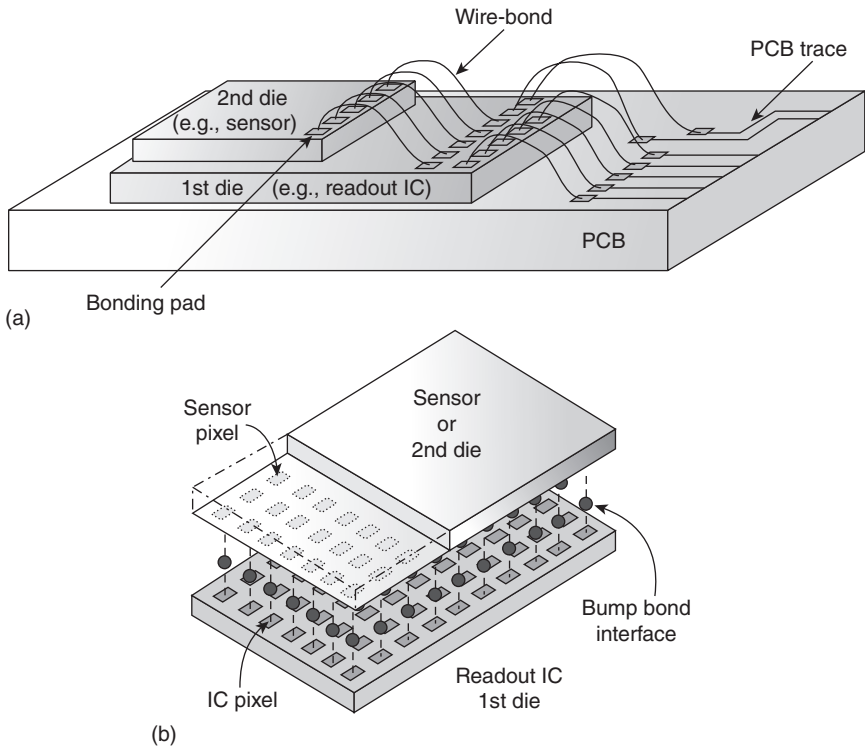
2.13 (a) A concept of stacked three dimensional wire-bonded IC. Wire bond pads are located at the periphery and (b) a diagram of three dimensional IC stacking using bump-bond array (flip-chip) technology. Bond pads occupy a significant part of the die surface (Arutinov, 2014).

Cu are normally used). The wire-bond interconnection density is not very high: pads are usually located at the periphery of the ICs with the pitch rarely falling below 75 µm (for very low pitch wire-bond interconnection see Zhong (2009). Higher flexibility, reliability and yield are the main advantages of this bonding technique. The disadvantages are low density of interconnections, high impedance caused by the required wiring (which yields large RC delays and high power consumption), and a large footprint.

The so-called flip-chip technology is another die-to-die bonding technique. The interconnection between vertically stacked ICs is provided by an array of metallic bumps located between the separately fabricated ICs (see Fig. 2.13(b)). All connection lines which are shared between the two dies have to be routed to the special bond pads located on the surfaces of the dies. The pads from the two ICs are then interconnected using an array of

metallic bumps as shown in Fig. 2.13(b). A good overview of different flip-chip techniques can be found in Rossi *et al.* (2006). The advantage over the previous stacking technology comes from short interconnection lengths between the ICs and the higher interconnection density. Disadvantages of this technology are firstly its high cost and secondly that, in general, no more than two dies can be stacked and bonded together.
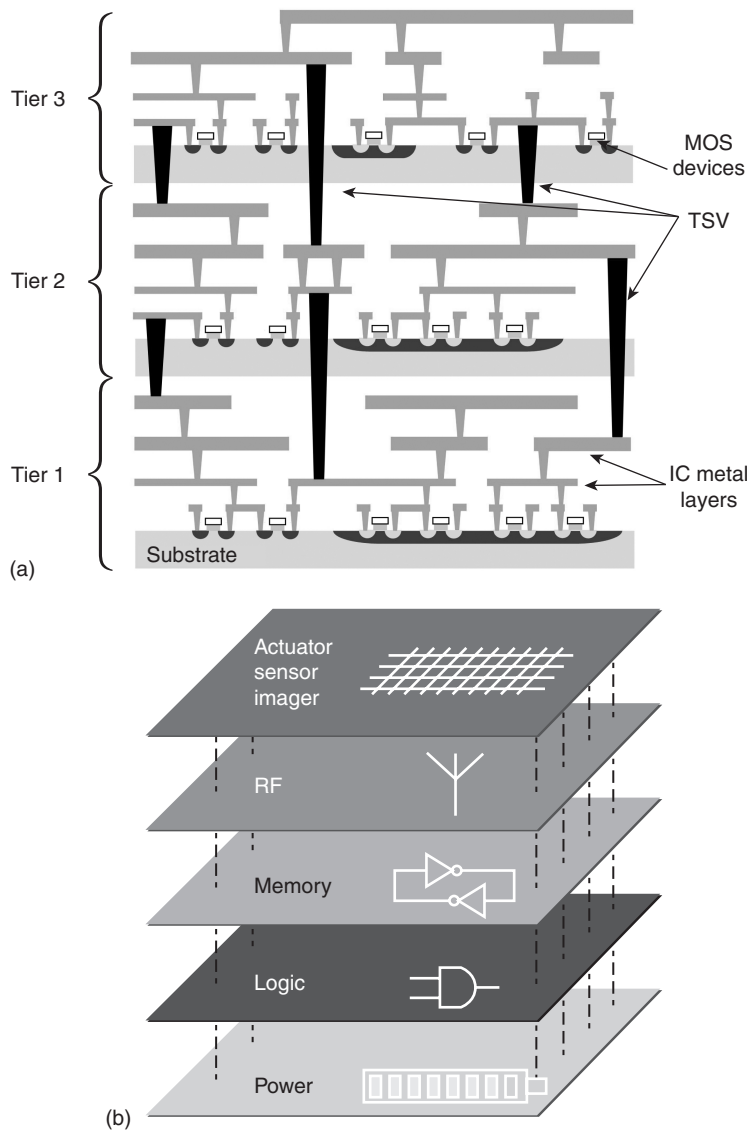
The flip-chip technique is also used in the development of imaging devices. The basic idea is to split the imager into two separate parts: the sensor part and the readout IC part, and interconnect them using this technique. The readout IC contains pixel related active readout and signal processing electronics. Every photoactive pixel is connected to the corresponding readout IC pixel (amplifier input) using a metallic (usually indium) bump bond. Modern bump-bonding techniques allow as low as $30\,\mu m$ bonding pitches (Chung and Kim, 2007). Such implementation allows having a pixel with near 100% fill-factor and combining sensor and image processing functionality into a single stacked IC.

A good example of such a development is the 4096 pixel photon counting readout IC – MEDIPIX (Bisogni *et al.*, 1998), developed at CERN (the European Organization for Nuclear Research). Every square pixel of the MEDIPIX IC has the size of $170 \times 170\,\mu m^2$ and contains an amplifier, a digitizer module, a digital counter, and some other signal processing and pixel controlling electronics. The readout mechanism of the IC is based on the shift register principle, consisting of the bits of the in-pixel digital counters.

The MEDIPIX chips have been mated (bump bonded) with different silicon as well as GaAs based photosensors. They can be used for X-ray direct conversion readout, and also for other ionization irradiations and charged particle detection in different applications, although initially developed for medicine.

A big step forward in vertical IC stacking was the introduction of the 3D integration technology that can integrate multiple dies (in this case called tiers or strata) into a single monolithic package. The basic idea of 3D integration is to split a sensor into several parts, produce them as separate tiers, stack them vertically, and interconnect them using a relatively new technology – the so called through silicon via (TSV) technique. The via (vertical interconnect access) represents a vertical pillar of conducting metal (e.g. copper or polycrystalline silicon) which is deposited into an opening etched through a die or a wafer to form an electrical and thermal connection and mechanical stability for the stacked tiers.

In addition to the TSV technology, the tiers of a 3D stack can be also interconnected using a wafer bonding interface which provides additional electrical interconnection and mechanical stability (see for example Bower *et al.* (2008)). Figure 2.14(a) shows several vertically stacked tiers

Tier 3

Tier 2

Tier 1

Substrate

(a)

MOS
devices

TSV

IC metal
layers

Actuator
sensor
imager

RF

Memory

Logic

Power

(b)

2.14 (a) A cross-section of a diagram of a 3D integrated circuit
consisting of three tiers (dies) stacked on top of each other and
interconnected using through silicon vias (black pillars) and (b)
technology mixing possibilities of 3D integration (Arutinov, 2014).

interconnected using the TSV technique. As can be seen from the figure, the vias (dark pillars) connect metal layers of two neighboring dies. The vias are usually produced either during the fabrication process of a die (or wafer) itself, or can also be implemented at later stages, after the IC fabrication, in a post-processing step. A typical via diameter starts from less than a micrometre and is not limited in size. A typical aspect ratio (regarding the vertical etching rate related to the lateral one) for TSV fabrication usually varies between 1 and 10. In general, TSV equipped wafers need to be thinned down in order to access the vias before the tier integration process can start.

In a well thought-out design of a 3D integrated IC, a circuit functionality can be split among several tiers in such a way that circuit blocks sharing a significant amount of electrical interconnection appear right on top of each other, and normally (in conventional planar technologies) required long traces can be substituted by short via based interconnections. This dramatically decreases parasitic delays and allows the implementation of very large input/output (I/O) buses, the implementation of which would be complicated if developed in conventional planar CMOS technologies.

Apart from the advantages on a circuit level, 3D integration technology also offers an interesting possibility for mixing tiers developed in different technologies with different feature sizes and functionalities (an approach also called a heterogeneous integration), a possible implementation of which can be observed in Fig. 2.14(b). By applying this approach, various actuators, biosensors, radio frequency devices, solar cells, or passive components can be incorporated into a single monolithic IC. At the same time, an integration of ICs developed in similar or even identical technology can dramatically increase the count of transistors per unit area even without shrinking the technology node: an approach called homogeneous integration.
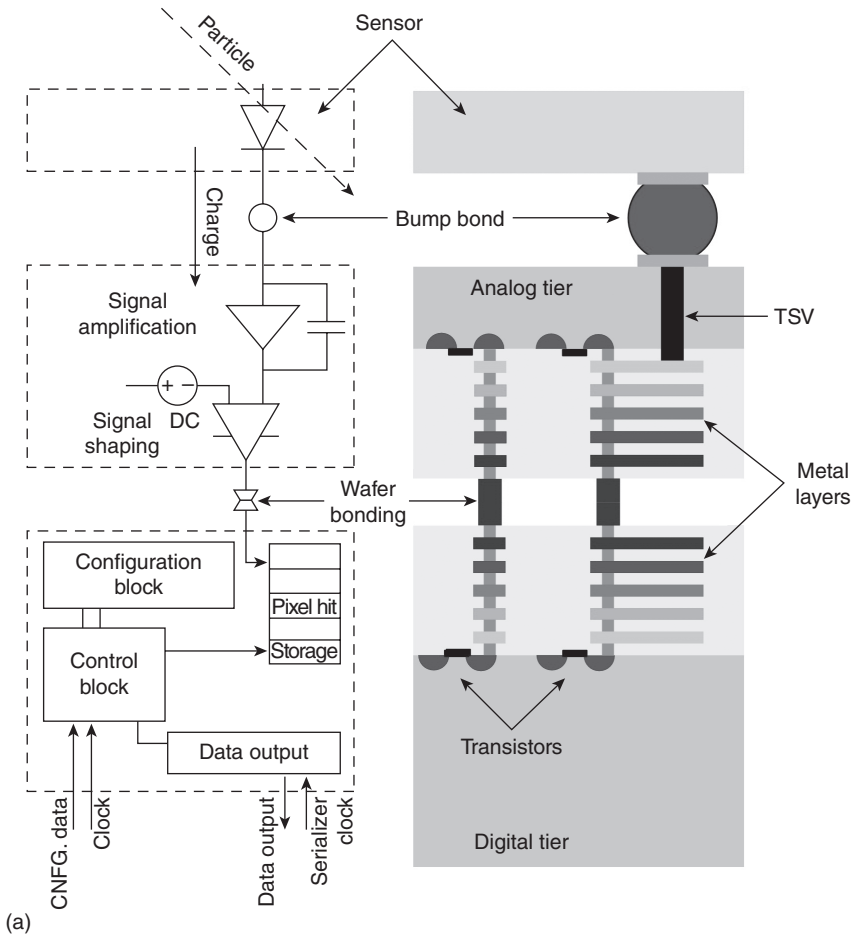
The 3D integration technology also introduced new challenges. Being an immature technology if compared with, for example, CMOS, it still lacks acceptance in the mainstream industrial applications. This leads to an absence of serious standardization, and due to this fact also of dedicated electronic design automation (EDA) tools. All of this makes this kind of development process quite complicated and for the time being not very efficient. Nevertheless, many research institutes and industrial companies across the world are starting to look in this direction and recognize the evident advantages this technology might bring. Smaller packaging size, the possibility of mixing different processing technologies for different parts of a sensor, high data processing performances, small footprint, higher speed and enhanced computing capabilities are all features offered by 3D integration.

Contemporary digital still camera sensors contain millions of pixels. Demands on image quality require powerful digital signal processing and
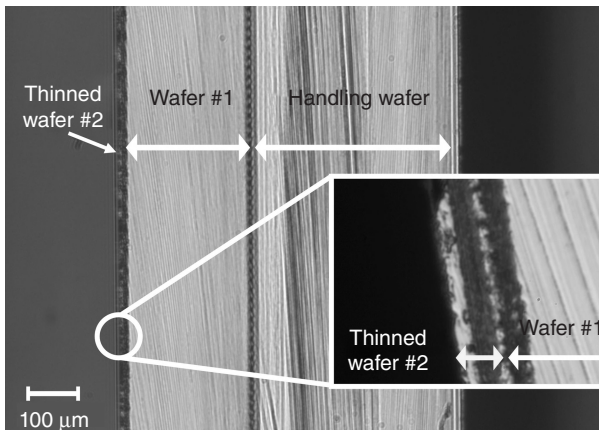
data acquisition. At the same time, pixel sizes rarely exceed a few square micrometres, a fact that prevents the implementation of complex circuits and functionalities in a pixel at acceptable fill-factors. A possible solution could be to apply a 3D stacking approach. Nevertheless, due to large bond pitches flip-chip techniques cannot be used for this task. On the other hand, 3D integration technology could become the technology of choice for solving this problem due to much smaller TSV pitches with results comparable to the pixel size of the camera sensors. As an example of a successful implementation of an image sensor developed in 3D integration technology, the new Sony IMX135 CMOS image sensor can be mentioned; introduced at the Content Strategy Conference (ConFab) in 2013 (Dick, 2013). The two tier 3D sensor consists of a back-illuminated image sensor developed in 90 nm technology bonded to an image processor fabricated in 65 nm CMOS technology.

Over the past few years, 3D integration technology also became a field of interest for the high energy physics community. The ability to mix different fabrication technologies for different sensor parts and be able to offer much higher functionalities makes this technology very attractive for the development of particle detectors. In the framework of the upgrade of the pixel detector (Wermes and Hallewel, 1998) of the ATLAS (The ATLAS Collaboration *et al.*, 2008) experiment at CERN, several two tier prototypes of the pixel detector readout chip have been developed in 3D integration technology.

The readout ICs were developed in the 130 nm CMOS technology of GlobalFoundries, whilst the 3D integration of the fabricated tiers was performed by Tezzaron Semiconductor. One of the prototypes developed, the FE-TC4 (Fig. 2.15(a)), represents a two tier pixelated sensor readout IC, consisting of 854 pixels with a pixel size of $50 \mu m \times 166 \mu m$. The 3D pixel functional diagram and its technology cross-section (out of scale) diagram are both shown in Fig. 2.15(a). One of the tiers provides the readout and analog signal processing parts of the pixel detector. The second tier provides the digital signal processing circuitry. After bonding, the back side of the tier providing the analog pixel functionality is thinned down to 12 μm (see Fig 2.15(b)) enabling access to the TSV structures, connected to the inputs of the analog pixels. The last step in the fabrication process consists of applying a metallization layer on top of the thinned tier of the 3D stack and preparing it for flip-chip bonding to the sensor, a separate array of photodetectors that might be fabricated in any technology of choice (including besides any silicon based developments, also III-V semiconductor or any other materials). Various tests of the prototype show positive and promising results with results comparable in performance to the sensor developments carried out in 2D CMOS technologies (Arutinov *et al.*, 2013), which makes this technology feasible for future imaging applications.

Particle

Sensor

Charge

Bump bond

Analog tier

TSV

Signal
amplification

Signal DC
shaping

Metal
layers

Wafer
bonding

Configuration
block

Pixel hit

Control
block

Storage

Data output

Transistors

CNFG. data

Clock

Data output

Serializer
clock

Digital tier

(a)

Thinned
wafer #2

Wafer #1

Handling wafer

Thinned
wafer #2

Wafer #1

100 μm

(b)

2.15 (a) Graphical representation of the pixel of the ATLAS pixel
detector readout IC prototype FE-TC4 developed in 3D integration
technology and (b) a photograph of the cross-section of the FE-TC4.
The bonding interface is zoomed (Arutinov, 2014).

In parallel to this, interesting lines of research nowadays also include: black-silicon, on-wafer optics, integration of laser modules and photodetectors on the same chip or in the same package for telecommunication applications, embedded CCD (CCD mechanism implementation using CMOS technology), and several others. The main challenges regarding CIS, especially those used in industrial applications or having high-performance characteristics, can be found in the existence of specialized CMOS processes and foundries willing to undertake major changes to their processes even for a reduced amount of chips per year. This is not a small problem for a steadily increasing number of small and medium enterprises (SMEs) with complicated requests and reduced budgets.

2.10   Conclusion

We have reached a point where further optimization of the CMOS technologies has to be fully exploited by advantageous sensor circuit designs and new emerging 3D integration technologies, as shown in Fig. 2.16. The latter also requires an exchange of information about CMOS processes
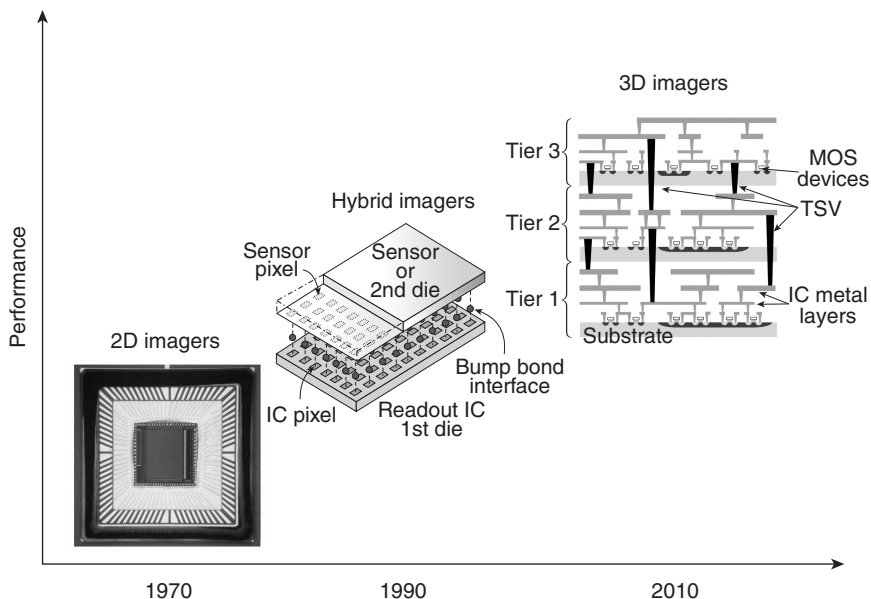
2.16 From 2D image sensors to 3D imagers (photograph by Stefan Bröcker).

used for developments of 2D image sensors or the different tiers to be used in 3D integration between the foundry and the design houses, and also between the foundry and the 3D integrator company specialized in post-processing, which is an issue in most cases. The trend is, nevertheless, that in future one single foundry also provides post-processing expertise.

Although much improved and already comparable with CCD performances, the 2D CIS still suffer from miss-match related photo-response non-uniformity (PRNU) and dark signal non-uniformity (DSNU) problems from pixel to pixel, that has never been a huge issue in CCDs. This is a drawback compensated by the *x-y* independent pixel addressing possibility present in CIS, and not available in CCDs. The much lower dark signals achieved in CCDs through process optimization have been recently matched by process modifications in specialized CMOS processes, but these possibilities are limited in CMOS technology, as all the other advantages of this technology must remain (yield, complex logic, on-chip processing, etc.). There are currently efforts being undertaken even to develop embedded CCD process modules in CMOS technologies to be able to exploit the advantages of both readout approaches, e.g. time-delayed integration (TDI) readout and column-wise analog-to-digital converter (ADC) with on-chip signal processing in a single integrated development. The problem concerning the shared silicon substrate among the photoactive and standard logic building blocks in the same circuit explained earlier in the text would nevertheless remain unsolved limiting the photosensing performance of the CCD part. A step forward in this direction would be to 3D integrate a real CCD photodetector part with all its performance advantages with an underlying CMOS fabricated readout and signal processing counterpart tier. Many of these issues will be discussed across the following chapters in this book. On the other hand, 3D integration brings a series of new challenges where image sensor performances are concerned that are yet to be solved. In both cases, there is much to be done in the future.

2.11   References

Amelio G. F., Tompsett M. F. and Smith G. E. (1970) 'Experimental verification of the charge coupled device concept', *Bell Syst. Tech. Journal*, 49(4), 593–600.

Arutinov D., (2014) 'Conceptual design and testing of the ATLAS pixel detector front-end integrated circuit in small feature size process and 3D technologies', Ph.D. Thesis, University of Bonn, Germany (in preparation).

Arutinov D., Barbero M., Gonella L., Hemperek T., Hgging F., Krger H., Wermes N., Breugnon P., Chantepie B., Clemens J., Fei R., Fougeron D., Godiot S., Pangaud P., Rozanov A., Garcia-Sciveres M. and Mekkaoui A. (2013) 'Experience with 3D integration technologies in the framework of the ATLAS pixel detector upgrade for the hl-lhc', *Nuclear Instruments and Methods in Physics Research*

Section A: Accelerators, Spectrometers, Detectors and Associated Equipment (in Press).

Bisogni M. G. *et al.* (1998) 'Performance of 4096 pixel photon counting chip', *CERN-EP-98-162*, CERN, Geneva, Switzerland.

Bower C., Garrou P. and Ramm P. (2008) *Handbook of 3D Integration: Technology and Application of 3D Integration Circuits*, Wiley-VCH, Weinheim.

Boyle W. S. and Smith G. E. (1970) 'Charge coupled semiconductor devices', *Bell Sys. Tech. Journal*, 49(4), 587–593.

Brennan K. F. (1999) *The Physics of Semiconductors with Applications to Optoelectronic Devices*, Cambridge University Press, Cambridge.

Bronzi D., Villa F., Bellisai S., Tisa S., Tosi A., Ripamonti G., Zappa F., Weyers S., Durini D., Brockherde W. and Paschen U. (2013) 'Large area CMOS SPADs with very low dark counting rate', *Proc. SPIE 8631, Quantum Sensing and Nanophotonic Devices X*, 86311B.

Burkey B. C., Chang W. C., Littlehale J. and Lee T. H. (1984) 'The pinned photodiode for an interline-transfer image sensor', *IEDM Techn. Digest*, 28–31.

Chamberlain S. G. (1969) 'Photosensitivity and scanning of silicon image image detector arrays', *IEEE J. Solid-State Circuits*, SC-4, 333–342.

Chen Y. *et al.* (2012) 'A $0.7e^-_{rms}$ temporal-readout-noise CMOS image sensor for low-light-level imaging', *Solid-State Circuits Conference Digest of Technical Papers* (ISSCC), IEEE International, pp. 384–386.

Chung Seung-Min and Kim Young-Ho (2007) 'Fine pitch and low temperature bonding using solder bumps and the assessment of solder joint reliability', *Materials Transactions*, 48(1), 37.

Compton A. H. (1923) 'A quantum theory of the scattering of X-rays by light elements', *Phys. Rev*, 21(5), 483–502.

Cova S., Longoni A. and Andreoni A. (1981) 'Towards picosecond resolution with single-photon avalanche diodes', *Rev. Sci. Instr.*, 52, 408.

Dautet H., Deschamps P., Dion B., MacGregor A., MacSween D., McIntyre R. J., Trottier C. and Webb P. P. (1993) 'Photon counting techniques with silicon avalanche photodiodes', *Appl. Opt.*, 32, 3894–3900.

Demtröder W. (1996) *Laser Spectroscopy: Basic concepts and instrumentation*, 2nd edn, Springer-Verlag, Berlin Heidelberg.

Dick, J. (2013) 'Inside today's hot products: what teardowns can reveal', Chipworks, *The Content Strategy Conference (ConFab) 2013*.

Durini D. (2009) 'Solid-state imaging in standard CMOS processes', Ph.D. Thesis, University Duisburg-Essen, Germany.

Durini D. and Hosticka B. (2007) "Photodetector structures for standard CMOS imaging applications', *Proc. IEEE PRIME 2007*, Bordeaux, France, pp. 193–196.

EMVA 1288 (2013) http://www.emva.org/cms/upload/Standards/Stadard_1288/EMVA1288-3.1rc.pdf

Fossum E. R. (1995) 'CMOS image sensors: electronic camera-on-a-chip', *IEEE IEDM Tech. Digest*, 17–25.

Fossum E. (1997) 'CMOS image sensors: electronic camera-on-a-chip', *IEEE Trans. Electron Devices*, 44(10), 1689–1698.

Fry P., Noble P. and Rycroft R. (1970) 'Fixed pattern noise in photomatrices', *IEEE J. Solid-State Circuits*, SC-5, 250–254.

Green M. A. and Keevers M. (1995), 'Optical properties of instrinsic silicon at $300 \, K$', *Progress in Photovoltaics*, 3(3), 189–192.

Guidash R. M. *et al.* (1997) 'A 0.6 μm CMOS pinned photo diode color imager technology', *IEDM Technical Digest*, 927–929.

Haitz R. H. (1965) 'Mechanisms contributing to the noise pulse rate of avalanche diodes,' *J. Appl. Phys.*, 36(10), 3123–3131.

Hecht E. (1989) *Optik*, translated by F. Siemsen (original title: *Optics*), Addison-Wesley, Hallbergmoos.

Horton J., Mazza R. and Dym H. (1964) 'The scanistor – a solid-state image scanner', *Proc. IEEE*, 52, 1513–1528.

Hosticka B. J. (2007) 'Analog circuits for sensors', *Proc. ESSCIRC / ESSDERC 2007 in Munich, Germany*, pp. 97–102.

Hynecek J. (1992) 'CCM – A new low-noise charge carrier multiplier suitable for detection of charge in small pixel CCD image sensors', *IEEE Trans. on Electron Devices*, 39(8), 1972–1975.

Intel Corporation (2013) *Laptop, Desktop, Server and Embedded Processor Technology – Intel* (online), URL: http://www.intel.com

Janesick J. R. (2001) *Scientific Charge-Coupled Devices*, SPIE Press, USA.

Janesick J. R. (2007) *Photon Transfer*, SPIE Press, USA.

Jerram P., Pool P., Bell R., Burt D., Bowring S., Spencer S., Hazelwood M., Moody I., Catlett N. and Heyes P. (2001), in 'The LLCCD. Low light imaging without the need for an intensifier', *Proc. SPIE*, 178.

Johnson J. B. (1928) 'Thermal agitation of electricity in conductors', *Physical Review of the American Physical Society*, 32, 97–109.

Karami M. A., Gersbach M. and Charbon E. (2012) 'A new single-photon avalanche diode in 90 nm standard CMOS technology', *Proc. SPIE*, 7780.

Kittel C. (1996) *Introduction to Solid State Physics*, 7th edn, John Wiley & Sons Inc., New York, p. 673.

Kosonocky E. F., Carnes J. E. (1971) 'Charge-coupled digital circuits', *IEEE Journal of Solid-State Circuits*, SC-6(5), 314–322.

Lai J. and Nathan A. (2005) 'Reset and partition noise in active pixel image sensors', *IEEE Trans. on Electron. Devices*, 52(10), 1–4.

Leach W. M. Jr. (1994) 'Fundamentals of low-noise analog design', *Proc. of the IEEE*, 82(10), 1514–1538.

Lee P., Gee R., Guidash M., Lee T. and Fossum E. R. (1995) 'An active pixel sensor fabricated using CMOS/CCD process technology', *Proc. 1995 IEEE Workshop on CCDs and Advanced Image Sensors*, Dana Point, CA, 20–22 April.

Lilienfeld J. E. (1928) 'Device for controlling electric current', *US Patent No. 1,900,018*, filed March 1928, patented 7 March 1933.

Magnan P. (2003) 'Detection of visible photons in CCD and CMOS: a comparative view', *Nuclear Instruments and Methods in Physics Research A* 504, 199–212.

McKelvey J. P. (1982) *Solid State and Semiconductor Physics*, Krieger Publishing Co., Malabar, p. 512.

Mendis S., Kemeny S. and Fossum E. R. (1993) 'A 128 × 128 CMOS active pixel image sensor for highly integrated imaging systems', *IEEE IEDM Tech. Digest*, 583–586.

Moore G. E. (1965) 'Cramming more components onto integrated circuits', in: *Electronics 38.8 (1965)*, URL: ftp://download.intel.com/museum/Moores_Law/Articles-Press_Releases/Gordon_Moore_1965_Article.pdf.

Nishina Memorial Lectures (2008), Springer Japan, pp. 99–114.

Nixon R. H., Kemeny S. E., Staller C. O. and Fossum E. R. (1995) '128 × 128 CMOS photodiode-type active pixel sensor with on-chip timing, control and signal chain electronics', *Charge-Coupled Devices and Solid-State Optical Sensors V, Proc. SPIE*, Vol. 2415, pp. 117–123.

Noble P. (1968) 'Self-scanned silicon image detector arrays', *IEEE Trans. Electron. Devices*, ED-15, 202–209.

Nyquist H. (1928) 'Thermal agitation of electricity in conductors', *Physical Review of the American Physical Society*, 32, 110–113.

Pain B., Yang G., Cunningham T. and Hancock B. (2000) 'Photodiode based CMOS APSs with zero lag, low noise, high-linearity design', *NASA Tech. Briefs*, ABP.

Rossi L., Fischer P., Rohe T. and Wermes N. (2006) 'Pixel detectors from fundamentals to applications', Springer Verlag, Berlin and Heidelberg.

Scholze F. *et al*. (2000) 'Determination of the electron-hole pair creation energy for semiconductors from the spectral responsivity of photodiodes', *Nucl. Instr. And Meth. In Phys. Res.* A, 439, 208–215.

Schroder D. K. (1998) '*Semiconductor material and device characterization*', 2nd edn, John Wiley & Sons, Inc., USA.

Schuster M. A. and Strull G. (1966) 'A monolithic mosaic of photon sensors for solid-state monolithic applications', *IEEE Trans. Electron. Devices*, ED-13, 907–912.

Singh J. (2003) *Electronic and Optoelectronic Properties of Semiconductor Structures*, Cambridge University Press, Cambridge, p. 532.

Spieler H. (1998) 'Introduction to radiation detectors and electronics', Lecture Notes – Physics 198, Spring Semester 1998, UC Berkeley.

Sze S. M. (2002) '*Semiconductor Devices. Physics and Technology*', 2nd edn, John Wiley & Sons, Inc., USA.

Teranishi N., Kohono A., Ishihara Y., Oda E. and Arai K. (1982) 'No image lag photodiode structure in the interline CCD image sensor', *Proc. IEEE Electron Device Meeting (IEDM)*, pp. 324–327.

The ATLAS Collaboration *et al*. (2008) 'The ATLAS Experiment at the CERN Large Hadron Collider', in: JINST 3 S08003.

Theuwissen A. J. P. (1996) '*Solid-State Imaging with Charge-Coupled Devices*, Kluwer Academic Publishers, Alphen van den Rijn.

Theuwissen A. J. P. (2008) 'CMOS image sensors: state-of-the-art', *Solid-State Electronics*, 52, 1401–1406.

Tompsett M. F., Amelio G. F. and Smith G. E. (1970) 'Charge coupled 8-bit shift register', *Appl. Phys. Letters*, 17, 111–115.

Wadsworth M. and McGrath R. D. (1984) 'Charge coupled devices for X-ray imaging', *IEDM Proceedings*, pp. 20–24.

Walden R., Krambeck R., Strain R., Mckenna J., Schryer N. and Smith G. (1970) 'The buried channel charge coupled device', *Bell Syst. Tech. J.*, 51, 1635–1640.

Wang X. (2008) 'Noise in sub-micron CMOS image sensors', Ph.D. Thesis, University of Delft, Delft.

Weckler G. P. (1967) 'Operation of p-n junction photodetectors in a photon flux integration mode', *IEEE J. Solid-State Circuits*, SC-2, 65–73.

Wermes N., Hallewel G. (1998) 'ATLAS pixel detector: technical design report', CERN, Geneva.

Wolfe C., Holonyak N. Jr and Stillman G. (1989) *Physical properties of Semiconductors*, Prentice-Hall Inc., NJ.

Wong H. S. (1996) 'Technology and device scaling considerations for CMOS imagers', *IEEE Trans. Electron Devices*, 43(12), 2131–2142.

Zappa F., Tisa S., Tosi A. and Cova S. (2007) 'Principles and features of single-photon avalanche diode arrays', *Sens. Actuators A, Phys.*, 140(1), 103–112.

Zhong Z. W. (2009) 'Fine and ultra-fine pitch wire bonding: challenges and solutions', *Microelectronics International*, 26(2), 10–18.

3
Charge coupled device (CCD) image sensors

M. L E S S E R, University of Arizona, USA

Abstract: Charge coupled devices (CCDs) have been the most common high performance imaging detector for nearly all scientific and industrial imaging applications since the 1980s. In this chapter the CCD imager and its use for high performance imaging are described. The focus is on scientific and industrial applications which require quantitative measurements of the incident scene, especially in low light level conditions and when spectral information is also required. The architecture and operation of CCDs as well as their quantitative characterization and most important operational parameters are also described.

Key words: detectors, charge coupled devices, scientific imaging, sensors.

3.1   Introduction

Charge coupled devices (CCDs) have been the most common high performance imaging detector for nearly all scientific and industrial imaging applications since the 1980s. While modern silicon processing techniques have allowed amazing advances in complementary metal-oxide-semiconductor (CMOS) imager capabilities, the use of CCD imagers is still a major component of the detector market, especially for high end applications. In this chapter the CCD imager and its use for high performance imaging are described. The focus is on scientific and industrial applications which require quantitative measurements of the incident scene, especially in low light level conditions and when spectral information is also required.

We first describe the various types of CCDs as differentiated by their pixel architecture. We describe charge transport methods (clocking) and briefly touch on the conversion of electrons into an output voltage which can be sampled and recorded by the camera electronics. We next describe the frontside vs. backside illumination modes and the advantages of each. For backside illuminated detectors we further discuss fabrication, quantum efficiency, backside charging, and antireflection coatings. We then discuss the major parameters of CCD performance along with the characterization of each. We have chosen to combine the description with characterization method because the evaluation of these parameters is of special concern for high performance imaging systems and must be well

understood when selecting devices for specific imaging applications. Finally we conclude with a discussion of future trends and their impact on high end CCDs.

3.2 Charge coupled device (CCD) design, architecture and operation

3.2.1 Basic photon detection for imaging

Silicon detectors are sensitive to photons of wavelengths where electrons can be created via the photoelectric effect. This occurs when the photon energy E_λ of wavelength λ_{cutoff} is greater than the band gap of silicon in order to excite an electron from the silicon valence band to the conduction band:

$$E_\lambda = \frac{hc}{\lambda} > E_{bandgap} \tag{3.1}$$

where h is Planck's constant and c is the speed of light. For silicon, $E_{bandgap} = 1.1\,eV$, so $\lambda_{cutoff} = \frac{hc}{1.1eV} = 1.12\,\mu m$. Silicon detectors are therefore excellent detectors for nearly all wavelengths shortward of about $1\,\mu m$, in the near-IR. See Janesick (2001) for an excellent description of all aspects of scientific CCD detectors. Dereniak and Crowe (1984) also provide an excellent discussion of basic optical radiation detection.

3.2.2 Design, layout and architecture

A CCD is an imaging detector which consists of an array of pixels that produce potential wells from applied clock signals to store and transport charge packets. For most CCDs these charge packets are made up of electrons which are generated by the photoelectric effect from incident photons or from an internal dark signal. Gate structures on the silicon surface define these pixels. A time variable voltage sequence is applied to these gates in a specific pattern which physically shifts the charge to an output amplifier which acts as a charge to voltage converter. External electronics (and often a computer) convert the output sequence of voltages into a two dimensional digital image.

Pixels are composed of phases which each have an electrical connection to the externally applied voltage sequence. Each phase acts much like a metal oxide semiconductor (MOS) capacitor. The array of pixels in each direction (rows or columns) has a repeating structure of these phases in which each phase of the same name has the same applied voltage or clock signal (see Fig. 3.1 for a schematic representation of a simple CCD). There
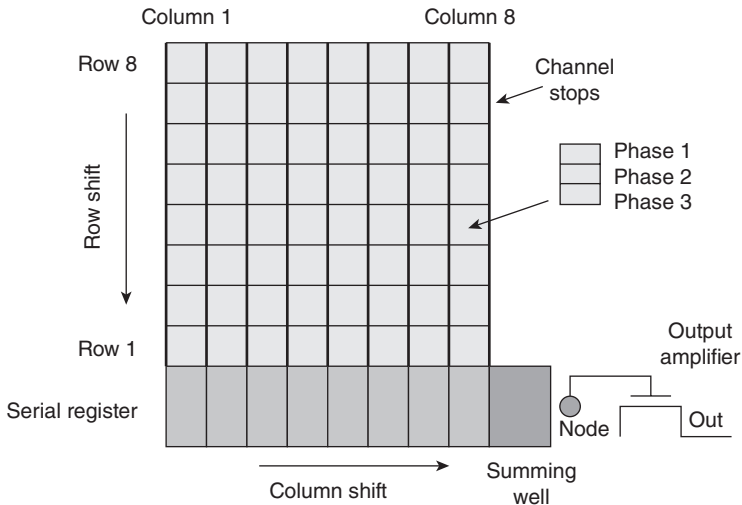
3.1 Layout of a typical 3-phase CCD.

are two, three, and four phase CCDs in fairly common use, although single phase devices also exist. As an example, a three phase device needs three different electrical connections for shifting in each direction (x/y, columns/rows, or parallel/serial), for a total of six applied clock signals. The distance from one potential minimum to the next defines the resolution of the detector, and is the pixel pitch. A three phase CCD therefore has phases spaced 1/3 of the pixel size apart. Typical CCD pixel sizes are 2–30 μm, with the smaller sizes being more common on the most recent sensors.

CCDs can be divided into two main types: linear arrays and area arrays. Linear arrays consist of one (or a few) columns of light sensitive pixels while area arrays consist of a 2D array of light sensitive pixels. Linear arrays are typically less expensive and read very fast. Area arrays are much more common for scientific and high end imaging and will be the focus of the remainder of this chapter (see Fig. 3.2). However, most properties of area arrays apply directly to linear arrays (see Theuwissen, 1995).

A full frame CCD uses the entire area (all pixels) to collect light. This is the optimal use of silicon area and the most common scientific detector. These detectors require a camera shutter to close during readout so that electrons are not generated when charge is transferred which would result in image streaking.

A frame store CCD has half of the pixels covered with opaque mask (frame store area) and half of the pixels open to incident light (image store) which collect photons during integration. This allows a very rapid shift (10^{-6}–10^{-4} sec) from image store to frame store after integration. If the shift
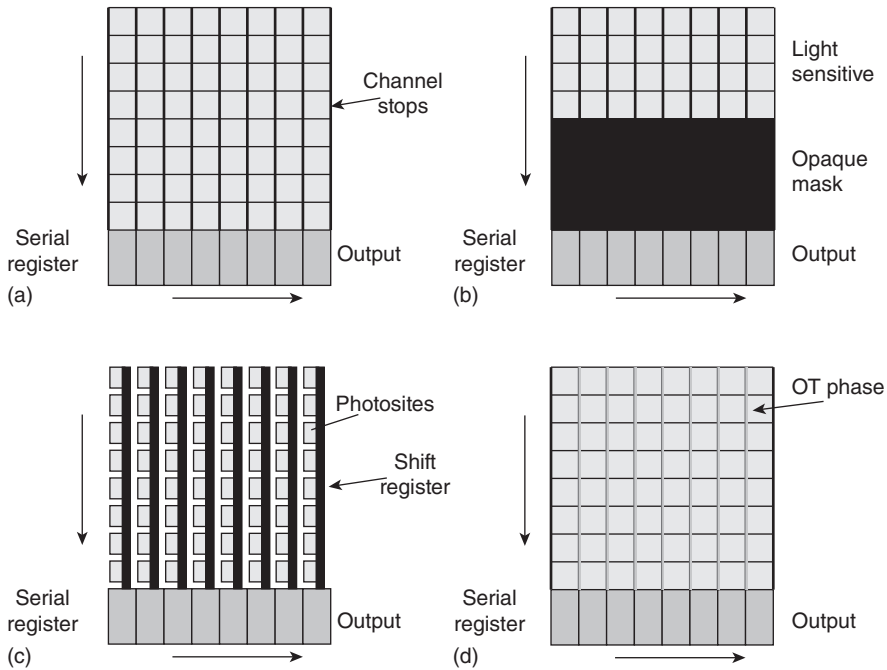
3.2 Common area array CCD architectures. (a) full frame CCD, (b) frame transfer CCD, (c) interline transfer CCD, (d) orthogonal transfer CCD.

is fast enough and the incident light is not too bright, there will be no image streaking and no external shutter is required. For frame transfer devices the image and frame parallel clocks must be driven separately, requiring slightly more complicated cameras than full frame devices. Integration can begin for the next image frame as soon as the frame transfer is complete and during the time the frame store section is reading out. This allows a faster frame rate than full frame devices with a slower pixel readout rate which usually provides lower overall noise. The main disadvantage of frame store devices is that only half of the silicon area is light sensitive.

Interline transfer CCDs have an opaque shift register alongside each column of photosensitive pixels so that a very rapid transfer can be made from the pixels to the shift register. The pixels in the shift register can then be read out while the next image frame is being exposed. This allows for a faster frame time and no external shutter is required. Interline transfer devices are more complex (in terms of fabrication) than the other types but are commonly used in high speed imaging and many commercial applications such as television. They are seldom used for high end scientific imaging.

An orthogonal transfer CCD (OTCCD) has its channel stops replaced with an actively clocked phase so charge shifting in both directions (along

rows and columns) may be achieved. If centroiding of a moving object in the scene is performed with another detector, the feedback can be used to clock the OTCCD in any direction to minimize image blurring. This is a useful function especially when making astronomical observations in which atmosphere motion (scintillation) blurs images. OTCCDs are therefore most useful for high resolution imaging, eliminating the need for tip/tilt mirrors and their associated optical losses which are more typically used to redirect the optical beam based on a feedback signal.

The orthogonal transfer array (OTA) is a monolithic device composed of (nearly) independent cells which are each an OTCCD (Burke *et al.*, 2004). The advantage of the OTA over the OTCCD is that the same detector can both provide the feedback signal and perform the data observation. When observing astronomical objects the Earth's atmosphere causes focused images to jump around, a phenomenon known as scintillation. By utilizing active optics in a telescope system, a movable steering mirror maintains the image centroid on the same detector pixels, resulting in a higher signal to noise ratio. OTA detectors eliminate the need for this steerable mirror by moving the electronic charge centroid from the image on the detector in the same manner as optically nearby guide stars are measured to move. OTAs have on-chip logic to address the OTCCD cells so that each cell can have independent timing. This allows some cells to be reading out while others are integrating. The integrating cells can be shifting small amounts in *x* and *y* based on the feedback signal obtained from the cells being read out at a higher frame rate. A common observing mode is therefore to read a few cells at high speed (many Hertz frame rate) and measure the centroid of guide stars. These centroids are then used to provide a feedback signal to shift the integrating cells which are observing objects of scientific interest. OTCCDs and OTAs were developed by Burke and Tonry at MIT/LL and the University of Hawaii (Tonry *et al.*, 1997).

3.2.3   Charge shifting and clocking

This section describes the basic operation of a CCD for making an exposure. The voltage sense described in this section applies to the most common type of CCD, n-channel CCDs which consist of p-type silicon and an n-type buried channel under the pixel gates. Fig. 3.3 shows a typically shifting sequence for a three phase CCD.

Integration during exposure

During integration (or light collection) the potential minimums are defined to collect electrons when a positive voltage is applied to one or two phases. The adjacent phases must be more negative to create a barrier to charge
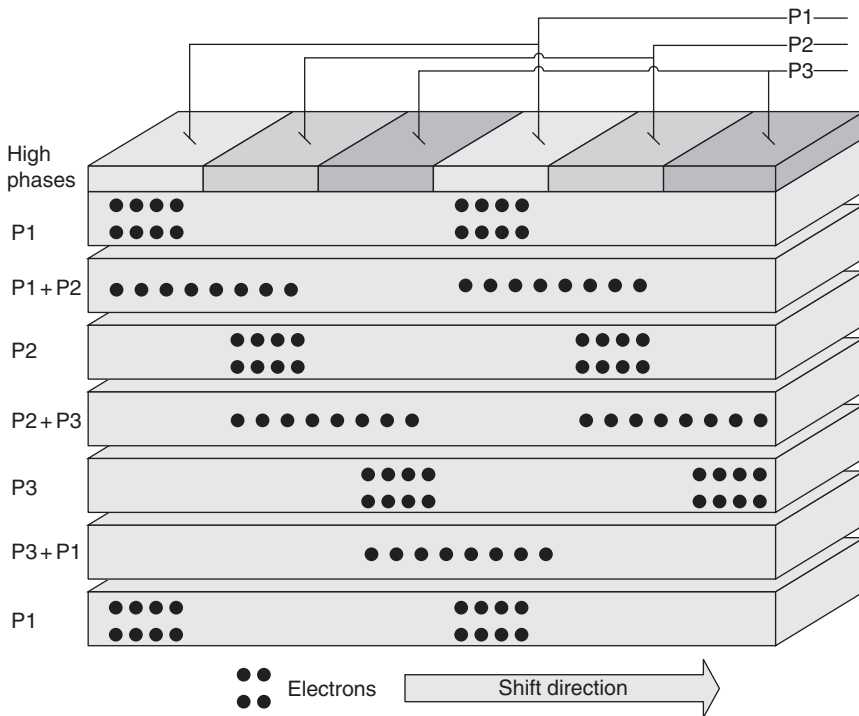
3.3 The three phase CCD shifting process. This schematic shows the shifting pattern of high voltages which moves electrons through a three phase shift register. The high (positive) voltages applied to each phase create potential minima in the silicon where electrons are confined.

spreading or image smear will occur. No shifting occurs during integration, only photoelectrons are collected. Typically a device must be cooled if integration is more than a few seconds or self-generated dark current will fill the potential wells and photogenerated signal will be lost in the associated noise. Channel stops (along columns) are created with implants during fabrication to keep charge from spreading between adjacent columns.

Shifting along columns after integration

Charge packets collected in the potential well minima are shifted when the minima (for electrons) are moved from under one gate to under the adjacent gate. This is performed by applying a voltage sequence to the phases to shift one row at a time toward the serial register. There may be multiple image sections on a device and so some charge packages may move in different directions toward their respective output amplifier.

Shifting along serial register

There is one serial (horizontal) register for each output amplifier and it receives charge from the associated columns. All charge is transferred from the last row of an image section to the serial register one row at a time. The serial register then shifts charge out to the amplifier at its end in the same manner as used for shifting charge along columns. Serial registers may be split so that charge from one half moves in one direction to an output amplifier and charge from the other half moves toward the opposite end and amplifier. Serial registers may even have multiple taps (output amplifiers) distributed along their length for frame rate operation.

Binning

The voltage timing pattern may be changed so charge from multiple pixels is combined together during transfer to the serial register (parallel binning) or to the output amplifier (serial binning). This decreases the spatial resolution of the detector by creating larger effective pixels which in turn allows higher charge capacity and so larger dynamic range. It also allows increased read out speed (higher frame rate) since every pixel is not individually sampled at the output amplifier which takes a significant amount of time. Serial register pixels are typically made with twice the physical size of image pixels to allow twice the charge capacity. Many CCDs have an output summing well which can be independently clocked as the last pixel of the serial register to aid in serial binning.

   Binning is also called *noiseless co-addition* since summing comes before readout, when read noise is generated. For a shot-noise limited imaging with uniform exposure levels, the effective signal to noise ratio (SNR) of the image is increased with binning:

$$SNR = \left[P_H P_V S(e^-) \right]^{1/2} \tag{3.2}$$

where $S(e^-)$ is the average unbinned signal in electrons per pixel and P_H and P_V are horizontal and vertical binning factors, respectively. Many cameras can be configured to vary binning in real time to optimize performance under various imaging conditions.

3.2.4   Charge sensing and amplifiers

The photogenerated electrons are shifted off the end of the serial to the output amplifier where their charge is sensed and an output voltage is produced (Fig. 3.4). The charge to voltage conversion occurs because of the capacitance of the output node according to the equation:

$$V = Nq/C \tag{3.3}$$
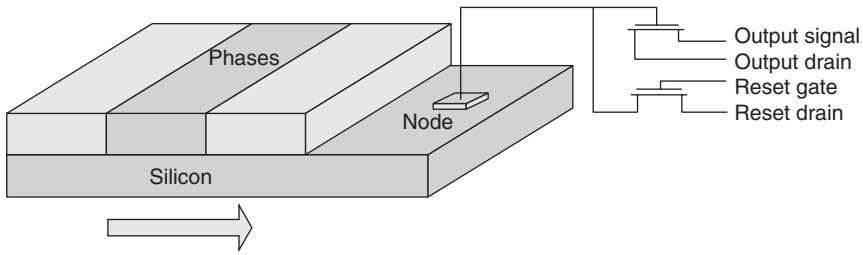
3.4 CCD output schematic. The reset transistor returns the node to a base voltage after each pixel is sampled. The output voltage is measured and digitized by the camera controller.

where C is the capacitance of the node (typically of order 10^{-13} F), N is the number of electrons on the node, and q is the electronic charge (1.6×10^{-19} C). Typically, a single electron produces from 1 to $50\,\mu$V at output depending on C. This voltage is buffered by the output amplifier (usually a metal-oxide-semiconductor field-effect transistor (MOSFET)) to create a measurable voltage across a load resistor located off the CCD as part of the camera controller. This output voltage is easily amplified in the controller and converted to a digital signal by an analog-to-digital converter to form a digital image. The node must be *reset* before each pixel is read so that charge does not accumulate from all pixels. This is accomplished with a separate on chip MOSFET called the reset transistor.

Some detectors such as electron-multiplying CCDs have internal gain which is developed in an extended serial register with a very high electric field within each extended pixel (Jerram *et al.*, 2001, Hynecek 2001). As the CCD shifts charge through this extended register, a small avalanche gain (1.01) is achieved. After ~100 gain stages, an electron packet larger than the read noise is generated and photon counting in low light level conditions is possible. There is noise associated with the gain process and so the expected signal to noise ratio must be carefully understood. But for some low light level applications with specific frame rate requirements these internal gain sensors may be the detectors of choice.

3.3　Illumination modes

CCDs are often categorized by being either front illuminated or back illuminated devices. This term indicates whether light from the imaged scene is incident on the front side of the detector (where the pixel structures and amplifiers are located) or the opposite or back side (Fig. 3.5).

Incident light

Pixel structures

Silicon

Electrons collected
in potential wells

(a)

AR coating

(b)
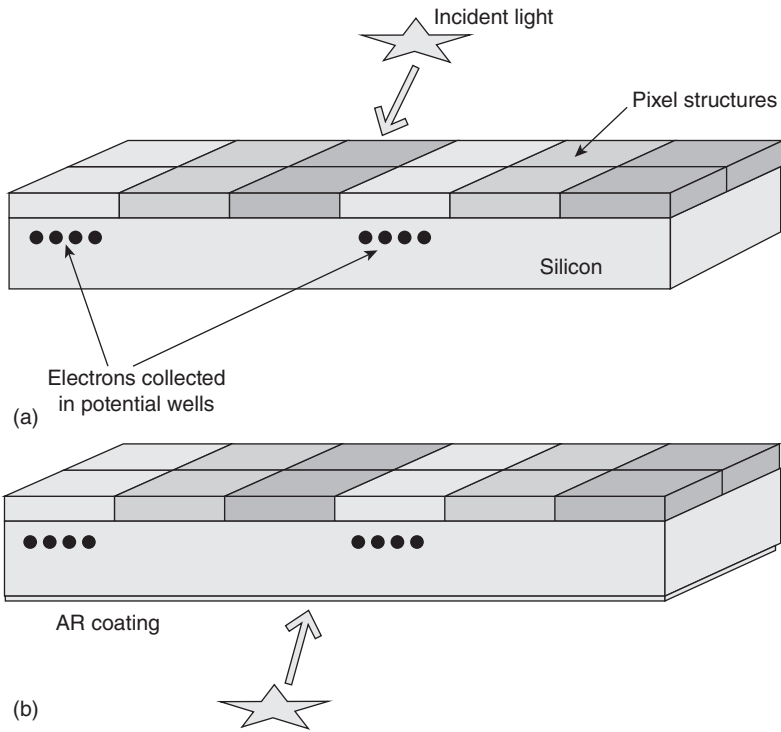
3.5 Detector illumination modes. A frontside detector is illuminated through the structures which define the pixels (a). A backside detector is illuminated directly into the silicon backside resulting in high quantum efficiency, especially with an added antireflection coating (b).

3.3.1   Frontside illumination

Front illuminated CCDs have photons incident on the gate structure or front side. They are considerably less expensive as they are used directly because they are fabricated from the silicon wafer (no additional processing steps). However, the front side gates absorb almost all blue and UV light and so they are not directly useful for imaging at wavelengths $\lambda < 400\,nm$. In addition, the physical gate structure causes many reflections, including complex quantum efficiency (QE) variations with wavelength due to interference between process layers (oxides/nitrides and polysilicon). There are several techniques to improve front illuminated device performance, such as indium tin oxide and thin polysilicon gates which are more transparent as well as open phases which have less frontside structures over photosensitive areas.

It is possible to increase front illuminated CCD efficiency by using microlenses to focus incident light on the most transparent part of each pixel.

These are microlenses typically made by applying photoresist to the frontside device surface, etching, and thermal processing to induce a lens shape. Microlenses do not usually have very good UV response but work well in the visible spectral region and are considerably less expensive to apply than the backside processing needed to create a back illuminated detector.

It is also possible to apply scintillators or other wavelength conversion coatings to the CCD frontside so that incident photons of higher energy are absorbed, causing re-emission of longer wavelength visible light which is detected with higher efficiency. This method is commonly used to make X-ray and UV sensitive frontside CCDs. The coatings can be applied directly to the imaging surface or even to a fiber optic boule which is optically bonded to the detector's front surface.

3.3.2   Backside illumination

Back illuminated devices require additional post fabricated steps (sometimes called thinning) which adds considerable cost. However, back illuminated devices are much more efficient in detecting light and usually have sensitivity throughout a very broad spectral range. For many high performance imaging applications, back illuminated CCDs are the detector of choice even though they are more expensive than their front illuminated counterparts. QE is limited only by reflection at the back surface and the ability of silicon to absorb photons which is a function of device thickness and operating temperature. An antireflection coated back illuminated CCD may have a peak QE > 98% in the visible.

Optical absorption and multiple reflections from frontside structures is avoided with backside devices although there is still interference fringing due to multiple reflections within the thin CCDs. This interference fringing is often worse for backside devices than it is for frontside devices. In recent years many back illuminated devices have been made much thicker than previously possible to increase absorption in the red and to decrease the amplitude of fringing.

Backside thinning

A back illuminated CCD must be relatively thin in order for photo generated electrons to diffuse to the frontside pixel wells and be collected under the sites where they were created. Most CCDs are built on epitaxial silicon with thickness layers which are 10–20 microns thick. The device must be thinned down to this thickness range so that the photogenerated electrons are absorbed and detected in this high quality epitaxial layer. This thinning process is difficult and expensive, leading to the much higher cost of backside sensors (Lesser, 1990).

Backside charging

After a CCD is thinned it requires an additional step to eliminate what is known as the backside potential well which will trap photogenerated electrons and cause an uncharged device to have lower QE than a front illuminated device. This backside well is caused by positive charge at the freshly thinned surface where the silicon crystal lattice has been disrupted and therefore has dangling bonds. The backside native silicon oxide also contains positive charge which adds to the backside well. This positive charge traps the electrons at the backside so that they are not detected in the frontside potential wells. Adding a negative charge to the back surface is called backside charging and leads to very high QE, especially when combined with antireflection (AR) coatings. Several different techniques have been used to produce high QE with backside devices, depending on manufacturing preferences. They can be divided into two classes, surface charging and internal charging. Surface charging includes chemisorption charging (Lesser, 2000; Lesser and Iyer, 1998), flash gates, and UV flooding (Janesick, 2001, Leach and Lesser, 1987). Internal charging includes implant/ annealing (doping) and molecular beam epitaxy (Nikzad *et al.*, 1994) and is more commonly used as it can be performed with standard wafer processing equipment (see Fig. 3.6 for a schematic view of backside charging).

AR coatings

When light is incident on the CCD backside some fraction of it reflects off the surface. This reflectance can be reduced with the application of AR coatings. An AR coating is a thin film stack of materials applied to the detector surface to decrease reflectivity. Coating materials should have
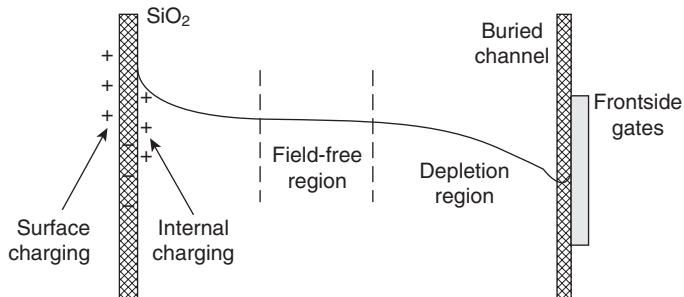
3.6 Backside charging schematic. Positive charge is added near the backside surface to create an electric field which drives electrons to the front side for detection. With no positive charge the native negative charges at the back surface will create a backside potential well which traps electrons.

proper indices and be non-absorbing in the spectral region of interest (Lesser 1987, 1993). With absorbing substrates which have indices with strong wavelength dependence (like silicon), thin film modeling programs are required to calculate reflectivity. The designer must consider average over incoming beam (*f*/ratio) and angle of incidence due to angular dependence of reflectivity.

Charge diffusion

For some devices the depletion region where electrons are swept to the potential well minima does not extend throughout the entire device. The region in a back illuminated CCD between the edge of the depletion region and the back surface is called the field-free region. Photogenerated electrons can diffuse in all directions in this region, reducing resolution through charge spreading. The diameter of the 'diffusion cloud' is found experimentally to be:

$$C_{\text{ff}} = 2x_{\text{ff}}\left(1 - \frac{L}{x_{\text{ff}}}\right)^{1/2} \qquad\qquad [3.4]$$

where x_{ff} is the field free thickness, and L is the distance from the backside surface to where the photoelectron is generated. It is important to minimize x_{ff} as much as possible or the modulation transfer function (MTF) is greatly reduced. This can be accomplished by increasing the resistivity of the silicon so that the depletion edge extends deeper into the device and by thinning the device as much as possible. Thick CCDs designed to have extended red response should always be fabricated on high resistivity silicon, typically > 1000 Ω-cm, and often have an applied internal electric field. Ideally x_{ff} is near zero and MTF is then determined mainly by the pixel pitch.

3.4   Imaging parameters and their characterization

In this section we describe the most common CCD parameters which affect imaging performance as well as the techniques used to characterize these parameters.

3.4.1   QE

QE is the measure of efficiency in which a CCD detects light. It is one of the most fundamental parameters of image sensor technology and provides the quantitative basis for selecting a frontside or backside device.

The absorptive quantum efficiency QE_λ is the fraction of incident photons which is absorbed in the detector and is given by:

$$\mathrm{QE}_\lambda = \frac{N_{\mathrm{abs}}}{N_{\mathrm{inc}}} = (1 - R_\lambda)\left[e^{\frac{-\alpha_\lambda}{t}} \right]$$ [3.5]

where R_λ is the reflectivity of the detector's incident surface, N_{inc} is the number of photons incident of the detector surface, N_{abs} is the number of photons absorbed in the detector, α_λ is the wavelength dependent absorption length, and t is the device (silicon) thickness. It can be seen from this equation that QE may be increased by:

- reducing surface reflectivity (reduce R_λ with antireflection coatings)
- increasing the thickness of absorbing material (increase t)
- increasing the absorption coefficient (decrease α_λ by material optimization)

Because nearly all CCDs are made with standard silicon semiconductor processing techniques, only options 1 and 2 are viable. AR coatings are commonly applied to nearly all backside scientific and industrial imaging CCDs (see Fig. 3.7 for some typical backside QE curves). Such coatings are discussed below but are not generally useful for frontside devices because of the complex optical and mechanical structures of polysilicon and glass passivation on the top surface of CCDs.

Back illuminated CCDs are in fact being fabricated considerably thicker than they were just a few years ago. This is to increase the QE especially in
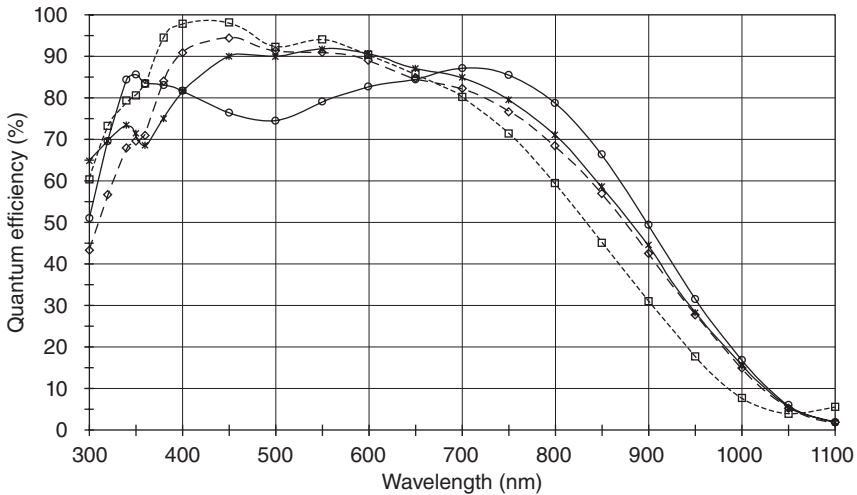
3.7 Typical backside quantum efficiency curves. The difference in QE is due to the application of different antireflection coatings to each detector as well as device thickness.

the red where the absorption length in silicon is longer. Modern silicon manufacturing allows higher quality devices to be fabricated with thick silicon due to the decrease in internal crystal defects which cause bad pixels and other imaging defects. An entire class of very thick detectors now exist which are still back illuminated but nearly the thickness of a standard silicon wafer. Pioneering work at the University of California's Lawrence Berkeley Laboratory has focused on the development of thick, high resistivity, back illuminated p-channel CCDs for high performance imaging (Holland *et al.*,2003).

Quantum yield

Related to QE is quantum yield (QY) which is the term applied to the phenomenon that one energetic interacting photon may create multiple electrons-hole pairs through collision (impact ionization) of electrons in the conduction band. This can cause the measured QE to appear to be higher than it is, even greater than unity. In fact QE remains less than one (the probability that a photon will be absorbed) but the number of resulting electrons detected can be much greater than one for a single photon. Since photon energy increases with decreasing wavelength, QY is only important in the UV and shorter spectral regions. An example, a commonly used diagnostic tool is an 5.9 keV Fe-55 X-ray source which produces ~1620 electrons per incident photon. For photon energies > 3.1 eV, or about 0.4 μm:

$$QY = \frac{E_\lambda}{E_{e-h}} \qquad\qquad [3.6]$$

where $E_{e-h,Si} = 3.65\dfrac{eV}{e}$ is the energy needed to produce an electron–hole pair (see (Janesick, 2001) for more details).

Characterization of QE is typically performed by illuminating the detector with a known light flux. The actual flux is usually measured with a calibrated photodiode having a known responsivity. Ideally the calibration diode is placed in the same location as the CCD to measure the actual optical beam flux. In some cases reasonable results can be obtained when the calibrated diode is placed in another portion of the beam and a scaling factor is used. In this case care must be taken to account for subtle variations in the beam such as spectral differences due to surface reflections and scattered light variations. Because the photodiode and CCD will not have the same spectral response these variations can be difficult to measure and remove. It is also important to control accurately the light source intensity which will often vary with AC line power. A common technique used to reduce this error source is to provide a feedback signal from the light source (using a diode)

and the light source power supply. The feedback from this circuit is used to stabilize the actual light intensity. Commercial power supplies with such features are available although some systems rely instead on chopping the beam between a reference diode and the device under test to account for temporal variations.

A commonly used QE characterization technique for CCDs is diode mode testing. In this case the detector is not operated as a CCD but in a static mode as a single diode. Because a CCD has a buried channel under the entire pixel array it may be treated as a single pixel of the total area of all the pixels combined. Typically the reset drain of the detector connects to the buried channel and a current is generated between this signal and the device ground. This current can be directly compared with the calibrated photodiode current for a simple but accurate QE measurement. This mode works at all temperatures and so can also be used when the device is cooled.

3.4.2   Read noise

Read noise is the fundamental uncertainty in the output of the CCD. It is often the dominant noise source in a high performance imaging system. Read noise is typically measured in electrons rms, but is actually a voltage uncertainty.

Nearly all high end cameras use correlated doubled sampling (CDS) to reduce read noise by reducing the uncertainty in absolute charge level at the output node. When the output node is reset, its final value is uncertain due to kTC or thermal noise. The CDS technique eliminates this uncertainty by sampling each pixel both before and after reset. Before shifting charge from a pixel onto the node, the node is reset with the on-chip reset transistor. The node voltage is sampled and recorded. The pixel to be measured is then shifted onto the node. The node is sampled again and the difference between the two samples is the actual charge in the pixel. Low noise MOSFETs with very low capacitance nodes and amplifiers, using CDS, can produce less than 2 electrons rms read noise with only one (double) sample per pixel.

Read noise is characterized by calculating standard deviation σ of a zero or bias frame (no light or dark signal). Noise is measured in units of digital numbers (DN) from clean subsections of image and so the system gain constant K (e/DN) is required. K is measured from the photon transfer technique or by measuring absolutely calibrated events such as those generated from Fe-55 X-rays. A common technique to measure the amplifier noise only without including other noise sources such as spurious charge is to clock charge backwards or away from the output amplifier and measure the standard deviation of the output. While this may not be the same value as measured in actual images it can allow better optimization of the amplifier noise which is a function of applied voltages.

3.4.3 Photon transfer and full well

The full well capacity of a pixel is the maximum number of electrons which a pixel can hold. It is determined by the pixel size and structure, the output amplifier, and the controller electronics. Pixel capacity is a function of area, so bigger pixels (or binned pixels) usually hold more charge and therefore have higher full well capacity.

Full well is often measured by making a photon transfer curve (PTC) plotting log noise (DN) vs log signal (DN). For this plot there are several regions of interest (see Fig. 3.8 for an example PTC):

- slope = 0 at low signal is the read noise floor
- slope = ½ (linear portion) is the linear photon which is shot noise dominated
- significant deviation from slope ½ occurs at the full well.

The system gain constant K is a critical parameter of an imaging system and is often measured from the photon transfer curve. It is the conversion constant from DN to electrons. K can be determined from photon statistics by analyzing two flat field images which are used to remove fixed pattern noise (see Janesick (2007) for extensive discussion and examples of photo transfer techniques).
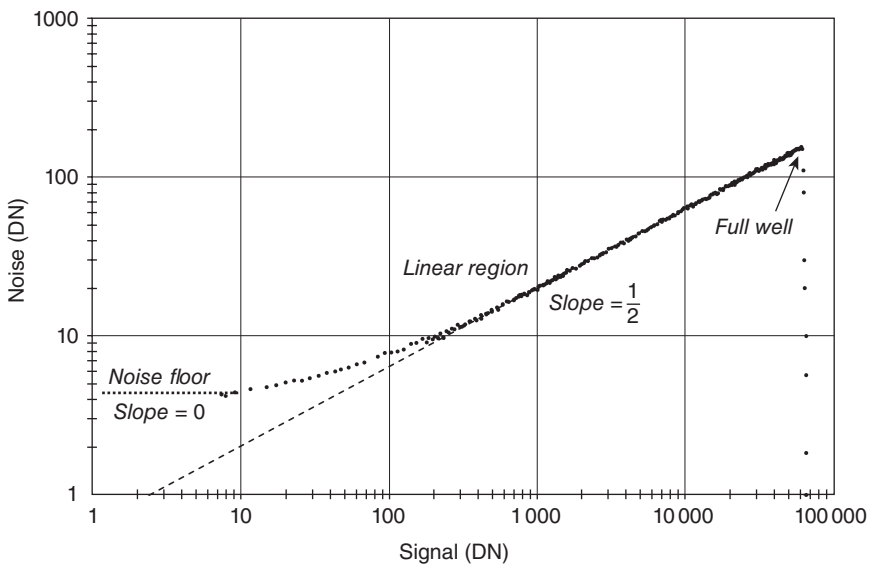
3.8 A photon transfer curve. The three regions described in the text can be seen in this PTC for a 2k × 2k pixel CCD.

3.4.4 Charge transfer efficiency

Charge transfer efficiency (CTE) is a fundamental imaging parameter for CCD systems. It is the efficiency in which charge is shifted from one pixel to the next. Good CCDs have CTE values of 0.999995 or better. CTE is often worse at lower temperatures. The charge in a pixel after N shifts is $S_N = S_i (CTE)^N$ where S_i is the initial charge in the pixel before shifting. As an example, a Fe-55 X-ray event (1620 electrons) in the far corner of a 4k × 4k device (8k shifts) will contain only about 1493 electrons when sensed at the output amplifier if CTE = 0.999990 (92%).

CTE is usually calculated by measuring the trailing edge of sharp images or the residual charge after reading the last row or column. A Fe-55 X-ray source can also be used in the lab to measure CTE quantitatively. Since each 5.9 keV X-ray produces 1620 electrons and the electron cloud produced is about $1\,\mu m$ in diameter, each event should appear as a single pixel point source. When CTE is less than one, pixels further from the readout amplifier (for horizontal CTE measurement) or farther from the serial register (for vertical CTE measurements) will be measured to have fewer electrons than closer pixels. By plotting pixel values verses location one can usually measure CTE to 1 part in 10^6. When CTE is poor a 'fat zero' or preflash, which adds a fixed amount of charge to each pixel before image exposure, may fill traps to improve CTE. There is of course a noise component associated with the added signal.

3.4.5 Linearity

A linear input signal should produce a linear output signal for a CCD. The difference between the actual and ideal output is the non-linearity of the system and can be due to silicon defects, amplifier, and/or electronic issues. Linearity is normally characterized by measuring the signal output generated as a function of increasing exposure level. Fitting a line to this data should produce residuals of less than 1% from just above the read noise floor to full well. Linearity is often measured from the photon transfer curve as well, although this is not strictly linearity as it shows variance vs signal and not output vs input signal.

3.4.6 Dark signal

Dark signal (or dark current) is due to thermal charge generation which occurs in silicon. The dark signal from a silicon pixel in electrons generated per second is given by:

$$D(e) = 2.5 \times 10^{15} A_{pix} D_{FM} T^{1.5} e^{-E_g/2kT} \qquad [3.7]$$

where D_{FM} is a silicon 'quality' parameter in nA/cm² @ 300 K, A_{pix} is the pixel area, E_g is the silicon band gap, T is temperature, and k is the Boltzman constant (Janesick, 2001).

The usual method of reducing dark current is to cool the detector. If very low light level measurements are to be made it is not uncommon to cool a CCD to −100 °C or lower. This reduces the dark signal to just a few electrons per pixel per hour. Higher light level measurements often use devices cooled with a thermoelectric cooler to −40 °C or warmer. While many commercial CCD systems operate with no cooling the dark current is so high that only limited quantitative measurements are possible since the dark signal overwhelms the incident signal.

Often the characterization of dark signal actually includes other signals such as optical glow from detector diode breakdown and dewar light leaks. Dark signal characterization is performed by taking multiple exposures and adding them together, usually with a median combine or clipping algorithm to reject cosmic rays or hot pixels. While such measurements are relatively straightforward for devices operated at room temperature or with only small amounts of cooling, many high performance detectors are cooled such that the dark signal is extremely small. Since dark signal for cooled detectors is of the same order as the device read noise, these measurements are very difficult to make accurately. Spatial variations in dark signal due to clocking artifacts and silicon processing variations can be larger than the mean dark signal values. Small light leaks also contribute to dark signal uncertainly and care should be taken to make sure the imaging system is completely dark. Even the hermetic connectors which bring signals through the dewar walls can introduce light of the same order as the actual dark signal. In addition, many materials fluoresce, especially after exposure to short wavelength radiation, and care must be taken to eliminate such additional sources of unwanted light. It is not uncommon to make dark current measurements using many hours of combined exposure time and with opaque caps over the camera window. Finally, accurate calibration of temperature sensors in the dewar as well as a good calibration of the difference between the silicon temperature and those sensors is important when characterizing absolute dark signal.

Dark noise

Dark noise is the uncertainly in dark signal and is approximately $\sqrt{N_{dark}}$. The usual practice is to reduce the device temperature until dark noise is less than read noise and so that dark signal is not the dominate noise source. This may occur at very low temperatures (−100 °C or less) for very low read noise devices which are used for long exposure time observations.

3.4.7 Fixed pattern noise

Fixed pattern noise in CCDs is due mainly to sensitivity variations from pixel to pixel. These variations may be due to QE differences (photo response non-uniformity (PRNU)), dark signal variations (dark signal non-uniformity (DSNU), or even fixed electronic variations which are synchronized with image readout. Typically images are 'flat field corrected' by dividing data images by a calibration image taken with uniform illumination, which increases noise in the resultant image by $\sqrt{2}$. DSNU and fixed noise patterns may be removed in a similar manner, by subtracting a master dark image which has a high enough signal to noise ratio to be statistically significant.

3.5   Conclusion and future trends

It is likely that the demand for CCDs will continue to decrease while the market for CMOS imagers increases. This is due mainly to the much greater number of fabrication facilities which can produce CMOS devices as compared to available CCD facilities as well as the low power and integrated image processing capabilities of CMOS devices. However, there are several areas where CCDs will likely remain dominant for many years.

First and foremost large pixel image sensors will likely remain CCDs as CMOS pixel sizes continue to shrink. Large pixels are required for most scientific and industrial imaging applications due to their larger dynamic range, although progress continues to improve the full well capacity of smaller pixels for all image sensors. Very large area scientific CCDs are in demand and continue to grow in size. Currently the largest CCD imager is a 10k × 10k pixel single die per 150 mm wafer (Zacharias *et al.*, 2007). The simplicity of CCD devices (fewer transistors and other structures compared to CMOS sensors) tends to produce higher yield at very large areas. It is certainly true that larger CMOS imagers can ultimately be produced since CMOS can be built on 300+ mm wafers, but yield is currently low and cost is very high.

Back illuminated CCDs for high QE and UV and X-ray applications dominate because backside CMOS processing is currently focused more on commercial applications which need very small pixels and therefore very thin devices for reasonable MTF. However, backside processing techniques used to make scientific CCDs are currently being applied to CMOS sensors for low noise and low light applications. These techniques will no doubt continue to increase the availability of very high performance CMOS sensors in the future. Similarly, enhanced red response requires relatively thick silicon (> 20 μm) and most CMOS processes are optimized for much thinner silicon. The low voltages used on CMOS imagers may limit the

depletion depth possible in thick sensors and so enhanced red QE detectors are likely to remain CCDs for some time.

3.6   References

Burke B, Tonry J, Cooper M, Luppino G, Jacoby H, Bredthauer R, Boggs K, Lesser M, Onaka P, Young D, Doherty P and Craig D (2004), 'The orthogonal-transfer array: a new CCD architecture for astronomy', *Proc. SPIE*, 5499, 185–192.

Dereniak E and Crowe D (1984), *Optical Radiation Detectors*, John Wiley and Sons.

Holland S, Groom D, Palaio N, Stover R, and Wei M (2003), 'Fully depleted, back-illuminated charge-coupled devices fabricated on high-resistivity silicon,' *IEEE Trans. Electron Dev.*, 50 (3), 225–238.

Hynecek (2001), 'Impactron-a new solid state image intensifier', *IEEE Trans. on Electron Devices*, 48, 2238.

Janesick J (2001), *Scientific Charge-Coupled Devices*, SPIE Press Monograph PM83.

Janesick J (2007), *Photon Transfer*, SPIE Press Monograph PM170.

Jerram P, Pool P, Bell R, Burt D, Bowring S, Spencer S, Hazelwood M, Moody I, Catlett N and Heyes P (2001), 'The LLLCCD: low light imaging without the need for an intensifier', *Proc. SPIE*, 4306, 178–186.

Leach R and Lesser M (1987), 'Improving the blue and UV response of silicon CCD detectors', *Pub. A.S.P.*, 99, 668.

Lesser M (1987), 'Antireflection coatings for silicon CCDs', *Optical Engineering*, 26, 911.

Lesser M (1990), 'CCD thinning, coating, and mounting research for astronomy', in Jacoby G, *CCDs in Astronomy*, BookCrafters, 65.

Lesser M (1993), 'Backside coatings for back illuminated CCDs', *Proc. SPIE*, 1900, 219.

Lesser M (2000), 'CCD backside coatings optimized for 200–300 nm observations', *Proc. SPIE*, 4139, 8.

Lesser M and Iyer V (1998), 'Enhancing back illuminated performance of astronomical CCDs', *Proc. SPIE*, 3355, 23.

Nikzad S, Hoenk M, Grunthaner P, Grunthaner F, Terhune R, Winzenread R, Fattahi M and Hseng H-F (1994), 'Delta-doped CCDs for high QE with long-term stability in the UV and visible wavelengths', *Proc. SPIE*, 2198, 907.

Theuwissen A (1995), *Solid-State Imaging with Charge-Coupled Devices*, Kluwer Academic Publishing.

Tonry J L, Burke B E and Schechter P L (1997), 'The orthogonal transfer CCD,' *PASP*, 109, 1154–1164.

Zacharias N, Dorland B, Bredthauer R, Boggs K, Bredthauer G. and Lesser M (2007), 'Realization and application of a 111 million pixel backside-illuminated detector and camera', *Proc. SPIE*, 6690, 8.

4

Backside illuminated (BSI) complementary metal-oxide-semiconductor (CMOS) image sensors

A. LAHAV, Tower Semiconductor Ltd, Israel and
A. FENIGSTEIN and A. STRUM, TowerJazz, USA

Abstract: This chapter reviews modern manufacturing techniques for backside illuminated (BSI) CMOS image sensors (CIS). It presents a thorough discussion regarding the advantages and disadvantages of the front illuminated CIS throughout different CMOS fabrication nodes and introduces a historical review, state-of-the-art discussion on the technological issues as well as on the applications and performance of back illuminated imagers. The chapter discusses all the relevant state-of-the-art issues of the imagers in both, commercial applications as well as high-performance scientific and industrial niche applications.

Key words: CMOS image sensors (CIS), backside illuminated (BSI) sensors, quantum efficiency (QE) enhancement, UV sensitivity, infra–red sensitivity, machine vision, global shutter sensors.

4.1   Introduction

The first backside illuminated (BSI) sensors were introduced by the scientific community in the mid-1970s. The main driving force was their superior quantum efficiency, especially in the UV and the blue spectrums. These sensors were produced in low quantities and with poor yields. Today the innovation in image sensor technology is driven by mass market applications. BSI technology was proven useful for coupling light into pixels smaller than 2 μm commonly used in every smart phone. This chapter is devoted to a review of BSI technology and focuses on its application for the high end market.

In this section we will describe briefly the challenges for scaled down pixels in a modern CMOS image sensor chip. We will review various technologies used to overcome these challenges and show why BSI is the most natural solution for such sensors. CMOS image sensors consist of an array of light sensitive pixels. Each pixel consists of a photo diode (PD), which is the light sensitive element, and several control transistors. The PD collects and stores the photo-carriers while the control transistors are used for setting the exposure time, transforming charge to voltage and for readout

98

control sequence. The pixel array is connected to the control circuit by several metallization layers. Each metallization layer is separated by inter-dielectric material. The interface between the chip and the outside world is the passivation layer which is placed above all the metallization layers and isolates the chip from environmental hazards.

In a front-side illuminated (FSI) sensor (see Fig. 4.1(a)), the light reaches the PD active region through the passivation, metallization and inter-dielectric layers. There are several loss mechanisms associated with the coupling of light from the front side of the sensor. First, the reflection of light from the passivation layer. Second, the reflection of light from the metal control lines surrounding the PD, and third, the shift and reflectance of light coming from large angles to the sensor, especially at the edges of the sensor. This angled light travels through the thick inter-dielectric layer and it may be collected by a neighboring pixel. These three mechanisms cause reduction in the maximum available photons or in other words, quantum efficiency (QE) reduction, and contribute to large crosstalk, and as a result, reduction of the sensor's signal to noise ratio (SNR).

A BSI sensor contains the same components as a FSI sensor but the sensor's metals are located behind the photodiode (see Fig. 4.1(b)). During manufacturing the wafer is flipped upside down so the metallization and passivation layers are now located beyond the PD with respect to light. The manufacturing of such a sensor is highly complicated since it requires processing steps from both sides of the wafer. The main technological challenges associated with BSI are bonding of the wafer to a holding wafer and then thinning of the original wafer to a thickness of a few microns. In addition, there are also some issues related to the optimization of the sensor's optical and electrical performance, such as inherent large color cross talk, which degrades the image quality, and excess dark current coming from the additional interface. In the past, these problems prevented the mass market manufacturing of BSI sensors. Consequently, until recent years, BSI was in use just in niche high end markets where low-light performance could not be compromised.

4.2   Challenges facing a scaled-down frontside illuminated (FSI) sensor

The typical back-end scheme of a CMOS image sensor is shown in Fig. 4.2(a). It is very common to design a sensor that uses three layers of metal in the pixel array area, and four in the periphery (digital and analog circuitry). The passivation layer is typically composed of oxide and nitride layers and is located above the last inter-dielectric layer (dielectric four, in the case of a four level metals scheme). The red, green and blue polymeric

On chip micro-lens

Passivation (and a reflecting coating)

Chip metal routing

Interaction surface

Incoming light

Light reflection from chip metal

Red filter · Green filter · Blue filter

Inter metal dielectric

Diode active region

Si substrate

(a)

Red filter · Green filter · Blue filter

Diode active region

(b)
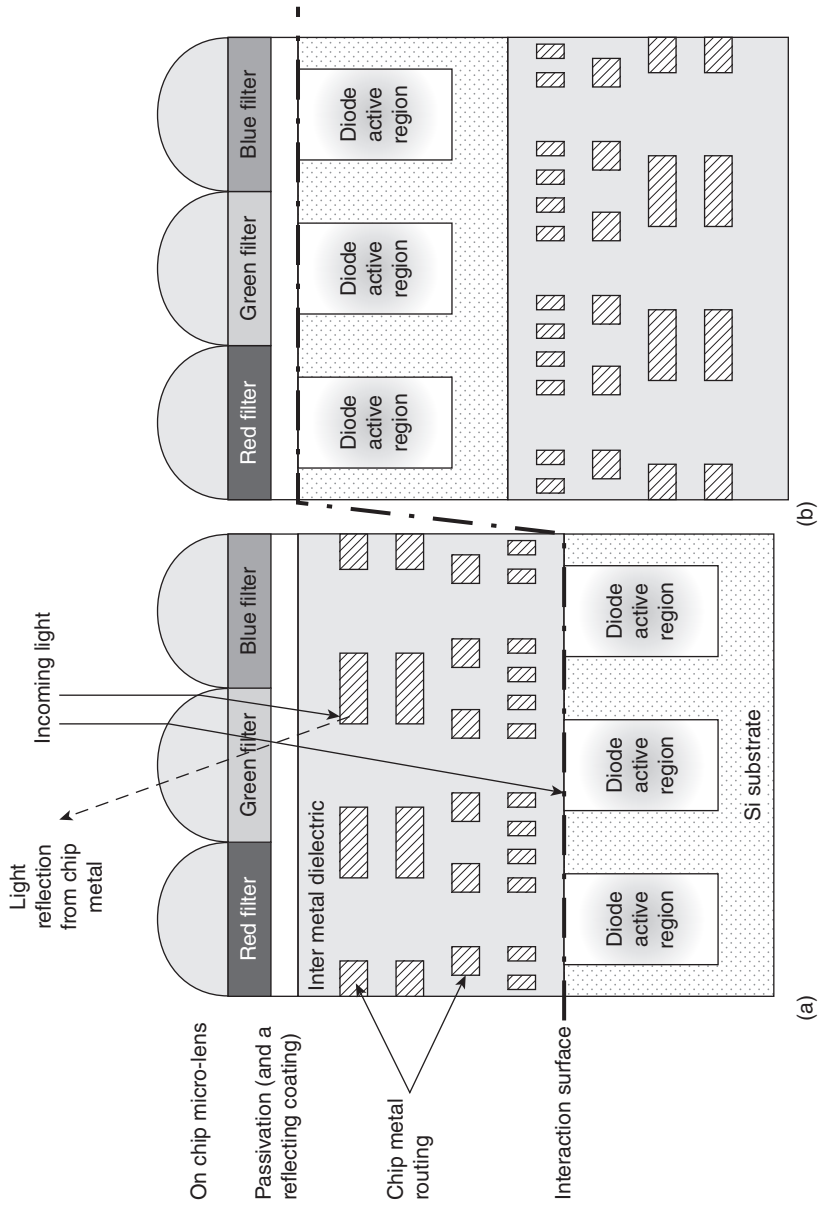
4.1 Illustration of (a) frontside vs (b) backside illuminated sensor.
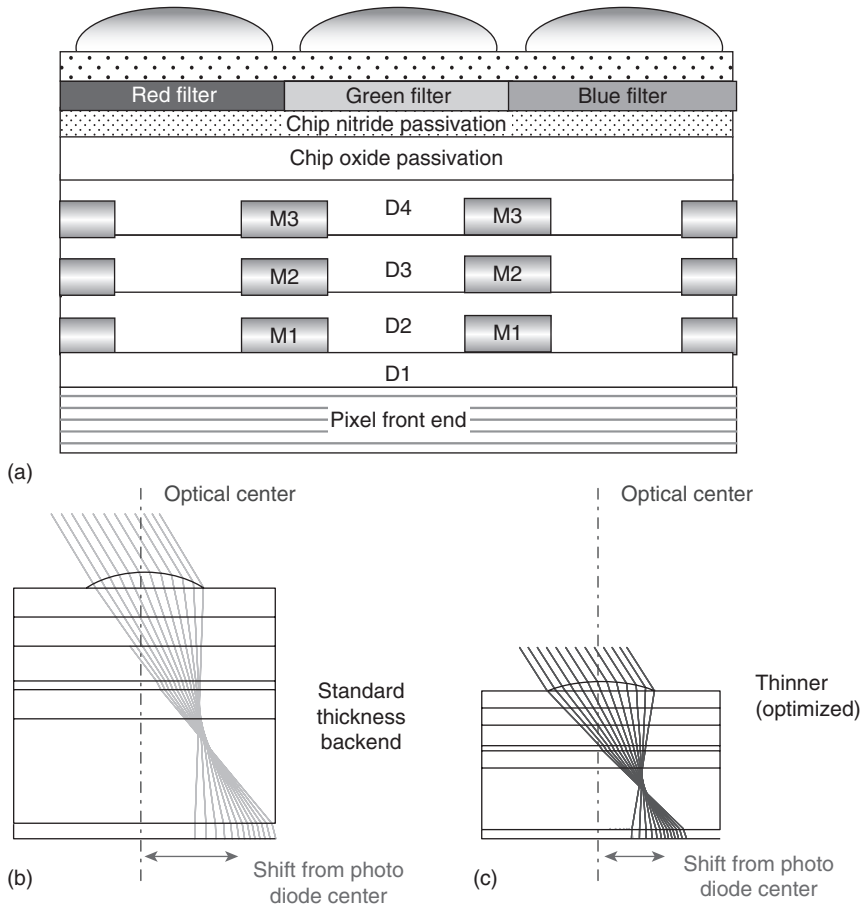
4.2 (a) Simplified schematics of a pixel metallization backend, (b) illustration of the optical power shift for angled light rays for a standard CMOS backend and (c) illustration of the optical power shift for angled light rays for an optimized CIS backend.

filters, which are responsible for the color reproduction of the images are located above the passivation layer (a single pixel without a color filter can only provide monochrome information). The micro-lens that is supposed to direct the light into the photodiode is located above the color filter. There are two fundamental problems that prevent light that hits the micro-lens from reaching the PD active region:

1. The physical size of the aperture on the first metallization that the light encounters (the last metal). For a small pixel, this aperture is rather small unless an advanced metal routing technology is used. If the

aperture is not large enough or if the micro-lens is not strong enough, the light will reflect back and will effectively be lost.

2.   The light reaches the micro-lens with a wide angle due to the aperture of the system lens. If the optical path from the bottom of the micro-lens to the diode is long, the light can easily reach a different pixel and cause a severe crosstalk problem.

The main optimization techniques traditionally used to solve these problems were improving the micro-lens performance (power) and reducing the total height of the metallization backend. The ability of the micro-lens to increase the number of photons available for detection becomes much more efficient when the micro-lens is placed closer to the active region of the PD. The difference between a pixel with thinner metallization and inter-dielectric compared with a pixel that uses standard metallization is shown in Fig. 4.2(b) and Fig. 4.2(c). The simple ray tracing illustration shows that a micro-lens that is placed over a thinner stack can focus light closer to the center of the diode than the same micro-lens placed over a thicker metallization backend.

A larger optical opening above the PD is key for achieving higher QE and better angular response. It can be achieved by using narrow control lines for the pixel. A pixel designer would use (in most cases) the narrowest line width that could be manufactured in a given set of fabrication tools. Therefore, moving from one pixel generation to the next (smaller pixel) will by definition require also moving to a smaller technology node. This approach was dominant for many pixel generations and thus, the technology node used for CIS moved from 0.5 microns down to 0.18 microns [1].

For the 0.13 um CMOS process node, the metallization is replaced by copper. The main reason is that copper allows higher current density and therefore narrower and thinner lines compared to aluminum. The image sensor industry does not use high current densities for the pixel control lines but could use the narrow and thin copper line for widening the optical opening above the PD. The copper inter-dielectric layers are also thinner and the total height of the backend can be reduced compared to an older generation of the fabrication process.

On the other hand, unlike the natural adoption of advanced aluminum process nodes for image sensor fabrication, moving to a copper backend created a major challenge for CIS. The main challenge is that during the fabrication of copper lines, there are alternate nitride and oxide layers for the inter-dielectric layers [2]. The alternating dielectrics with different permittivity cause light in a certain wavelength to be reflected just like light hitting an interference filter. Thus, in order to use a copper back end for image sensors, the nitride layers in the optical path to the silicon surface

have to be removed [3]. The immediate result is that shrinking pixel dimensions could not be done just by moving to a more advanced fabrication node, and the gap between the modern image sensor process and standard logic fabrication becomes wider.

For fabricating CIS with pixel sizes below 3 μm, creative solutions and non-standard backend modules were developed. One example is recess of the metallization stack above the pixels array [4]. This starts by designing a pixel using only two metal lines for operation, with an opening as large as possible above the diode. This task requires either a very advanced fabrication process (90 nm and below) or development of an additional module of a local interconnect metallization below the first metallization layer [5]. In this solution the array periphery is still being designed by using four levels of metallization. The inter-dielectric layer between the last metal and the second metal is then selectively etched over the pixel array area. The result is that color filters and micro-lenses are placed very close to the top metal of the pixel (second metal) instead of the top metal of the sensor (fourth metal) and this significantly increases the pixel performance, especially for angled light.

Another example is the 'light pipe' solution [6, 7] which is illustrated in Fig. 4.3. A 'light pipe' (or 'light guide') is fabricated by etching a deep via from the passivation layer down to the diode surface. After the etch process, the via is filled with spin on glass (SOG) [8] or with a special polymer having a high refractive index [9]. These three dimensional structures, if designed properly, can trap the light beams inside the light pipe thus reducing or even eliminating completely the color cross talk. There are many obstacles for successfully integrating light pipes into a reliable sensor; the most significant ones are the relatively high aspect ratio etch with uniform depth over many millions of pixels passivation of the pipe walls, and filling the pipe with high refractive index material. This process is not scalable, namely, each pixel pitch will require a different etch recipe and sometimes a different fill recipe as well. Thus, for each pixel generation, this process will require re-optimization.

In Fig. 4.4, there is a good example of optical path optimization made by Aptina. The chip was taken out of a Lenovo phone and was analyzed by Chipworks [10]. The chip was fabricated using 65 nm copper metallization. The tight metallization enables pixel routing using only two levels of metals. In addition to the recessed backend, Aptina also used light pipe technology to reduce the optical crosstalk even further. The light pipe etch serves also for removal of the nitride–oxide–nitride layers which comes as a standard at this process node.

In the previous paragraphs we showed the enormous efforts invested by different companies in coupling the light into the pixel; thinner and non-standard metallization backend processes to decrease the optical path, light

Incoming light

On chip micro-lens

Red filter   Green filter   Blue filter

Passivation (and a
reflecting coating)

Inter metal
dielectric

Chip metal
routing

Interaction surface

Diode
active
region

Diode
active
region

Diode
active
region

Si substrate
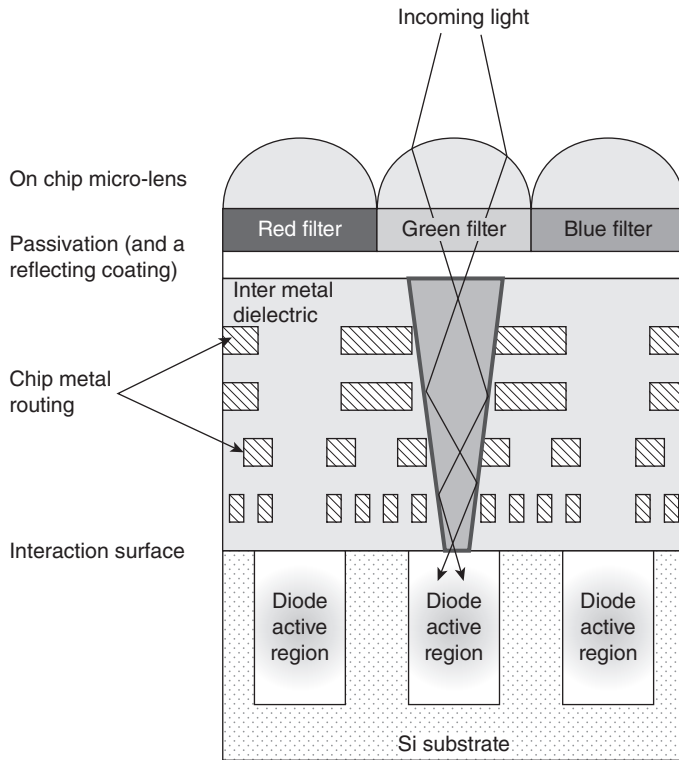
4.3 Illustration of the light pipe concept. Deep vias etched through the inter-metal dielectric layer above each diode. The vias are filled with a slightly higher refractive index than their surroundings. The u-lenses are modified to ensure good coupling of the photons into the light pipe. Due to the higher refractive index in the pipe, photons are guided from the chip top surface to the diodes.

guides and a recessed backend. These solutions needed re-optimization over and over again for each new pixel dimension.

The alternative approach is to try to couple the light from the back side of the sensor. This approach could in principle save the effort of customizing the backend process and allow using advanced CMOS fabrication technology. The other benefit from BSI is an increased QE due to optimization of an anti reflecting coating (ARC) layer. This led many market leaders to seek a wafer scale BSI process that is compatible with modern fabrication tools (unlike the BSI process which was used for scientific applications and other niche markets). In 2007, OmniVision (OVT) was the first to introduce a BSI 5 mega-pixel sensor for the cellular phone market.
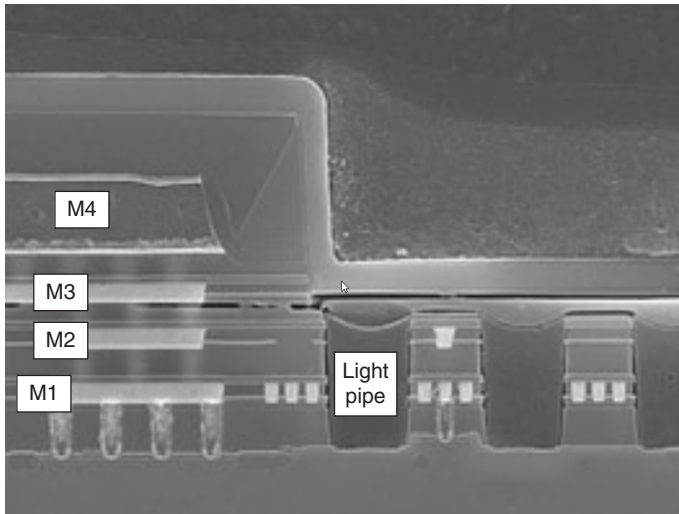
4.4 An Aptina sensor taken from a LENOVO phone (courtesy of Chipworks (10)). The sensor is fabricated using a Cu backend. There are two levels of metals in the pixel array and four levels of metals in the array periphery. In addition to the recessed backend, Aptina is using a light pipe to improve the pixel optics. The light pipe also solved the potential reflectance issue coming from the nitride–oxide–nitride interference by etching them away during the light pipe etch.

4.3 Basics of backside illuminated (BSI) sensor process integration

Typical integration scheme of a modern, wafer level BSI sensor was presented by Wuu from TSMC and Rhodes from OVT at IISW 2009 [11], [12]. Their flow is illustrated in Fig. 4.5. The flow used by others is similar in principle but each sensor manufacturer takes somewhat of a different approach to solve the same physical problem. Most of these flows were developed with small pixel dimensions in mind. But other sensors with relatively large pixels, such as the ones used for cinematography, or UV sensitive sensors for machine vision or medical applications, could also benefit from wafer level backside illumination process. These special sensors are less cost sensitive but still can benefit from modern, high yield, processes. In this section we will review the most common technique to integrate the full sensor and the common difficulties that can arise at each process step. The impact on the pixel performance will be discussed as well.
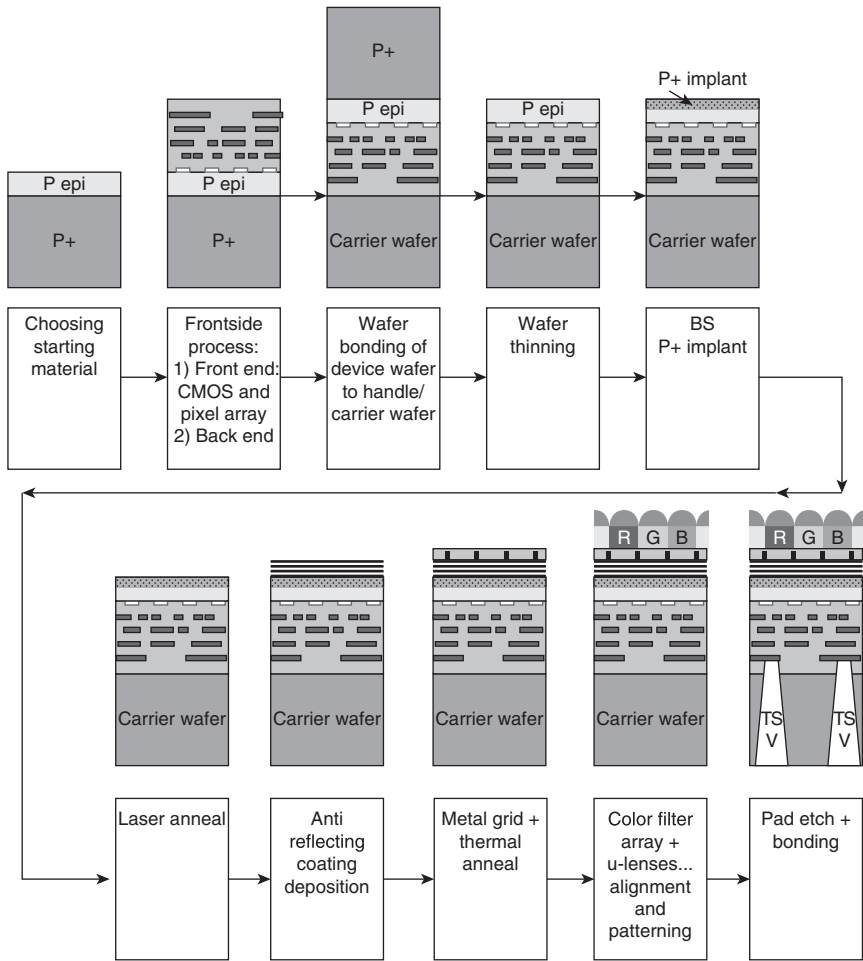
4.5 Typical BSI complete process flow (inspired by TSMC and OmniVision from IISW 2009 [11, 12]).

4.3.1   Starting material selection

The sensor fabrication process begins by starting material selection. This is a very important technological choice which will have many integration consequences in later fabrication steps. OVT and TSMC chose to start with a P on P+ doped wafer. The P on P+ doped wafer is available with a proven supply chain, relatively cheap and widely used in the CIS industry for FSI devices. The other choice is to use silicon on insulator (SOI) wafers [13] which were used, for example, by Sony in an early BSI process generation [14] and by TowerJazz for devices targeted at high end markets [15]. The

main advantage of SOI substrates over bulk material is that the buried oxide (BOX) of the SOI acts as a built-in etch stop layer as well as a protection to the active silicon layer during the backside thinning process as will be discussed in detail later. The main disadvantages are the cost and availability of such wafers, compared to bulk wafers which make it less [16] attractive for high volume production for the mobile phone market.

4.3.2   Frontside processing

Next is the frontside processing of the image sensor. This stage is very different from one sensor manufacturer to another. Each company has chosen a different technology node to fabricate the frontend and uses a different set of manufacturing tools. Nevertheless, there are many similarities which come from identical optimization targets of the BSI pixel design. One of the main physical limitations, which are inherent to a BSI sensor, is that light is absorbed far from the PD and the resulting photoelectron may diffuse and be collected by an adjacent pixel. This is illustrated in Fig. 4.6.
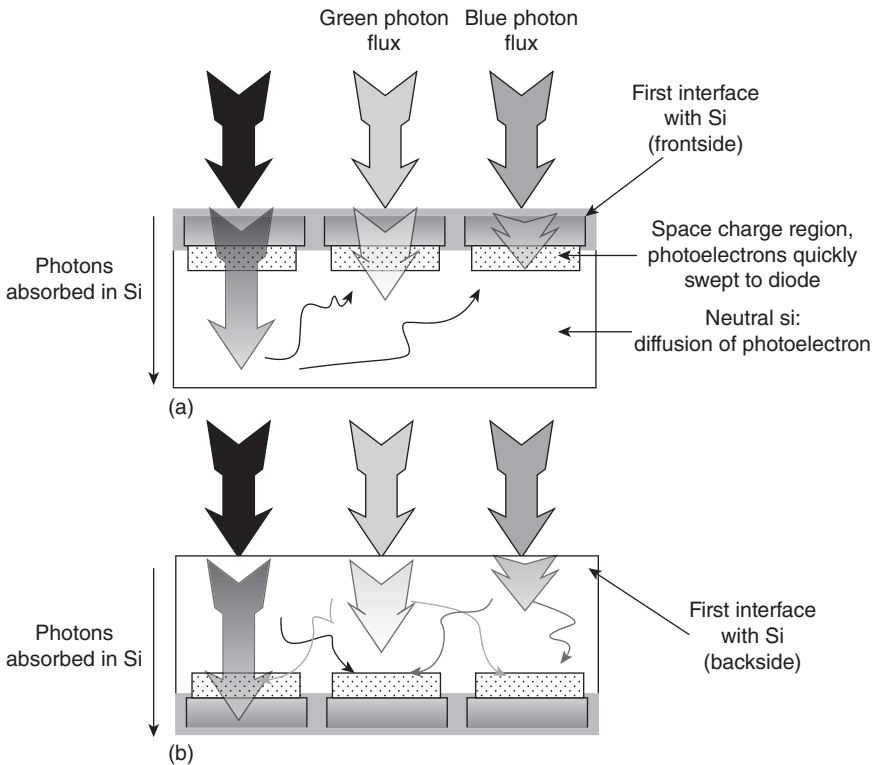
4.6 Illustration of photon absorption in FSI pixel compared to BSI.

The increased electrical cross talk reduces the modulation transfer function (MTF) for monochrome sensors and degrades the post color interpolation SNR [17]. This is also a known problem on FSI devices but in the latter case, it is limited to red photons which are absorbed deep below the PD [18, 19]. In BSI pixels, the problem becomes severe since a significant amount of green photons ($\lambda = 550$ nm, ~30% absorption in the first 0.5 µm) and the blue photons ($\lambda = 470$ nm, ~70% absorption in the first 0.5 µm) are strongly absorbed [20, 21].

The optimization path taken by OVT in solving this problem was to create the diode implants much deeper than the ones used for the FSI device. This same route was also reported by ST Microelectronics [22] who reported implants energies of at least 800 keV and for some cases, even more than 1 MeV, to create the deep diodes which are suitable for green photoelectron absorption. The main problem in using this method is the difficulty to achieve fast transfer of electrons from the deep diode to the sampling node which causes increased noise, image lag and non-linear response at low illumination levels. An alternative similar solution is to use a deep P+ implant to try to isolate the diodes [23]. Two further alternatives which rely on deep implants raise additional problems such as the difficulty to anneal the silicon defects which are caused by the high energy implants. These defects, without a proper anneal, may cause high dark current (DC) and blemish. As is very well known, proper anneals require a long, high temperature process that is not compatible with modern CMOS technology nodes.

For companies that choose SOI as their starting material there is an additional problem to solve during frontend fabrication. One of the most fundamental problems that was solved in the early days of silicon imagers was gettering of the metal elements during fabrication [23, 24]. A P/P+ doped wafer usually contains oxygen precipitates in the P+ layer and even heavily doped polysilicon on the backside of the wafer. These well-known methods are used to getter or 'capture' the metal atoms far away from the diode collection area. Sensors manufactured without this gettering technique (or others) will produce images that exhibit very poor performance in the dark due to dark current generation, centered around the metal atom. However, when producing the image sensor on an SOI substrate, these known methods cannot be used because heavy metals cannot pass through the buried oxide film. Therefore, the oxygen precipitates or polysilicon layer on the backside of the wafer cannot prevent heavy metal contamination near the active region of the PD and are basically useless. A different and new gettering method for heavy metals, suitable for pixels in SOI is required.

There are many published techniques for achieving good gettering in SOI substrates. We usually classify the techniques into three groups:

1. The 'local gettering' technique that relies on local gettering centers in close proximity to the pixel active region [25, 26]. Implementation of this technique is achieved through modifications to the front end flow.
2. The 'vertical gettering' technique which relies on gettering centers created in the epi active region or in some cases in the silicon substrate [27].
3. The 'post process' technique which relies on (i) creating gettering centers after bottom oxide removal, (ii) special thermal treatment which moves metallic atoms into the gettering centers, and finally (iii) removal of the damaged layer [28].

We will briefly point out two examples of these techniques which will emphasize the severity of the problem and the efforts needed to successfully solve it. The first example of 'vertical gettering' is from Sony in a 2010 patent [29]. Sony suggests using gettering centers on the SOI substrate part. A path for the metal contamination in the active epi layer to the silicon is created by damaging the SOI oxide layer which separates the epi from the Si substrate. Implementation of such a technique requires very strong cooperation of many parties including SOI and epi layers vendors. Another profound change to a standard front end process can be found in a 'lateral gettering' example which was patented by ST Microelectronics in 2006 [30]. Here, ST Microelectronics suggests gettering the metallic atoms in polysilicon filled deep trenches. The patent application does not specify where this gettering center should be placed, but a similar module was already reported [19] by ST Microelectronics in their 1.4 micron FSI pixel technology. The extra process modules are added immediately after the shallow trench isolation is completed and use 45 nm photolithography and etch tools. The aspect ratio for the etched trench is higher than 1 to 25. ST Microelectronics reports that special passivation was used on the deep trench isolation (DTI) walls before gap filling, in order to reduce DC to the level of the DC achieved without the DTI. To conclude this subject, in Fig. 4.7, grey level images (50% saturation at 30 fps) taken at wafer level for two different BSI products which ran at TowerJazz on SOI starting material are illustrated. In Fig. 4.7(a) the process does not include a special gettering process and exhibits many bright points. Clean images can be seen in Fig. 4.7(b) for an optimized gettering process.

4.3.3 Bonding

The next process step, taken after finishing the frontend flow including all of the metallization layers, is the bonding of the device wafer to an additional handling wafer. Bonding is a crucial step which is required in order to supply mechanical support for the next steps of fabrication. Direct bonding

Y/X	0	1	2	3	4
1					
2					
3					
4					
5					

(a)   Wafer notch   (b)
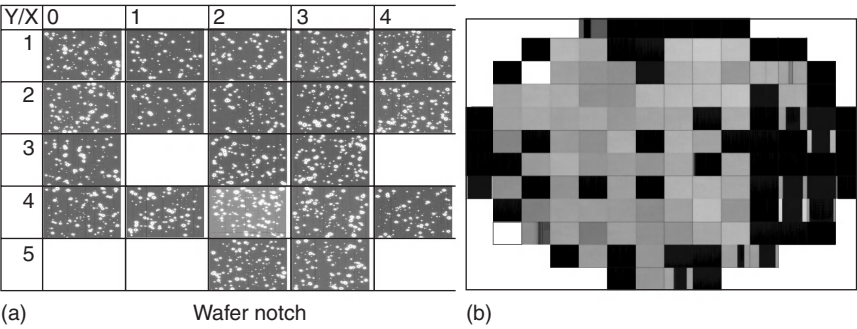
4.7 Dark image taken from a full wafer with BSI sensor manufactured by TowerJazz (a) grey image wafer map for early process version with no proper gettering technique and (b) grey image wafer map for process which includes gettering technique.

Deposit silicon oxide   Si oxide

Si wafer

Chemo-mechanical polish
CMP polish
Si wafer

Plasma activation
Plasma
Si wafer

Activate (terminate) with chemical species allowing formation of chemical bonds a1 CMOS compatible temperatures

Place surfaces together
Si wafer

Si wafer

Spontaneous chemical reaction with byproducts diffusing away from bond interface
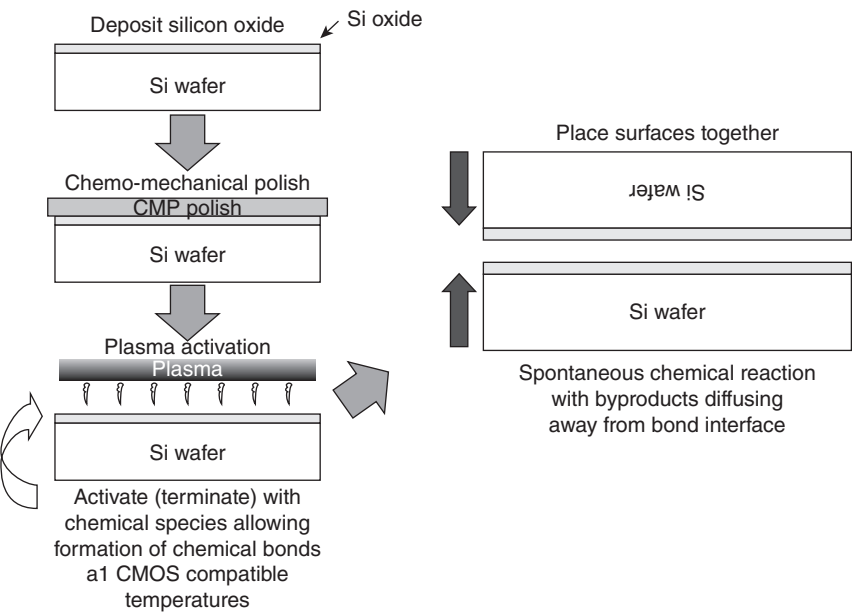
4.8 Modern wafer to wafer bonding (ZiBona® process flow) – taken from Ziptronix presentation (Direct Bonding TM).

of two wafers requires stringent control of surface properties like micro-roughness, wafer flatness, and surface chemistry. The main parameters which are affected by poor control of wafer uniformity and flatness are voids in the boundary interface, mechanical stress imposed on the device wafer which can easily cause alignment errors in the subsequent fabrication steps, and bonding strength and reliability [31–33]. Figure 4.8 illustrates a modern wafer to wafer bonding technique known as direct bonding™ from

Ziptronix [33–34]. In this method, the device wafer and the handling wafer surfaces are cleaned using special chemical mechanical polish(CMP) and then plasma activated. The actual bonding occurs when the two surfaces are placed one on top of the other.

4.3.4 Wafer thinning

After bonding, the wafer is flipped and ready for the next process flow on the wafer backside. The first process on the backside is wafer thinning. The quality of the thinning has a direct impact on the overall performance of the BSI sensor and is considered one of the most important challenges of BSI integration. In the case of an SOI wafer being chosen, the wafer thinning is quite simple and involves removing the silicon and stopping on the BOX. Immediately after, the BOX itself is removed (now the active silicon is used as an etch stop) and the wafer is ready for the next steps of fabrication. These steps are well described in Fig. 4.9.

The simplicity of the wafer thinning step in the case of SOI is the main reason for choosing this starting material over the P/P+ standard wafer since thinning of a P/P+ wafer is very challenging. In such case, there is no etch stop layer and thus, one of the most common thinning methods is to use the concentration difference between the P and P+ layer as an etch stop interface. The final silicon thickness after thinning target is between 2 to 3 microns, for pixel sizes under 2 microns, so the thinning of a P/P+ wafer usually starts with mechanical grinding from initial thickness of about 700 microns down to 20 microns, and only then comes the special etch, that

Handle wafer	
Buried oxide layer (BOX)	Buried oxide layer (BOX)
SEED (P type)	SEED (P type)
P- epi Full front end process (pixels, transistors...)	P- epi Full front end process (pixels, transistors...)
Metallization, inter dielectric and passivation	Metallization, inter dielectric and passivation
2nd handle wafer	2nd handle wafer

(a)   (b)
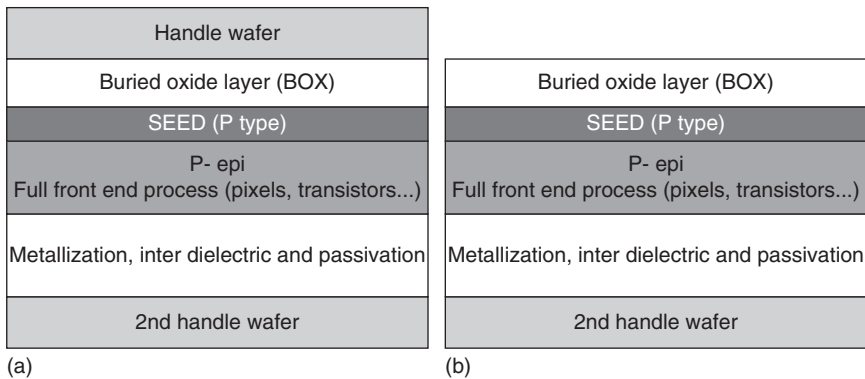
4.9 (a) SOI wafer after wafer bonding and flipping and (b) SOI wafer after removing the first handling wafer and stopping on the SOI bottom oxide.
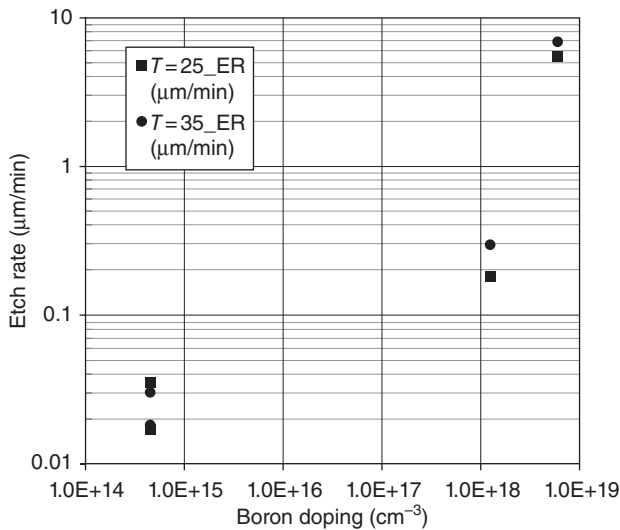
4.10 Etch rates differences with respect to the doping level of a bare silicon wafer.

removes the silicon to the final thickness target. In Fig. 4.10, the etch rates of an hydrofluoric–nitric–acetic (HNA)-based solution is plotted versus the doping level of a bare silicon wafer. It can be seen that the etch rate is negligible until a doping level of 10^{18} cm^{-3} atoms. The doping concentration of the P active layer of a P/P+ wafer is in the order of 10^{15} while the P+ concentration level is in the order of 10^{19}. This, in principle, should supply enough margin for thinning process control.

In practice, there are many practical problems with utilizing a good wet etch thinning. The first problem is the out diffusion of the boron from the P+ layer to the P active layer. At the start of the frontend fabrication process, the boundary of the P+ and the P layer is well defined and controlled by the wafer manufacturer. But this clear boundary becomes somewhat fuzzy since, during the frontend processing, there are many long high temperature anneals that drive the boron from the P+ layer into the P layer. As a result, when the thinning step arrives, there is no longer a clear boundary that can serve as an etch stop mechanism. The second problem is that the edges of the wafer are exposed to the etchant from both the top of the wafer and from its side. This causes an increased etch rate in the wafer periphery compared to its center. In Fig. 4.11, a picture of a 200 mm wafer that was etched without any edge protection is shown. It can be easily seen that the active silicon is fully etched at the periphery. Protecting the edges during etch is more complicated than one may think, since standard
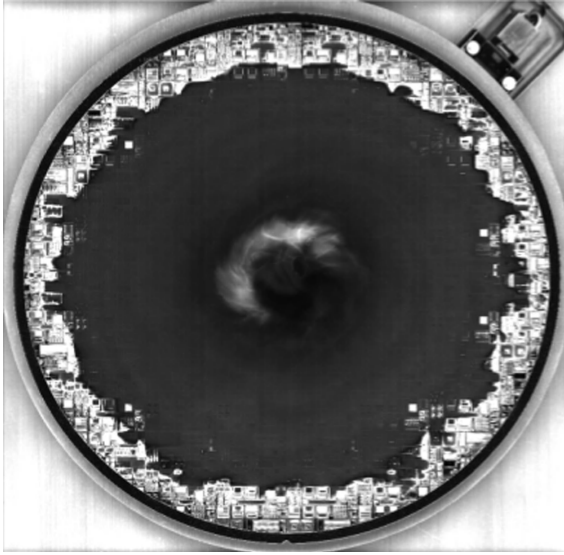
4.11 A 200 mm P/P+ wafer after HNA etch without edge protection. The P active layer thickness is 5 microns and the HNA process started after grinding of the silicon down to 20 microns. At the edge of the wafer there is no active silicon left and the first inter-dielectric layer of the metallization is fully exposed.

photo resists which are regularly used in modern silicon fabrication for such purposes, are also etched by the HNA-based solution. Moreover, protecting the wafer edge requires non-standard exposure since the photo resist material has to be completely removed from all the wafer area apart from its edges, a task that cannot be performed by using standard photolithography mask and a photo scanner tool. Finally, the last problem is that the surface roughness after etch is not sufficient and a special CMP process is required in order to get the required surface roughness for the next processing step and for a proper device, as shown in Fig. 4.12.

The difficulties in achieving good thickness control on a P/P+ wafer is the main reason why today SOI starting material is used for most BSI high end sensors. The optical area of such sensors is usually relatively large and even a moderate change in the active silicon layer thickness will show up as optical response variation along the sensor. In cinematography, it is customary to have two or more cameras on the shooting set. This dictates very strict demands on the allowed variations between sensors which is very hard to achieve without precise control of the thickness of the active silicon layer. Nevertheless, TSMC reports impressive thickness control of less than ±0.1 micron achieved on 200 mm bulk wafers.

1:Y = 0.1045 µm

(c)

T200209/#03, backside, center

TSL 24.0 kV 0.9 mm × 150 k SE 4/4/12 07:19   300 nm

(a)

(b)
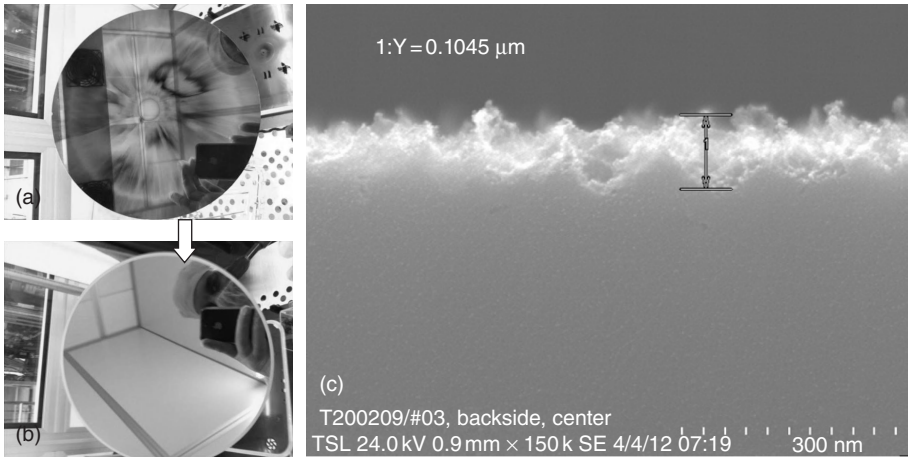
4.12 (a) 200 mm wafer surface after HNA thinning, (b) 200 mm wafer surface after HNA thinning and CMP and (c) X-section of the device wafer after HNA thinning showing the surface roughness.

4.3.5 Surface passivation

The next process step is backside surface passivation. The FSI method for surface passivation is based on surface pinning using high dose implant, activated and annealed in relatively high temperatures. However, this method cannot be applied to the BSI wafer at the final stages of manufacturing since metallization layers already exist on the wafer and therefore, the wafer cannot withstand these high thermal cycles. There are at least three widely accepted methods which are used to passivate the back surface of the BSI wafer. A good review of the different methods being used can be found in a paper by Bedabrata Pain from IISW 2009 [13]. We will discuss briefly these three common methods.

Laser annealing

This method involves the following steps:

- The backside of the wafer is implanted by shallow p-type implant (for n-type photodiode which is manufactured on a p-type substrate).
- The implant is annealed and activated by a powerful laser beam pulse. The pulse length is in the order of several hundred nsec (<200 nsec). The heat generated by the pulse is enough for local annealing of implant damages and activating of the traps [35].
- Since the pulse is very short, the temperature is the metallization area of the wafer does not change.

This method is widely used for commercial manufacturing of BSI CMOS image sensors. Apart from using relatively high cost dedicated equipment which is not standard for regular CMOS foundries, the main problems associated with this method are:

- Difficulties in annealing large sensors – multiple shots of laser on one die are required in order to cover the sensor area. This can cause later artifacts in the image.
- Need to optimize the implant + anneal to have good quantum efficiency in short wavelengths

Passivation using low temperature growth of highly doped epi-silicon layers

This method was invented by the California Institute of Technology (CALTECH) and Jet Propulsion Labs (JPL) [35, 36] and was widely used for scientific applications [37] of BSI close coupled device (CCD) and CMOS sensors. The main advantage of this method is the high quantum efficiency which can be achieved for short wavelengths. The main disadvantage is the requirement of low temperature epitaxial growth that can be carried out only by using a molecular beam epitaxial (MBE) growth machine which is not standard in the silicon industry, is very costly and has low throughput.

Passivation using dielectric with negative fixed charge

Recently [38], Guy Meynants of CMOSIS, suggested using Al_2O_3 as a passivation layer. The Al_2O_3 is deposited over SiO_2 and has some inherent fixed negative charge. This negative charge creates a mobile hole accumulation layer close to the interface between the Si and the oxide and can, in some conditions, pin the interface and deactivate the charge. The main advantage of this method is the relative simplicity compared to the previously described methods and the compatibility with standard silicon foundry equipment. The main disadvantage is the difficulty to get a controlled high amount of fixed charge in the Al_2O_3 layer. Here, the thin oxide layer just above the active silicon needs to be of good quality and have a low level of defects.

4.3.6 Light coupling

The next process steps are dedicated to the light coupling into the sensor. It is customary at this step to create:

- an anti-reflecting coating
- a metal grid
- a color filter array (CFA) and u-lenses.
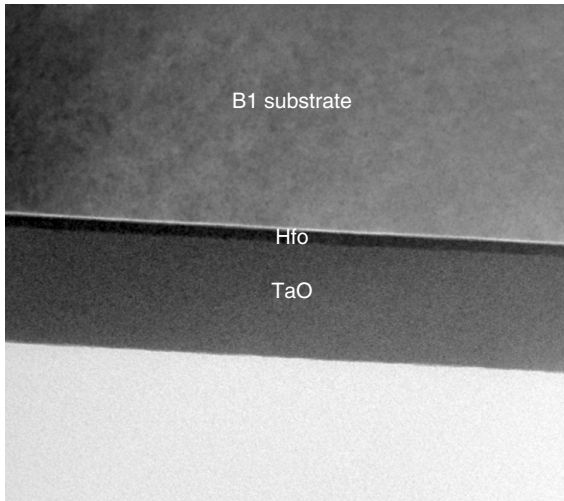
4.13 Anti-reflecting coating example of 1.55 μm generation from Sony BSI (courtesy of Chipworks [40]).

Anti-reflecting coating

One of the big advantages of a BSI sensor is the freedom to use a dedicated optimized ARC to increase QE for a given application. This is the main driving force for high end sensor designers to move into BSI. High volume sensor manufacturers tend to use HfO, SiON and SiO$_2$, but more exotic material systems like MgF$_2$ and Ta$_2$O$_5$ are commonly used as well. A nice example of ARC deposited on a Sony 1.55 BSI sensor can be seen in Fig. 4.13.

Metal grid

Deposition and patterning of a thin metal grid at the physical boundaries of each pixel is a common practice for most high volume manufacturers. For a small pixel, there is a significant amount of light arriving at the pixel periphery. These photons tend to increase the cross talk between pixels and can also hit the color filter array in places where the color filters are not defined which causes increased photon response non uniformity (PRNU).

Color filter array (CFA) and u-lenses

Color filter array and u-lenses patterning is very similar to the front end case. The main difficulty here is to align the optical stack and the metal grid to the pixel active diode area. Typically, the Si thickness after thinning may

Color to deep trench bar in bar
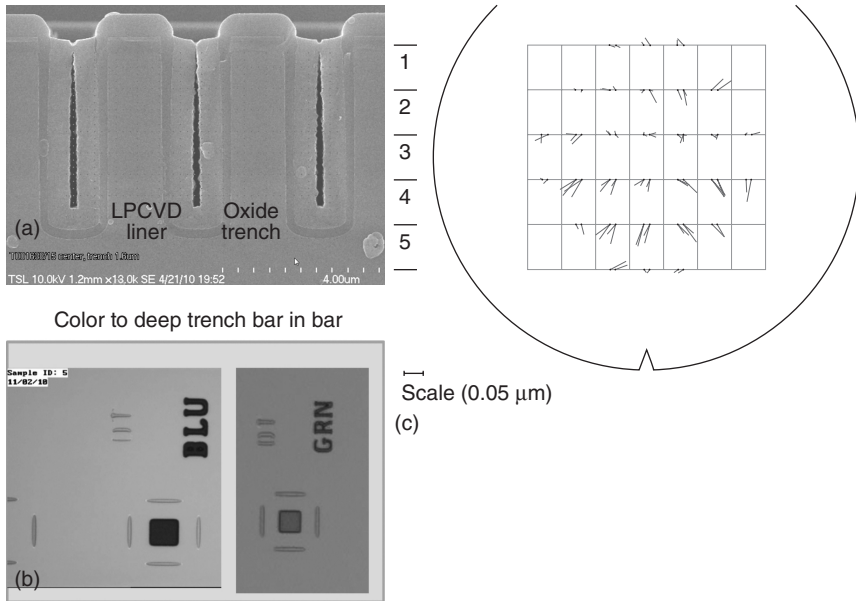
4.14 Alignment of the color filter array to the frontend using deep trenches (a) SEM pictures of the deep trenches alignment mark, (b) blue and green color filter alignment to the deep trench mark and (c) typical wafer map which measure the misalignment between the deep trench mark and the color filter.

be over 2 um for a high volume small pixel sensor and over 5 um for a high end sensor. Hence, alignment of backside structures to the pixel active area needs to upgrade to an IR alignment tool or devolving, a solution which can work with the standard fabrication tool. At TowerJazz, for example (see Fig. 4.14), deep trench marks formed at the frontside of the wafer, and subsequently exposed at the bonded backside after thinning, are used for backside alignment.

4.4   Interface solutions to BSI sensors

Finally, a few words on the interface of the backside illuminated sensor to the outside world. For a conventional wire bonded frontside sensor, the passivation layer above the last metallization is removed above the pads. This is a similar/identical process to a regular CMOS chip. The pads are then bonded and wired using standard wire bond (usually gold) techniques. For BSI sensors, the pads are 'hidden' by the thick supporting wafer and the standard pad opening is not applicable and therefore, some alternative solutions are used. A solution known as through silicon Via (TSV) where a deep via is etched through the support wafer to the 'hidden' pads is known
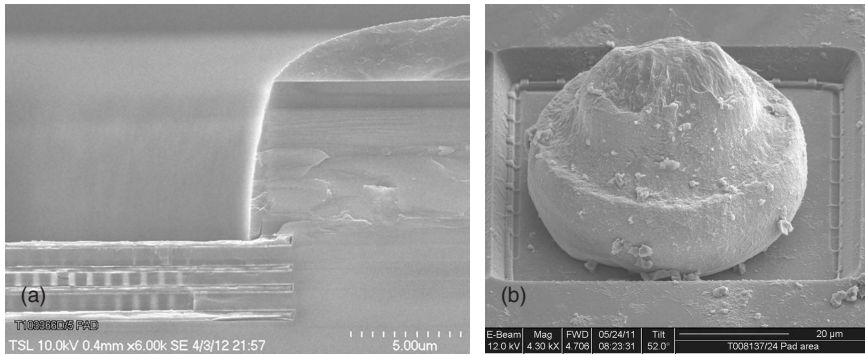
4.15 Pad opening scheme (a) cross section showing deep etch from the sensor backside which stops at M1 of the pad structure and (b) arbitrary unit bonding in the center of the backside pad.

to be used by some mass market players [39], and also for FSI wafer level packaged (WLP) devices. A more straightforward solution is to open the pads from the backside by etching first the anti-reflective dielectric coating, through the epi active layer and the first frontside inter-dielectric, stopping at the M1 pads as shown Fig. 4.15.

A discussion on the most advanced method for interfacing a backside illuminated sensor to the outside world concludes the chapter. In this method, the carrying wafer which was previously discussed as a dummy wafer, becomes an electrically active circuit. Some circuits which used to be printed on the image sensor are now fabricated on the carrier wafer. The image sensor wafer is fabricated as before. The last metal in each wafer is designed to connect the electrical circuit from the image sensor wafer to the readout integrated circuit (ROIC) wafer. The wafers are aligned frontside to frontside, bonded and annealed to create the full active device and the backside flow continues as described before. Contacts can be either TSV on the ROIC wafers or on the imager wafer or regular wirebond as described above. The process of electrically bonding the two wafers is schematically described in Fig. 4.16. Sony has already announced [40] an image sensor chip which is using some readout circuits in the carrier wafer. The next evolutionary step in this technology, considered to be the 'holy grail' for pixel and camera designers all over the world, is to use the same technology with one or two connections per pixel between the image sensor wafer to the ROIC wafer. A cross section of a wafer to wafer bonding with 2 um separation is shown in Fig. 4.17. Some prototypes of a 3 Meg imager were reported by Ziptronix [41]. If this technology finally reaches the market, the possibilities are endless, especially for high end sensors. For example, one can design the image sensor wafer with a dedicated process with long thermal anneal which will result in a very clean image due to low
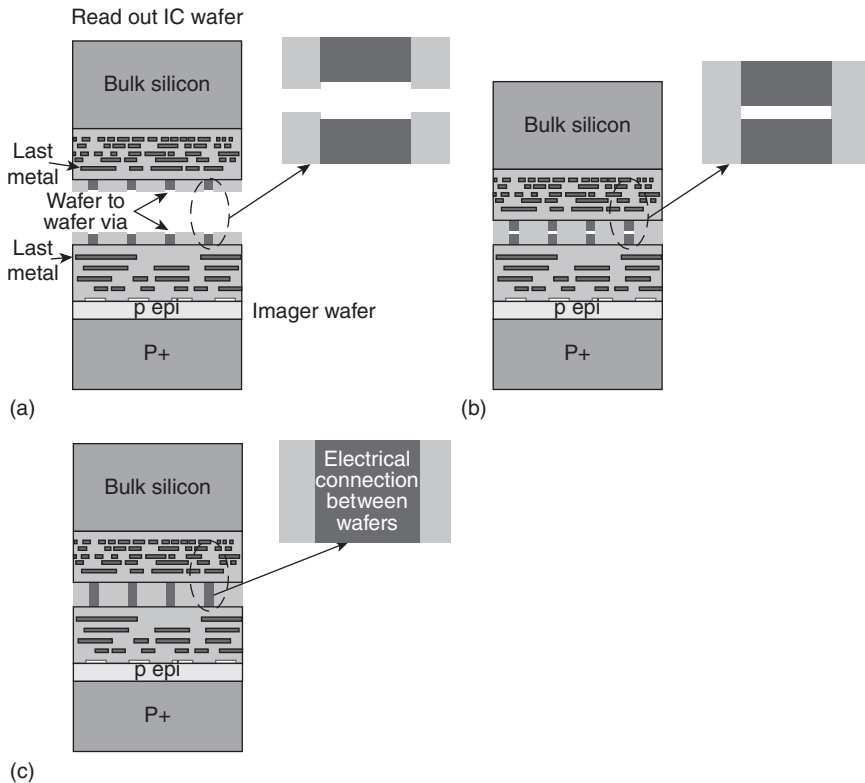
4.16 Illustration of a typical 3D bonding between imager wafer and an ROIC wafer (a) alignment of imager wafer and ROIC wafer (each wafer surface has been prepared carefully before alignment), (b) wafers are now carefully placed one on top of the other. The surfaces of the waters are cleaned and sometimes activated by plasma. As a result strong oxide to oxide bonding between the two wafers is achieved, (c) thermal treatment causes the metal to expand and create electrical connection between metals.

dark current and low 1/f noise. This cannot be achieved in a traditional image sensor due to many compromises imposed by the fast logic process [42]. Application of this technology includes a better global shutter pixel with very low parasitic light sensitivity, low noise application and sophisticated high dynamic range sensors like time to saturation sensors or well cycling sensors [43].

4.5   Conclusion

In this chapter we reviewed some of the major challenges and trends in the fabrication of backside illuminated sensors. We started by describing the
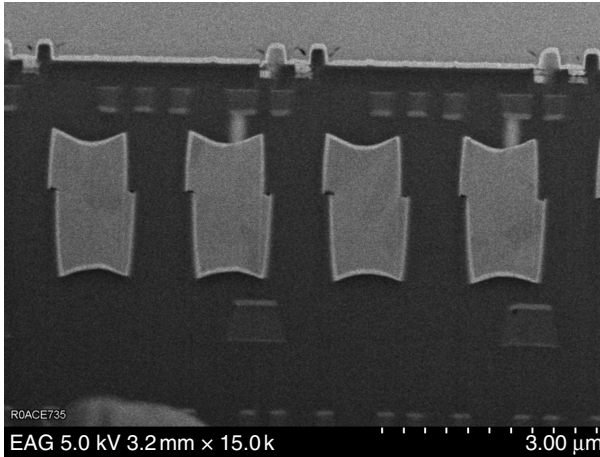
4.17 Cross section of a wafer to wafer electrical test structure with 2 µm separation between the vias (courtesy of Ziptronix [35]).

advantages of a BSI sensor over an FSI sensor. We showed that the possibility to use advanced materials to create a highly optimized anti-reflecting coating is the main driving force for BSI technology for scientific and high end applications. On the other hand, for pixels sizes <2 um intended for the mass market, BSI is a natural choice and a necessity. It might even turn out to be a cheaper solution than the other complicated front-end methods which are used to couple light into these pixels.

Our discussion continued with a detailed look at the fabrication of a wafer level backside illumination sensor. We reviewed the pros and cons of SOI starting material which is widely used for scientific applications and showed that the gettering issue needs to be addressed properly in order to achieve an acceptable working sensor.

Finally, we discussed the different options of connecting the sensor to the outside world. The exciting concept of an image wafer bonded to an ROIC wafer was brought forward. We predicted that BSI sensors with a pixel to pixel contact to an ROIC wafer could enhance high end image sensor performance such as high quality global shuttering and sophisticated high dynamic range sensors.

4.6   References

1. G. Agranov, T. Gilton, R. Mauritzson, U. Boettiger, P. Altice, J. Shah, J. Ladd, X. Fan, F. Brady, J. McKee, C. Hong, X. Li and I. Patrick, 'Optical-electrical characteristic of small, sub-4 µm and sub-3 µm pixels for modern CMOS image sensors,' in *IEEE CCD and AIS Workshop*, 2005, pp. 206–209.

2. D. Edelstein, J. Heidenreich, R. Goldblatt, W. Cote, C. Uzoh, N. Lustig, P. Roper, T. McDevitt, W. Motsiff, A. Simon, J. Dukovic, R. Wachnik, H. Rathore, R. Schulz, L. Su, S. Luce and J. Slattery, 'Full copper wiring in a sub-0.25 µm CMOS ULSI technology,' in *Electron Devices Meeting, 1997. IEDM '97. Technical Digest., International*, 1997, pp. 773–776.

3. M. Cohen, F. Roy, D. Herault, Y. Cazaux, A. Gandolfi, J. P. Reynard, C. Cowache, E. Bruno, T. Girault, J. Vaillant *et al.*, 'Fully optimized Cu based process with dedicated cavity etch for 1.75 µm and 1.45 µm pixel pitch CMOS image sensors,' in *Electron Devices Meeting, 2006. IEDM'06. International*, 2006, pp. 1–4.

4. J. Adkisson, J. Gambino and Z. He, 'CMOS imager array with recessed dielectric,' US Patent 7781781B2, 2010.

5. H. E. Rhodes, 'Local interconnect structure for a CMOS image sensor and its manufacturing method,' US Patent 7345330, 2008.

6. J. Gambino, B. Leidy and J. Adkisson, 'CMOS imager with copper wiring and lightpipe,' in *International Electron Devices Meeting (IEDM)*, 2006.

7. G. Agranov, R. Mauritzson, J. Ladd, A. Dokoutchaev, X. Fan, Z. Li, R. Yin, V. Lenchenkov, S. Nagaraja, W. Gazeley, J. Bai, H. Lee, A. D'Anna and G. De-Amicis, 'Pixel continues to shrink. Pixel development for novel CMOS image sensors,' in *International Image Sensor Workshop*, 2009.

8. H. Reznik, R. Edelstein, M. Shach-Caplan, F. Dulberg, V. Kamenetsky, K. Karaste and J. T. Rantala, 'Integration and characterization of spin on dielectric materials in image sensor devices,' *Proceedings of SPIE*, 2008.

9. J. Gambino, K. Ackerson, B. Guthrie, B. Leidy, W. Abadeer, S. Mongeon, D. Meatyard, J. Adkisson, R. J. Rassel, J. Ellis-Monaghan and M. Jaffe, 'Reliability of CMOS image sensor with polymer lightpipe,' in *International Image Sensor Workshop*, 2007.

10. Chipworks, http://www.chipworks.com/blog/technologyblog/2012/05/23/a-survey-of-recent-image-sensor-pixel-structures/ (accessed: 29 January 2013).

11. S. Wuu, 'BSI technology with bulk Si wafer,' in *Symposium on Backside Illumination of Solid-State Image Sensors*, 2009.

12. H. Rhodes, 'Mass production of BSI image sensors: performance results,' in *Symposium on Backside Illumination of Solid-State Image Sensors*, 2009.

13. B. Pain, 'Backside illumination technology for SOI-CMOS image sensors CMOS image sensors,' in *Symposium on Backside Illumination of Solid-State Image Sensors*, 2009.

14. S. Iwabuchi and Y. Maruyama, 'A back-illuminated high-sensitivity small-pixel color CMOS image sensor with flexible layout of metal wiring,' *Solid-State Circuits Conference, 2006*. ISSCC 2006.

15. R. Edelstein, O. Katz, B. Lavi, I. Aberman, S. Rosenthal, M. Shadmi, S. Arad, N. Golan, M. Shach Caplan, M. Massalha, C. Lagahe-Blanchard, L. Marinier, R. Fontanière, A. Castex, M. Broekaart, M. Martinez, N. Milhet, A. Rigny and C. Pelissier, 'Process integration aspects of back illuminated CMOS imagers using smart stacking TM technology with best in class direct bonding,' *International Image Sensor Workshop*, 2011.

16. A. Getman, T. Uvarov, Y. I. Han, B. Kim, J. C. Ahn and Y. H. Lee, 'Crosstalk, color tint and shading correction for small pixel size image sensor,' in *International Image Sensor Workshop*, 2007, pp. 166–169.

17. O. Yadid-Pecht and R. Etienne-Cummings, *CMOS Imagers: From phototransduction to image processing*. Springer, 2004.

18. J. Shah, M. E. Jernigan and R. I. Hornsey, 'Characterization and deblurring of lateral crosstalk in CMOS image sensors,' *IEEE Transactions on Electron Devices*, 50(12), pp. 2361–2368, December 2003.

19. A. Tournier, F. Leverd, L. Favennec, C. Perrot, L. Pinzelli, M. Gatefait, N. Cherault, D. Jeanjean, J.-P. Carrere, F. Hirigoyen, L. Grant and F. Roy, 'Pixel-to-pixel isolation by deep trench technology: Application to CMOS image sensor,' in *International Image Sensor Workshop*, 2011.

20. J. Nakamura, *Image Sensors and Signal Processing for Digital Still Cameras*. CRC, 2005.

21. 'Optical-properties-of-silicon,' http://pveducation.org/pvcdrom/appendicies/ optical-properties-of-silicon (Accessed: 4 February 2014).

22. J. Michelot, F. Roy, J. Prima, C. Augier, F. Barbier, S. Ricq, P. Boulenc, Z. Essa, L. Pinzelli, M. Gatefait and J. Broquin, 'Back illuminated vertically pinned photodiode with in depth charge storage,' in *International Image Sensor Workshop*, 2011.

23. J. Prima, F. Roy, P. Coudrain, X. Gagnard, J. Segura, Y. Cazaux, D. Herault, N. Virollet, N. Moussy, B. Giffard and P. Gidon, 'A 3 mega pixel back-illuminated image sensor in 1T5 architecture with 1:45 μm pixel pitch,' in *International Image Sensor Workshop*, 2007, no. 1.

24. T. Reiner and P. J. Coppen, 'Characterization of leakage current with phosphorus gettering,' in *Proceedings of 1978 International Conference on the Application of Charge Coupled Devices*, 1978.

25. N. Teranishi, 'Dark current and white blemish,' in *Image Sensors*, 2012.

26. D. J. S. Chepis *et al.*, 'Method and structure for front-side gettering of silicon-on-insulator substrates,' US Patent US005646053A, 1997.

27. R. Capedelli, L. Di Turi and D. Faralli, 'Process for manufacturing a SOI wafer with improved gettering capability,' US Patent App. 11/347, 801, 2006.

28. Shinji Uya *et al*, 'Backside illuminated imaging device, semiconductor substrate, imaging apparatus and method for manufacturing backside illuminated imaging device,' US Patent 7781715 (Accessed: 24/10/2010).

29. 'SOI substrate and method for producing the same, solid-state image pickup device and method for producing the same, and image pickup apparatus,' US Patent 810199, http://www.google.com/patents/US8101999?pg=PA1&dq=Ta kizawa+image+pickup+soi&hl=en&sa=X&ei=RgMIUavLO6mC4gS8pYCQAg &ved=0CDMQ6AEwAA#v=onepage&q=Takizawa image pickup soi&f=false (Accessed: 29 January 2013).

30. Patent US20060194409 – 'Process for manufacturing a SOI wafer with improved gettering capability.' US Patent 20060194409, http://www.google.com/patents/ about?id=VHSYAAAAEBAJ (Accessed: 29 January 2013).

31. G. Gaudin and G. Riou, 'Low temperature direct wafer to wafer bonding for 3D integration: Direct bonding, surface preparation, wafer-to-wafer alignment,' in *IEEE 3D Systems Integration Conference 3DIC 2010*, 2010.

32. C. Lagahe-Blanchard and B. Aspar, 'Wafer stacking: Key technology for 3D integration,' in *2009 IEEE International SOI Conference*, 2009, pp. 1–4.

33. C. Blanchard and I. Radu, 'Smart stacking TM technology: An industrial solution for 3D layer stacking,' in *EEE International Conference on IC Design Technology*, 2011.

34. P. Enquist, '3D integration applications for low temperature direct bond technology,' *Low Temperature Bonding for 3D Integration (LTB-3D), 2012 3rd IEEE International Workshop on IEEE*, 2012.

35. F. R. K. Huet, C. Boniface, J. Venturini, Z. Ait Fqir Ali-Guerry, R. Beneyton, M. Marty and D. Dutartre, 'High performance and high yield junction formation with full device exposure laser thermal annealing,' in *International Image Sensor Workshop*, 2011.

36. M. E. Hoenk, P. J. Grunthaner, F. J. Grunthaner, R. W. Terhune, M. Fattahi and H. F. Tseng, 'Growth of a delta-doped silicon layer by molecular beam epitaxy on a charge-coupled device for reflection-limited ultraviolet quantum efficiency,' *Applied Physics Letters*, 61(9), pp. 1084–1086, 1992.

37. B. Pain, T. Cunningham, S. Nikzad and M. Hoenk, 'A back-illuminated megapixel CMOS image sensor,' *IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, 2005.

38. G. Meynants, 'Method of manufacture of a backside illuminated image sensor,' US Patent App. 12/939, 599, 2010.

39. R. Fontaine, 'A review of the 1.4 μm pixel generation,' in *International Images Sensor Workshop*, 2011.

40. Sony, 'ISX014, CMOS image sensor (SOC) with built-in high-performance image-processing engine and camera system functions,' 2012, http://www.sony.net/Products/SC-HP/cx_news/vol70/pdf/isx014.pdf.

41. 'Ziptronix,' http://www.ziptronix.com/ (Accessed: 4 January 2013).

42. A. Lahav and A. Fenigstein, 'The megapixel race: a chip designer's point of view,' *EDN (Electrical Design News)* 53 (12), pp. 57, 2008.

43. D. Yang and A. El Gamal, 'Comparative analysis of SNR for image sensors with enhanced dynamic range,' *Proceedings of SPIE*, 1999.

5

Circuits for high performance complementary metal-oxide-semiconductor (CMOS) image sensors

B. CHOUBEY, University of Oxford, UK and
W. MUGHAL and L. GOUVEIA, University of Glasgow, UK

Abstract: Complementary metal-oxide-semiconductor (CMOS) image sensors have become the mainstay of imaging technology with applications ranging from low end requirements such as that of toys to high end scientific applications. This has been achieved by several advances including noise reduction, pixel count, capture speed enhancement, global shuttering, power reduction as well as dynamic range improvement. This chapter reviews circuits utilised to provide these improvements in CMOS image sensors.

Key words: pixel size, fixed pattern noise, correlated double sampling, ADC, pixel speed, wide dynamic range, threshold comparison pixels, logarithmic pixels, global shutter, dark current.

5.1   Introduction

Complementary metal-oxide-semiconductor (CMOS) image sensors were initially projected to be a low cost though poor image quality replacement for charge coupled devices (CCDs) for imaging systems. However, research and development, particularly in process modifications as well as novel interface circuits have significantly improved the performance of these devices to a level where they find applications in high performance imaging requirements such as biomedical imaging, high end digital cameras as well as scientific instrumentation. In this chapter, we review circuits used in some of the high performance applications.

The chapter has been divided to cover broad areas of performance. In the first section, we review approaches to increase the number of pixels and consequently the resolution of the image sensor. CMOS imaging historically lagged the CCDs on account of its poor noise performance. Section 5.2 studies various noise sources and the approaches to reduce them. Most commercial image sensors require video frame rate abilities. However, scientific imaging often requires frame rates of thousands of frames per second. Furthermore, with increasing number of pixels, the speed of the sensor as well as interconnects have to be enhanced to meet the specification.

124

Techniques to achieve this have been studied in Section 5.3. This section also provides a brief analysis of circuits to achieve global shuttering. One of the largest markets for image sensors is that of hand-held devices including the mobile phone, where battery life is of paramount importance. This also means that the image sensor should consume as low power as possible. Section 5.4 therefore studies techniques to reduce the power consumption of image sensors. The limited dynamic range of image sensors often leads to inability to record very bright or very dim scenes. Circuit techniques to enhance these are presented in Section 5.5. Finally, Section 5.6 presents a brief overview of a number of other high performance sensors including those with large formats as well as low light sensitivity.

5.2   High resolution image sensors

CMOS image sensor designers have been pushed by market forces to increase the spatial resolution in order to improve image quality as well as to digitally zoom. The latter is of particular significance in mobile cameras, where a lens assembly to provide optical zoom is difficult to integrate. Still cameras are being designed with over tens of millions of pixels. Even video applications have demanded an increase in sensor resolution. For example, HDTV 1080p has a resolution of about 2 million pixels (1920 × 1080). Similarly, ITU approved high-definition standards UHDTV-1 (4k HDTV) and UHDTV-2 (8k HDTV) aim to transmit 8.3 and 33.2 million pixels, respectively [1, 2]. Spatial resolution is a measure of the capability of an image system to correctly distinguish details (contrast levels) of adjacent areas. Different components of the system including the optics as well as the sensor have different measurement methods including resolution test charts and modulation transfer function (MTF) [3]. When the resolution of a system is detector limited, a key factor to increase spatial resolution of a sensor is to reduce the separation between pixels (increasing the spatial sampling frequency). In most of the monochromatic image sensors, the separation between pixels corresponds to the pixel size, therefore resulting in the reduction of the sensor size. In addition, for most applications the image sensor size is generally fixed by historic standards which include the lens size. Therefore, increasing the number of pixels has been one way to increase spatial resolution. This has led to a pixel race leading to an exponential rise in pixel count. Even large scale production applications such as mobile cameras, which would benefit from a reduction in the sensor size, are involved in this race.

5.2.1 Pixel size limits

Increasing the number of pixels for the same imager size has led to a reduction in pixel size. In fact it has decreased to levels where it has been

argued to be already below the optical limit determined by the len's optical resolution. A perfect lens, with no aberrations and a circular aperture creates a circular pattern known as airy pattern, with the airy disk being the first circular area of this pattern where most of the light should be focused. Using a first order approximation and considering paraxial rays, the diameter of the airy disk can be given by:

$$D_A = 2.44 \, \lambda \, F/\# \qquad\qquad [5.1]$$

where λ is the light wavelength and $F/\#$ is the f-number of the lens, with the expression valid for $F/\# > 3$. A simple calculation with an F/2.8 lens leads to a minimum pixel size of around 3.7 μm. Pixels smaller than the optical limit can suffer from light crosstalk and, therefore, suffer from contrast losses.

Nevertheless, current pixel dimensions for high resolution mobile phone cameras have reported pixel pitch as low as a micrometre [4]. It must be noted, however, that this reduced dimension can provide opportunities for oversampling and compensation for the Bayer pattern of colour filter, normally used to obtain colour image. Each Bayer kernel contains four pixels (one for the blue, one for the red and two for the green filter). Therefore, for blue and red colours, the distance between colour is twice the pixel size as shown in Fig. 5.1. However, the optical resolution is not the only limiting factor for the pixel size. In general, reducing the pixel size degrades the capacity of the photodetector to store charge. This leads to poor sensitivity and therefore, poor signal-to-noise ratio (SNR) and dynamic range (DR) [5].

5.2.2 Pixel size improvements

The effect of CMOS pixel shrinkage on the well capacity can be reduced by increasing the pixel fill-factor. One technique to do so is to share the non-detector components with neighbour pixels. In a 4-T buried active pixel sensor (APS); the pixel, the reset, the source follower and the row selection transistors can be shared as they are only used at specific times and are typically disconnected from the photodiode by the transmission gate at other times. Depending of the number of photodiodes sharing these transistors, different terminologies for pixels are adopted: 2.5T (two photodiodes) [1], 1.75T (four photodiodes) [6] as in Fig. 5.2 or 1.5T (six photodiodes). The row selection transistor can also be eliminated if the configuration in Fig. 5.2(b) is used [7]. In these pixels, all floating diffusions (FDs) voltages rather than the ones on the selected row are set to a low value by setting Vrst high. The low voltage switches off the respective source-followers transistors and therefore the selected row pixels can buffer their values due to the winner-take-all effect. In this case the four shared photodiode configuration creates an 1.5T equivalent pixel.
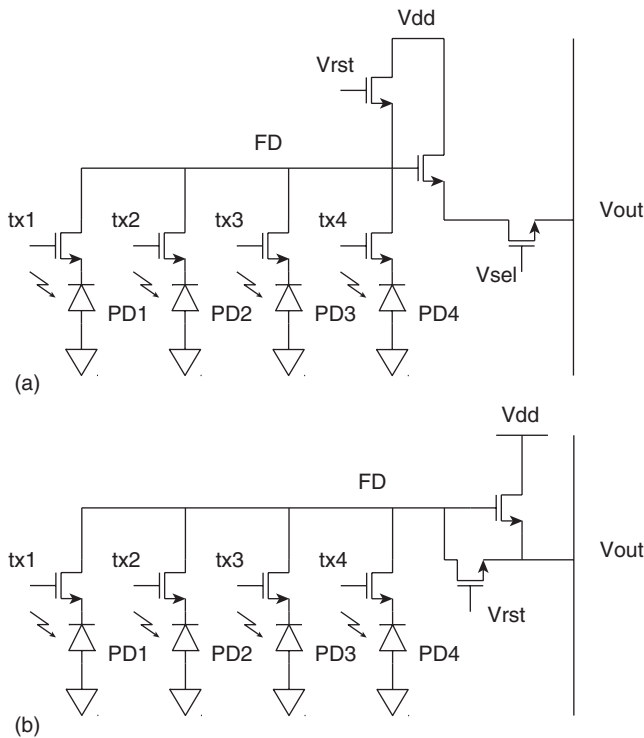
5.1 Pixels with shared transistors to reduce the pixel size. (a) Four pixels sharing common readout circuits and (b) pixel architectures with no row select transistors.
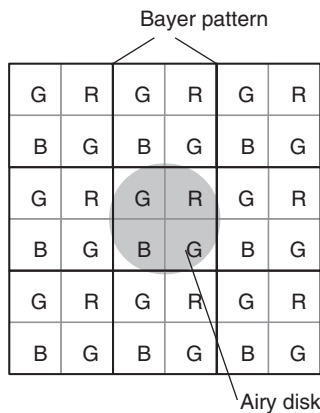
5.2 A typical Bayer pattern as used in CMOS image sensor with an airy disk.

Another approach to increase the fill-factor of the pixel is to use micro-lenses and light-guides (lightpipes). Micro-lenses are placed on top of each individual pixel while light-guides are designed to fill the gap between metal stack. In both cases the light is focused on the photo-detector reducing light crosstalk and scattering.

Until recently, most CMOS image sensors were categorised as front-side illumination (FSI) sensors. In these sensors the light had to cross the space between metal and polysilicon stacked layers before it reached the photo-detector resulting in significant light scattering. In addition, the presence of transistors reduces the available semiconductor area for photon collection and therefore the quantum efficiency of the pixel. This can be resolved by manufacturing image sensing chips, where the substrate is thinned and the sensor is illuminated from its backside illumination (BSI). A backside illuminated sensor can provide almost 100% fill-factor. However, BSI is a costly process and can suffer from crosstalk.

An alternative technique to enhance pixel count without physically shrinking the pixel is to capture all three required wavelengths for colour imaging from a single pixel. The differential depth of absorption of light in silicon can be used to create a stacked pixel, where blue, green and red spectral intensities can be captured from the same site [8]. Photons corresponding to blue colour and therefore lower wavelength are more likely to generate electrons closer to the surface of the semiconductor. On the other hand, at deep locations, most electrons generated are due to the red colour photons. Therefore, a pixel can be built with three photodiodes stacked on the wafer; one for each colour. However, such pixels require a non-standard CMOS fabrication process and impurities deposition on specific depths.

5.2.3 Windowing and binning

Although the principal reason for increasing the number of pixels is to increase the spatial resolution of the image sensor, other functionalities have also improved as a consequence. One of them is image windowing or digital zooming. Windowing is the capability of cropping an image to a smaller region of interest. This is easily performed on CMOS image sensors due to their flexible addressing scheme. The optical resolution is the same as for the whole image, but it allows for smaller images and the capture of more frames per second.

Increasing the resolution of the sensor can lead to image quality degradation due to low well capacity. When conditions require a greater well capacity, such as to increase the sensitivity in low light scenes, pixel binning can be performed. Pixel binning is the capability of combining the

photodiode charges from different pixels thus increasing the well capacity. For instance, a 2×2 binning would increase the well capacity nearly four-fold at the expense of equivalent resolution losses.

5.3   Low noise complementary metal-oxide-semiconductor (CMOS) image sensors

CMOS image sensors are inherently noisy and therefore require correction to produce a noise free image. Any noise in the pixel diminishes its intensity and contrast differentiating ability, as the sensor would only be able to faithfully record contrasts higher than its inherent noise. The captured image from these sensors suffers from temporal noise due to thermal, flicker and shot sources. In an APS, thermal noise originates from the reset and charging of photodiode capacitance with a small contribution from the readout circuitry [9]. Shot noise has its origins in the inherent noise in the DC power supply as well as in the photon shot noise present in the incoming light. The latter is generally considered to be the noise limit for any imaging system. Furthermore, digitisation of the pixel response also introduces quantisation noise to the image. The image sensor is also prone to interference from digital parts of the chip.

In addition to the temporal noise, image sensors also suffer from non-temporal and spatial fixed pattern noise (FPN). The desire for higher resolution has led to significant reduction in pixel size. However, the majority of the pixel area has to be devoted to the photodiode to increase the amount of light captured and therefore the quantum efficiency. This often leads to the use of the smallest geometry transistors in a pixel, that are prone to high mismatch due to process imperfections. The small size of the pixel also means that layout techniques used to reduce mismatch in typical circuits, such as fingering and common centroids are impractical in pixels.

The resultant variations between the characteristics of individual pixels with additional variations from the readout and the interconnects lead to differences in the response of the image sensors to a uniform stimulus. This leads to the presence of a spatial but temporally fixed pattern noise in images acquired by these detectors [10, 11]. Figure 5.3 shows the image of a laboratory scene captured before and after FPN correction illustrating the effects of FPN in real life scenes. Columnar strips of noise due to the inter-column variations are also visible in the image. A prominent noise observable is that of columnar strips. More importantly, human subjects have better perception of fixed pattern noise than temporal noise [12] and it is therefore of paramount importance to reduce this to less than the contrast ability of the human eye of 1%.
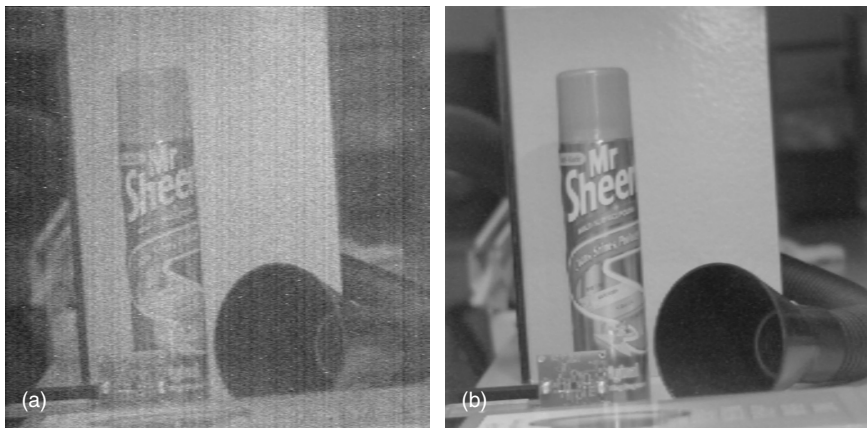
5.3 (a) An image showing the effects of FPN, captured by Fuga camera from Fillfactory and (b) the same image after correction for FPN.

5.3.1   Noise sources

To further understand various sources of noise in a pixel, let us visit the signal chain in a typical APS with a conventional photodiode shown in Fig. 5.4. The photodiode captures the light and in general is the largest physical structure in a pixel. The current generated in this diode consists of the thermally generated leakage current [13] and an optical term produced by the photon falling on diode surface [11]:

$$I_{PD} = I_{therm} + Q_D G_A A_D G_L L_O \qquad [5.2]$$

where Q_D is the quantum efficiency of the material used to make the photodiode, G_A is the gain factor related to the photodiode area compared to the total area of the pixel, and thus accounts for the amount of light captured by the photodiode or lost due to non-photosensitive parts of the pixel, A_D is the area of the photodiode, G_L is the gain of optical assembly and L_O is the input intensity.

The dark current I_{therm} depends on the diode area and often its geometry. Therefore, any variation in the photodiode area leads to fixed pattern noise, which will affect the performance of the image sensors in low intensity images. These variations are often referred to as the dark current non-uniformity (DCNU) in order to differentiate it from intensity dependent fixed pattern noise. It is also worth noting that the diode current is temperature dependant, which leads to changes in FPN values with temperature and in general the FPN is expected to increase with any increase in temperature [14].
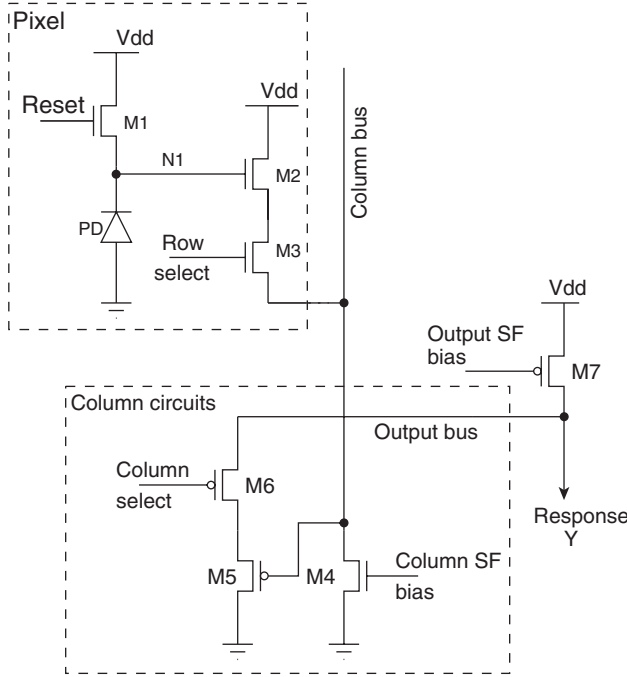
5.4 A conventional CMOS APS pixel with typical column circuitry to enable two stage addressing mechanisms [11].

At the start of the frame, the diode voltage is reset by closing the switch M1. During this time, the temporal noise in the diode is a white Gaussian process with two sided power spectral density of [9]:

$$S_I(f) = q i_D \ A^2/Hz \tag{5.3}$$

where i_D is the diode current. A similar expression is valid for the noise in the transistor, M1, due to noise in its subthreshold current. As a rough approximation of achieving steady state during the reset process, the average reset noise power can be obtained by integrating the two noise sources in a circuit where the transistor channel resistance charges the diode capacitance, C with this noise. The reset noise power obtained can be:

$$V_n^2 = kT/C_{PD} \tag{5.4}$$

where k is the Boltzmanns constant, T is the absolute temperature and C_{PD} is the capacitance of the diode. However, steady state is generally not achieved during the small reset time and thus reset power can be approximated to be $kT/2C_{PD}$. After a small reset, the diode is disconnected

from the power supply V_{dd} by switching off the transistor M1. This allows the diode capacitance to be discharged by the photocurrent. The principal source of temporal noise during this region of operation is the shot noise due to diode current, which is often very small compared to the signal. The noise power is again inversely proportional to the diode capacitance.

After a fixed integration time, pixel voltage is readout using the in-pixel source follower (M2), row-select switch (M3), column source follower (M5) and column-select switch (M6). Using a first order model, the output of the image sensor after two stages of readout can be expressed as [11, 15]:

$$V_Y = V_{dd} - V_{T,M1} - \frac{t_{int}}{C_D}(I_{dark} + Q_D G_A A_D G_L L_O)$$

$$- \sqrt{\frac{\beta_{M4}}{\beta_{M2}}}(V_{GS,M4} - V_{T,M4})\{V_{T,M2} + V_{T,M5}\} \qquad [5.5]$$

$$- \sqrt{\frac{\beta_{M7}}{\beta_{M5}}}(V_{GS,M7} - V_{T,M7})$$

where $V_{T,Mx}$ is the threshold voltage of x transistor, C_D is the diode capacitance and other terms have their usual meaning. The readout circuits add flicker noise as well as thermal noise to the pixel output due channel resistance in the switches.

Terms related to transistor M7 are common to all pixels, while those related to transistors M4–M6 are common to all pixels in a column. It is well known that in small geometry devices, the threshold voltages as well as the transconductance parameters vary from transistor to transistor due to a number of factors, principal among which are the lateral diffusion of the source and drain implants and the field oxide encroachment in the metal-oxide-semiconductor (MOS) channel [16]. Additional sources of variations include local mobility fluctuations, oxide granularity, oxide charge, gate dielectric thickness, device orientation and shape. The variations in the threshold voltage, V_T and transconductance, β of the transistors have been observed to be Gaussian distribution with zero mean and a variance dependent on the device area [16]:

$$\sigma^2(\Delta V_T) = \frac{A_{V_T}^2}{W_{eff} \cdot L_{eff}} \qquad [5.6]$$

$$\frac{\sigma^2(\Delta\beta)}{\beta^2} = \frac{A_\beta^2}{W_{eff} \cdot L_{eff}} \qquad [5.7]$$

where, A_{V_T} and A_β are process dependent terms, while W_{eff} and L_{eff} are the effective width and length of the transistor, respectively.

From Equation 5.5, it can be observed that any mismatch in transistors will lead to variations in the response of a pixel to uniform light. Furthermore, transistors M4–M6 are shared by all pixels in a column and therefore they will only introduce column-to-column noise. With column parallel analogue to digital converters (ADCs) used to increase the speed of operation of image sensors, mismatch due to column readout circuits are removed; however, analogue components of the ADC such as the comparator introduces new sources of FPN. With ADCs used to increase the speed of operation of image sensors, mismatch due to column readout circuits are removed; however, analogue components of the ADC such as the comparator introduce new sources of FPN. In addition to these first order models, there are other noise sources including higher order effects such as the body bias effects as well as switch resistances which affect the performance of CMOS image sensors. Furthermore, sources of FPN are also dependent on the temperature and therefore the net FPN is temperature-dependent as well [14, 17].

5.3.2 Noise correction circuits

In principle, FPN can be reduced by improving the manufacturing process. However, this is very costly and impractical, if not impossible. Circuits to improve the variations after manufacturing have been proposed by using floating gate transistors or circuits to compensate for mismatch [18, 19] with limited results. Consequently, post-processing techniques have been devised to correct for the FPN as well as temporal noise after an image has been acquired. FPN in APSs can be modelled with an additive and a multiplicative term:

$$Y_i = O_i + G_i X_i \qquad [5.8]$$

Here, the subscripts identify an individual element of the array and the parameters can be identified using equation 5.5 as:

$$O_i = V_{DD} - V_{T,M1} - \frac{t_{int}}{C_{PD}}(I_{dark}) -$$

$$\sqrt{\frac{\beta_{M4}}{\beta_{M2}}}(V_{GS,M4} - V_{TM4}) - \{V_{T,M2} + V_{T,M5}\} \qquad [5.9]$$

$$- \sqrt{\frac{\beta_{M7}}{\beta_{M5}}}(V_{GS,M7} - V_{T,M7})$$

$$G_i = \frac{t_{int}}{C_D} Q_D G_A A_D G_L \qquad [5.10]$$
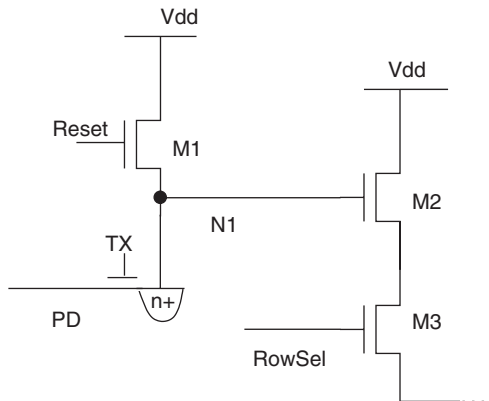
5.5 An APS with buried photodiode to support correlated double sampling.

Furthermore, it has been observed that the additive noise is dominant in APSs and its correction is often sufficient to obtain good quality images. The additive noise sources are also present in the reset level of the pixel. Therefore, recording the reset level and subtracting it from the response of the pixel after the integration time leads to removal of a significant portion of the FPN. More importantly, the principal temporal noise source is that of reset noise, which is additive to the signal. Therefore, the reset level and the integrated output are correlated with regards to the temporal noise as well. Their difference will consequently remove most principal sources of noise.

The preferred approach to record reset levels is to store it in a memory for subtraction from the pixel's response after the integration period leading to correlated double sampling (CDS) [20]. A widely used technique to store the reset level of the pixel is to use a pinned or buried photodiode, which provides a capacitance inside each pixel to store the reset value. One such example of the pixel is shown in Fig. 5.5.

The modified pixel operates in a similar fashion to the conventional diode pixel, except that node N1 holds the reset charge, while the switch TX separates it from the photo-generated charge in the diode. During the readout, the reset levels are first readout followed by the signal level by using the switch TX. Banks of analogue memory can be used in the column circuits to store these two or they can be directly digitised for simpler difference operation in the digital domain. It is worth noting however, that with reduction in the feature size of pixels, multiplicative fixed pattern noise sources have been increasing requiring further correction. This is often achieved using an offline correction in the digital domain.

The correlated double sampling, however, is not suitable for pixels requiring continuous operation. Scene based fixed pattern noise corrections have therefore been suggested to reduce the noise of such sensors. There are two broad class of approach, based on the assumptions being made about the scene. The first group of techniques make a statistical assumption about the scene and utilises it to correct for variations. For example, if it is assumed that the statistics of observed scenes, primarily the mean and standard deviation, are constant with a significantly large number of scenes, then the mean and standard deviation of the responses of individual pixels to these large numbers of scenes would become their offset and gain, respectively, of the pixel [11]. These techniques, however, require exposure to several scenes and their continuous operation can introduce ghosting after-effects due to expectation of nonstationarity of the objects. Alternatively, global motion between frames captured by an image sensor has also been used to reduce the FPN [21, 22]. Herein, an accurate estimation of motion and precise correspondence of all pixels in any motion sequence is used to determine individual parameters of each pixel, which can be used to correct for variations. It is worth noting, however, that such techniques would be of limited use in the absence of motion in the field of view of all pixels. They also suffer from high computational complexity though faster correction than statistical techniques.

An APS with correlated double sampling is able to effectively reduce the reset noise, which is the largest source of noise in the circuit. With the reduction of reset noise, flicker noise in the readout circuits, particularly the source follower, becomes predominant. Flicker noise can be minimised by suitably selecting the size of the source follower device. In a number of designs, the pixel output is amplified by a column amplifier before digitisation. These stages of amplification and digitisation introduce further noise of their own. However, a high gain column amplifier will limit the effect of noise introduced on the later stages. To better appreciate this, let us consider a block diagram of the signal chain as shown in Fig. 5.6.

The input referred noise can be expressed as:

$$V_{n,i}^2 = \frac{1}{G_s} V_{n,i}^2 + \frac{1}{G_c} \left(V_{n,i}^2 + V_{n,i}^2 \right) \tag{5.11}$$

With a source follower gain of less than 1, it can be appreciated that the effect of later stages can be reduced by using a high gain column amplifier. This arrangement may suffice for low signal levels. However, a high gain column amplifier may saturate the output for high signal levels. Therefore, it is advisable to have two or more column amplifiers with different gains. The high gain amplifier may be used for low output levels, while high signal levels should be measured through a low gain amplifier. High
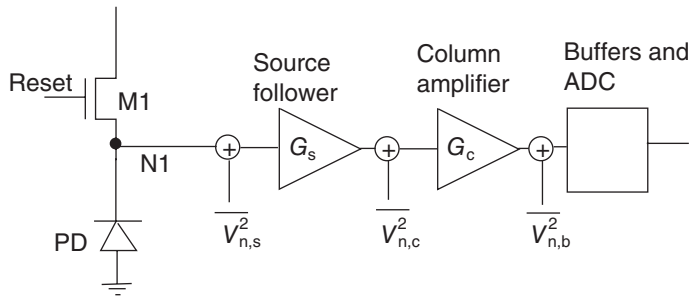
5.6 Noise sources in the signal chain of an image sensor.

signals already have a high SNR so additional noise in later stages does not affect them.

The noise can be further reduced by averaging the signal over a period of time. While continuous averages are difficult to achieve, one may provide for multi-sampling of the output signal to further reduce the temporal noise to acceptable levels. Use of column parallel ADCs as in high speed image sensors can be used to reduce noise at higher frame rates. Different ADCs have different noise performance and should be considered when computing the noise performance of the sensor. By applying all of these techniques, image sensors have been reported which claim to have sub-electron noise floors.

5.4   High speed image sensors

Increasing the speed of frame capture is valuable in several applications including those in automotive, scientific, military, medical and entertainment industries. High speed image sensors are also required for machine vision, high-definition video (HDTV, UHDTV) as well as three dimensional imaging. Applications demanding speed spanning from hundreds to millions of frames per second have been reported. Furthermore, spatial resolution requirements also vary according to the application. Therefore, the speed of an image sensor is usually expressed in throughput units of pixels or bits per seconds (pix/s or bps). High speed image sensors have been designed by reducing the congestions or removing the various bottlenecks in the signal chain of an image sensor.

5.4.1 Signal chain

Light falling on several pixels is converted into a train of digital output by an image sensor. Most, if not all image sensors are essentially systems with a large number of parallel inputs and one or few outputs. An ideal and very

fast image sensor would function like the human eye where a large number of parallel outputs are recorded from the sensor (the retina). However, with two dimensional sensor design with limited three dimensional integration, if any, digital image sensors have to multiplex the input signal through circuits inside each pixel as well as in each column to enable sequential readout of output data. This multiplexing introduces a fundamental limit to the speed of image sensor.

This also means that the interface to the image sensor should work at speeds higher than the product of the number of pixels and the speed of an individual pixel. This is comparatively simple to achieve on account of improvements in the transfer of digital data. It, nevertheless, introduces a bottleneck at the conversion stage between analogue pixel output and digital sensor output. Fig. 5.7 shows various architectures utilised to reduce the effect of this bottleneck.

Early image sensors including CCDs and early CMOS image sensors provided the output through a single analogue port. Pixel values were read sequentially and buffered for further processing outside the image sensor chip itself, the first of which was conversion to digital data using an external ADC. One of the earliest improvements to this design was an ADC inside the sensor die [23] as shown in Fig. 5.7(a). This integration of an ADC was also one of the leading advantages of CMOS image sensors over CCDs. Although the readout process is still sequential, the input capacitance of an on-chip ADC is much smaller than an external chip, which leads to faster settling times and therefore fast transfers and higher speeds.

An extension to this design which leads to further increase in the speed of the image sensors has been the incorporation of an ADC in each column of the sensor as in Fig. 5.7(b). Ideally, one should use a dedicated ADC for each column [24]; however, different ratios of the columns per ADC have been implemented as well [25]. This introduces digital parallelism into the readout signal chain, speeding up the overall sensor operation. Such schemes are therefore referred to as column pixel sensors.

Further improvements on the throughput of all the above architectures can be achieved by allowing for multiple parallel outputs. The sensor array can be split into top and bottom halves with readout circuitry being placed on both sides [26]. A greater number of outputs can be implemented by partitioning the pixel array even further depending on the desired throughput, power consumption, location of system bottleneck and manufacture costs.

Following the trend of pushing up the signal conversion to as early as possible, ADCs have also been implemented inside each pixel as in Fig. 5.7(c). This allows for higher throughput since the parasitic load of other pixels is removed from the analogue signal chain. However, pixel level ADCs need to be as small, simple and power starving as possible for area
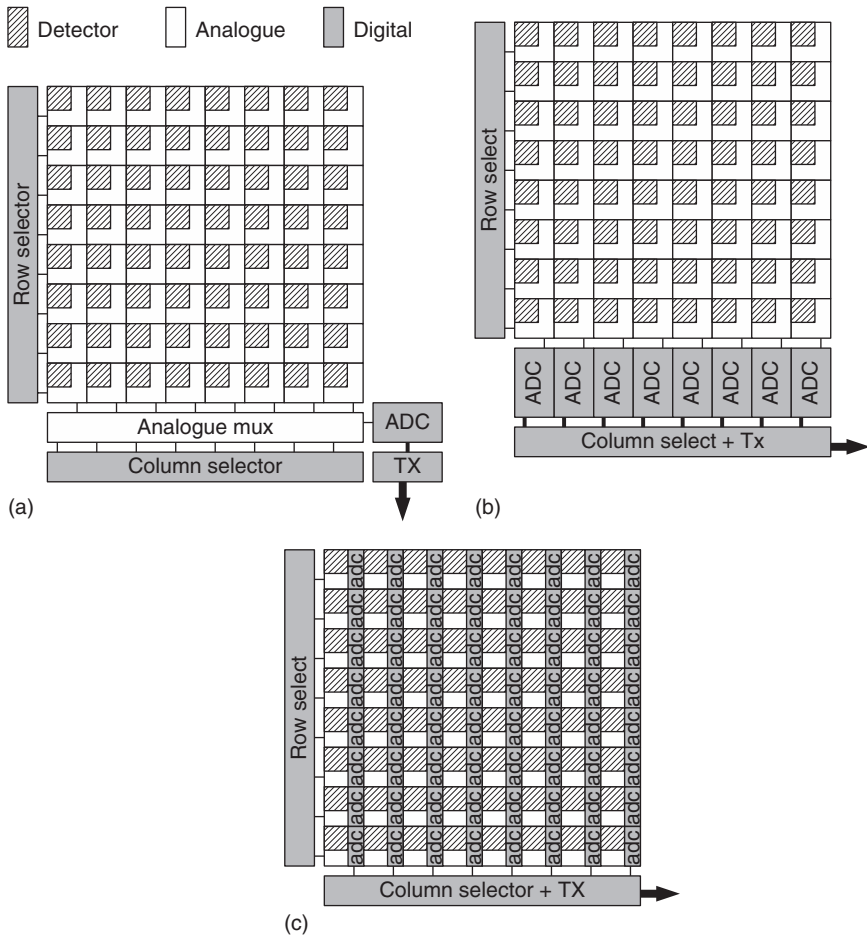
5.7 Different architectures to increase speed. (a) Typical image sensor with one ADC for the whole array and with analogue multiplexing of pixel's output, (b) column parallel design with an ADC per column to reduce the signal delay and (c) architectures with ADC inside each pixel to further increase speed.

and noise considerations. Considering the reducing pixel pitch on account of the demands of the higher resolution, this is often very difficult to achieve leading to some compromise with either the resolution or the fill factor and hence the quantum efficiency. Sensors with ADCs inside each pixel have been referred to as digital pixel sensors and often integrate only parts of an ADC inside each pixel, leaving common parts to be implemented at the end of the column or shared among a number of pixels.

5.4.2 Digital conversion and output

In addition to the placement of the signal conversion stage, the methodology used for the signal conversion also affects the speed of image acquisition and hence that of the sensor. A number of ADC topologies have been proposed and used in image sensors.

Early CMOS image sensors preferred the use of successive approximation register (SAR) ADC architecture. Although not as fast as flash or pipeline ADCs, it offers simpler implementation. A SAR ADC, as shown in Fig. 5.8(a), works by first sampling the pixel output and then changing the comparator reference accordingly. The reference is set using a digital-to-analogue converter (DAC) whose input is a digital register. The conversion covers all the N_b register bits, from the most significant bit (MSB) to the least significant bit (LSB). Therefore, a conversion lasts for approximately $N_b X t_{DAC}$ seconds, where t_{DAC} is the DAC settling time. SAR ADCs can also be used as a column ADC [24], although its large area due to the capacitor scaling, shown in Fig. 5.8(b), is a limitation. A 14-bit SAR ADC (with minor modifications) was reported to finish a conversion in 1.7 µs [26]. This can be contrasted with the needs of a UHDTV-2 image sensor, which requires every row readout to be completed in 1.9 µs.

Figure 5.9(a) shows yet another popular ADC topology used in image sensors [27]. The ramp or single-slope ADC (SS ADCs) normally provides a slower conversion rate; however, it requires a small area and generally provides improved linearity characteristics [28]. While in the SAR ADC the reference is adjusted accordingly to the register value; in the SS ADC, the reference applied to the comparator is a single-slope ramp [29].
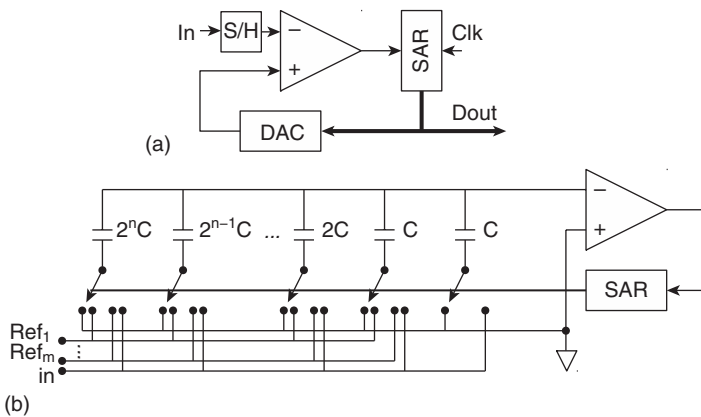
5.8 (a) A block diagram of successive approximation ADCs and (b) a typical circuit implementation.
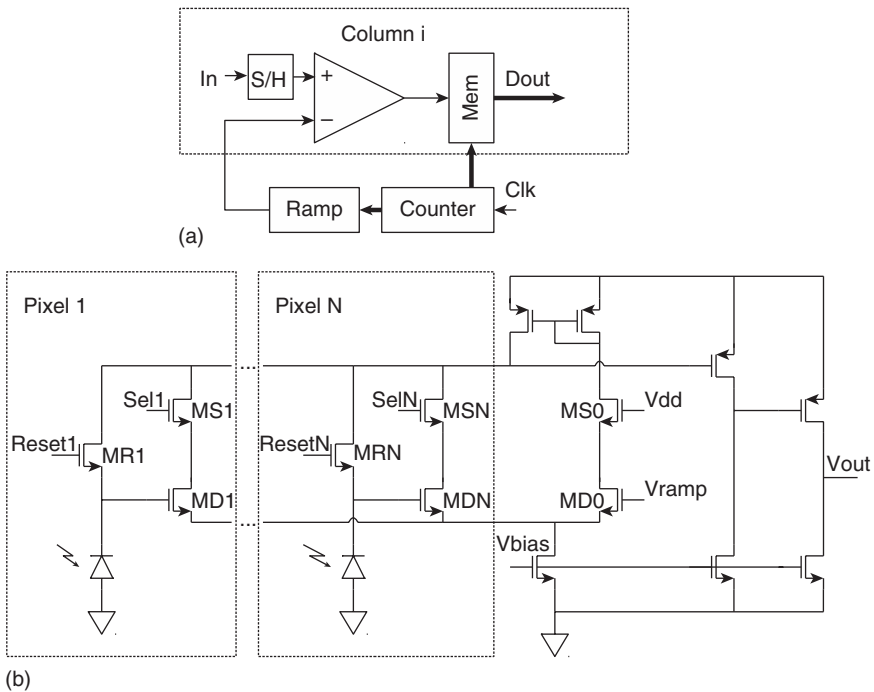
5.9 (a) A block diagram of ramp ADC and (b) a typical circuit implementation.

Simultaneously, a counter is increased at a fixed rate until the instant the ramp crosses the reference value, locking the counter. Therefore, the counter final value is a function of the sampled input. In another approach, a fixed reference is used and the sample input is integrated, generating a variable rate single-slope ramp. The conversion speed is then a function of the counter's clock frequency $f_{clk} = 1/T_{clk}$ and is given by $2^{Nb-1} \times T_{clk}$ seconds. Such designs tend to better support column parallel architecture, as the ramp generation circuitry can be made external to the columns thereby reducing the size requirement per column. Figure 5.9(b) shows a typical example of such an implementation [30]. Ramp ADCs have also been used as in-pixel ADCs. This is again due to the possibility of designing the sensor with a small number of the ADC's components required inside each pixel while the bulk of the ADC can be shared between neighbouring or column pixels.

Another alternative for signal conversion is to use time-based ADCs [31], the majority of which are based on pulse frequency modulation or pulse width modulation (PFM or PWM). These ADCs use a comparator with a fixed reference and a memory to register the amount of comparator-generated (PFM) or clock (PWM) pulses. In addition to these standard

ADC architectures, high speed image sensors have also used variations of the aforementioned topologies such as algorithmic/cyclic ADCs or dual-slope/multiple-slope ADCs [29]. Others implement different ADC topologies including sigma-delta converters[32].

The final part of the signal chain in an image sensor is the local storage and subsequent transmission to an external device. The optional former can be achieved in a chip with local dynamic random-access memory (DRAM) or static random-access memory (SRAM) [30]. Serial readout of the imaging data has been used in a number of sensors; however, parallel readouts offer higher throughput [33]. Signalling standards developed for high speed data transmission such as low voltage differential signalling (LVDS) [34] and scalable low voltage signalling (SLVS) [35] have been used in image sensors. Both of these utilise a pair of wires for each bit transmitted. A fixed current (typically 3.5 mA) is injected on one of the wires (according to the bit value) and a matching resistance provides the receptor with a low voltage (350 mV) signalling the bit value.

Signal conversion to digital domain increases the overall speed of an image sensor; however, there are applications where even the pixel level conversion fails to meet the speed specifications. One solution for such extreme high speed applications is to implement analogue memories to provide temporary frame storage [36]. Although the overall average throughput remains the same, the use of these memories allows for a burst mode operation, where a small number of frames can be fast captured and slowly readout. These buffers enable a burst mode operation even up to 1 tera-pixels/s for 128 frames against a 780 Mpixels/s on continuous mode.

5.4.3 Pixel improvement

As for the pixel itself, the photo-detector area for high speed applications is usually larger than for high-resolution digital photo-still image sensors. The large pixel size is required to allow a measurable charge integration during the short available integration time [28].

In very fast applications, such as time-of-flight 3D vision, it is often necessary to change the photodiode diffusion vertical profile to allow for fast charge transfer from the photodiode to the sensing node (floating diffusion) [37]. Furthermore, the layout of the photodiode can also result in charge transfer improvements. Examples include T-shaped detectors [38] and horn shaped photodiodes [37] with reported charge transfer times which are 500 times faster than a rectangular shape. In such applications, it is convenient to not only achieve fast transfer of charge but also some form of in-pixel computation, as the readout time may not be fast enough. For example, the time of flight has been calculated by adapting the photodiode diffusion gradient [39]. This reduces the charge transfer time from the

5.10 A scene showing the effects of rolling shutter on fast moving scenes.

photodiode to the floating diffusion but also provides two separated sensing nodes. Sequential charge transfer to these sensing nodes and using the difference of charge provides a simple estimate of time of flight.

5.4.4   Global shutters

CMOS APS is inherently a rolling shutter pixel system, where every row is read and reset in turn. This, however, introduces smear in fast moving scenes as shown by the example in Fig. 5.10. A number of alternative CMOS pixels have been proposed which are inherently globally shuttered [40]. It is also possible to add additional switches and diffusions to obtain global shutter in an APS.

Figure 5.11 shows a 5-transistor circuit which provides a global shutter by adding an additional diffusion node [41]. The circuit provides a global reset through the TG switch for all photo-diodes. At the end of the integration frame, the integrated charge is transferred to the readout node through the switch TX. However, this disables the correlated double sampling and is therefore is prone to temporal noise. The addition of further switches and diffusions can be made to introduce correlated double sampling in this pixel [42, 43]

5.5   Low power image sensors

The large majority of image sensors are used in handheld or mobile devices. These include digital still cameras, video cameras or mobile phone cameras.
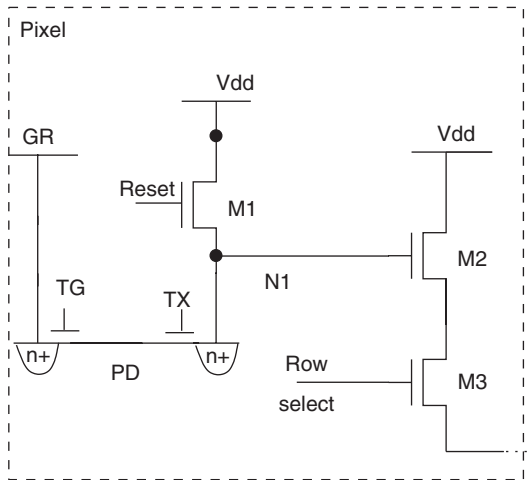
5.11 A five-transistor global shutter circuit.

In these applications, battery life is limited and therefore image sensors are expected to consume as low power as possible [44]. Furthermore, the demands of multi-mega pixel image sensors in even mobile phone cameras, puts further stress on the power dissipation of image sensors. In addition, there are applications such as a camera in a pill, cameras in space or energy scavenging cameras, where the demands for lower power usage are even more stringent. CMOS image sensors offer two to three orders of power reduction over CCDs, a fact which was crucial in their rapid success. However, new applications as well as increasing demand from existing applications for better battery life has led to even more aggressive power reduction. Furthermore, packaging and cooling costs also place demand on power supply management even when power itself is not a concern. It is also worth noting that any increase in the chip temperature of an image sensor increase the leakage in the photo-diode. This in turn leads to poorer low light performance. Therefore, the image sensors need to manage their power budget effectively.

One approach to reduce the power dissipation of any CMOS chip is design and built in lower dimension processes with lower power supplies. Technology and voltage scaling has indeed been one of the driving factors behind the reducing power cost of image sensors. CMOS imagers can benefit from technology scaling by reducing pixel size, increasing resolution and integrating more analogue and digital circuits on the same chip with the sensors. However, being primarily an analogue circuit block, they have lagged a few generations behind the state of the art digital technology, as image sensor design often requires thick gate transistors, high threshold voltage transistors, additional poly layers

and proper body biasing. Furthermore, implementation of a low power image sensor in lower dimension processes is challenging on account of reduced photoresponsivity, dark current and increased leakages. In addition, technology scaling with aggressive supply voltage reduction affects SNR, leakage current as well as the dynamic range of an image sensor. Technology such as silicon-on-insulator (SOI) have also been suggested to produce low power image sensors by reducing leakage current [45, 46]

Alternative approaches at architecture or algorithmic level are possible to reduce the power dissipation of image sensors [47]. Particularly, with application specific image sensors, one has the option of suitably defining the capture and processing algorithm to reduce the power dissipation, similar to those of digital integrated circuits. This may be achieved by shutting down parts of image sensors. For example, if a region of interest is well defined, one may capture the image from this region alone and shut other regions to reduce power consumption. Furthermore, one may sample different regions of an image sensor at a different rate to reduce the power consumption. As an example, the central field of view may be sampled more frequently than peripheral regions if one has prior information of more activity in the centre. Previous frames can be used to predict any region of interest or even the whole frame. Efficient partition of computation between analogue and digital blocks can further optimise the power dissipation. Finally, one can always transfer power hungry functions away from the image sensor chip.

Circuit modifications can also be undertaken to reduce the power dissipation in the image sensor [48]. Reduction of any leakage including that in digital circuits will reduce the overall power consumption of the image sensor. One can introduce 'sleep' and 'active' modes of operation by suitably placing switches to minimise power consumption of inactive regions of the chip. Multi-threshold voltage circuits or dynamic threshold circuits can also be used to further reduce power dissipation in image sensors [49]. Pixel level ADCs can be used to achieve pixel-level parallelism, which in turn can reduce power dissipation [50]. Another attractive analogue design technique to reduce power consumption is to operate some if not all transistors in the subthreshold region [50, 51]. Pulse width modulation by directly encoding the photocurrent into free running pulses has also been shown to reduce power dissipation [52, 53]. PWMs offer the advantage of being applicable in even low voltage technologies as they are generally independent of power supply.

5.6   Wide dynamic range sensors

A typical CMOS APS as well as CCDs have a limited dynamic range of 40–70 dB, which is in turn due to their limited well capacity. However,
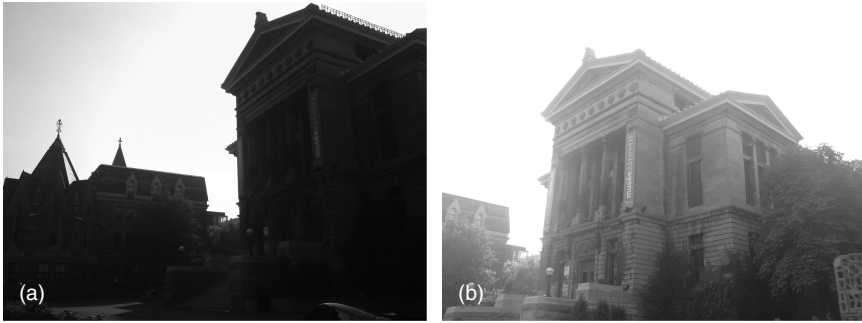
5.12 An image showing the effects of limited dynamic range of active pixel sensor. (a) Shows loss of dark information with a short integration time and (b) shows saturation in brighter areas with high integration time.
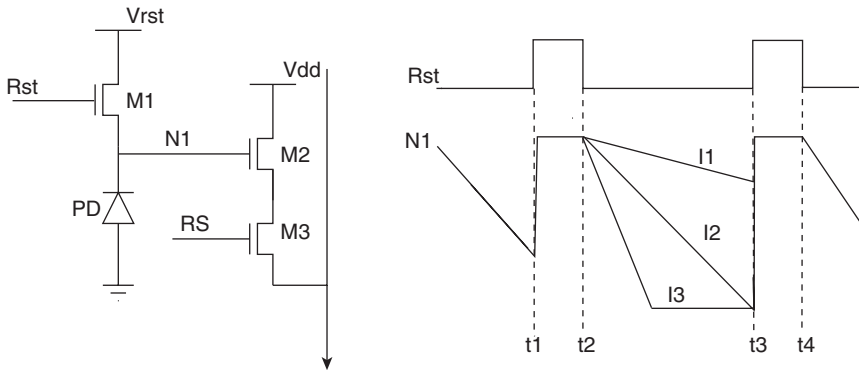
5.13 APS and its operation.

illumination available in nature and captured by the human eye spans over 120 dB of intensity variation. Even typical real world scenes can have dynamic intra-scene ranges extending to five orders of magnitude, from 1 lux in shadow to $1-10^5$ lux of bright sunlight. The limited dynamic range of the CMOS APS and CCDs therefore can lead to disappointing and often disastrous results in a number of captured scenes with a typical example shown in Fig. 5.12. A number of circuits have therefore been proposed and used for extending the dynamic range of CMOS APSs as well as for new modes of imaging leading to capture of a wide dynamic range of intensities [54, 55].

Before discussing these approaches, let us revisit an APS to understand its limited dynamic range. Figure 5.13 shows an APS and a typical signal diagram for three different photocurrents. For very high input intensity, I3,

and for all currents higher than I2, the pixels saturate to a constant output. This limiting level of photocharge is often referred to as the saturation charge, Q_{sat} or the well capacity of the pixel.

The lowest intensity which a pixel can record is determined by the thermally generated leakage current of the photodiode. These two values determine the dynamic range of the pixel. In addition to these two values, the dynamic range of a pixel is also affected by the residual temporal noise. It is also worth noting, however, that the lowest and the highest level of intensities faithfully recorded can be changed by altering the integration time of the pixel. However, this does not change the dynamic range as any change in integration time changes both levels by the same factor.

5.6.1 Logarithmic sensors

A simple technique to capture the wide dynamic range of intensities is to use a logarithmic amplifier inside each pixel to compress the input photocurrent. Figure 5.14 shows a simple implementation of this technique using an nMOS transistor in weak inversion to build the required logarithmic amplifier.

The resulting circuit is similar to that of the APS and shares the same readout mechanism of source-follower and switch. Transistor M1, however, now operates in weak inversion and converts the photocurrent into a logarithmic voltage:

$$V_{S,M1}(or V_{G,M2}) = V_{G,M1} - V_{T,M1} - \frac{nkT}{q}\log\left(\frac{I_{DS,M1}}{I_{DSO,M1}}\right) \qquad [5.12]$$
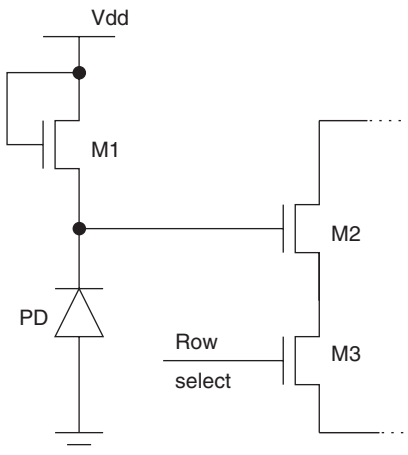
5.14 A typical logarithmic pixel using an nMOS in weak inversion.

where $I_{DSO,M1}$ is the current flowing in the device M1, when the gate-source voltage of the device equals its threshold voltage and therefore the device is making its transition from weak inversion to moderate inversion, q is the electronic charge constant and n is the subthreshold slope.

This circuit has the advantage of simultaneous capture and compression of intensities, thereby reducing complexity in later chains of readout including the number of bits required for data conversion. Such pixels also provide true random addressability on account of their continuous operation. However, the continuous operation of these pixels make them highly susceptible to fixed pattern noise, as any double sampling operation is difficult to undertake.

Several approaches have been suggested to reduce the fixed pattern noise by utilising reference values for additional fixed pattern noise correction though with limited success [56–59]. These reference values are often either the dark response of the pixel or its response to a very high photocurrent. Off-chip techniques have also been proposed where rather than generating a reference at every frame, a reference reading from the array illuminated with uniform light or a white paper is stored in an off-pixel memory. Post-fabrication correction for non-uniformity using hot carrier degradation [19] or compensatory circuits [18] have also been proposed. The former suffers from long stressing times and a requirement for high driving signals. The latter generally requires complex circuits inside the pixel which reduces the fill factor and hence the quantum efficiency of the pixel. Figure 5.15 shows
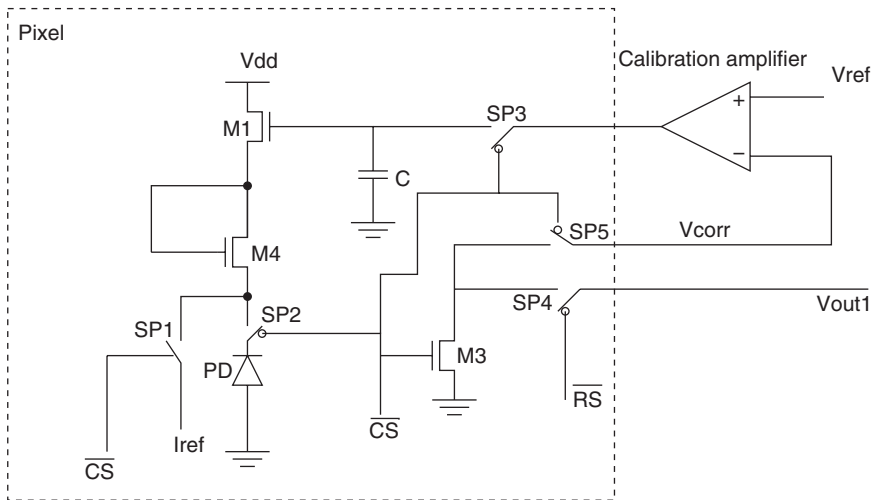
5.15 In-pixel calibration circuit proposed by Loose and co-workers [18].
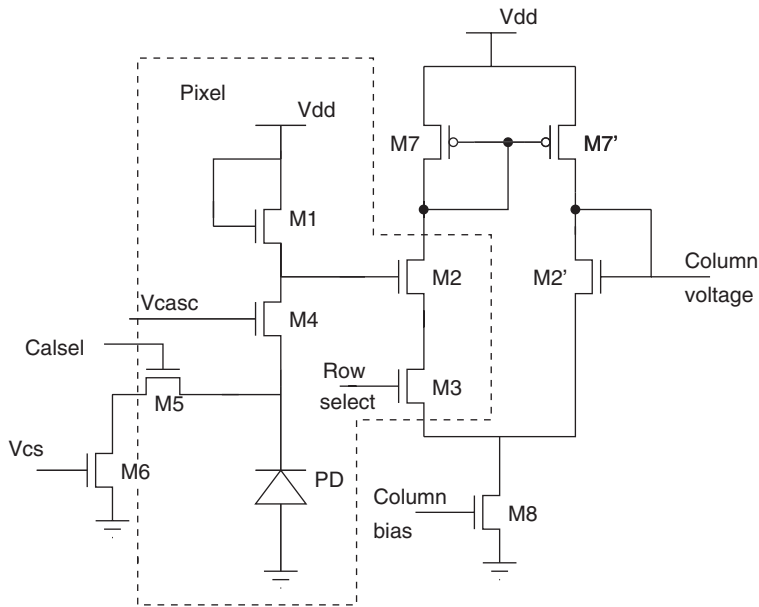
5.16 A logarithmic pixel with one stage of a differential amplifier readout and circuit for electronic calibration [60].

an example of a compensatory circuit, where the gate voltage of one of the load transistors if adjusted during the calibration process by using external reference voltages, thereby compensating for variations [18].

Most of these techniques only reduce the additive fixed pattern noise; however, the logarithmic pixel has small gain on account of small sub-threshold slope of the load device. This means that even a small multiplicative fixed pattern noise will reduce the image quality [11, 60]. Therefore, logarithmic pixels often require multi-parametric calibration to significantly reduce the fixed pattern noise. Figure 5.16 shows one pixel circuit which can correct for two or even three parameters leading to contrast performance matching the human eye. In a two-parametric calibration scheme, two well-spaced reference currents are used to record the response of the pixel and to extract its parameters. These parameters are then used to correct for individual variations in the pixel's response to input light. The technique can be further enhanced to compensate for leakage current variations by using the dark response of the pixel.

5.6.2 Pixels with combined linear and logarithmic response

Logarithmic pixels suffer from limited gain, particularly in low light regions. Furthermore, transition from high light to low light is problematic on
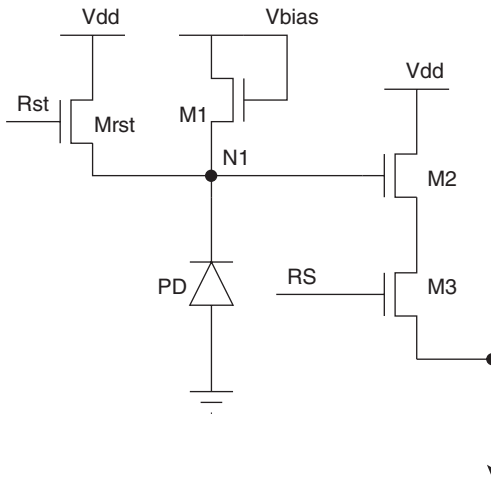
5.17 A pixel with combined linear and logarithmic response.

account of the long settling time of these pixels. Circuits have therefore been proposed which aim to utilise the fast settling time of linear pixels and the wide dynamic range response of logarithmic pixels. A simple approach is to use an active pixel sensor as a linear pixel in one frame and as a logarithmic pixel in second frame. These two outputs can then be merged using a threshold voltage to separate the two response regimes [61]. Alternatively, a smooth transition can be achieved between the two regions of operation in a single frame by designing a pixel which incorporates a logarithmic device as well as a reset switch as shown in Fig. 5.17. [13, 62].

The pixel operates as a linear pixel at the start of the frame. It is reset using the Mrst device; however, the node N1 is pulled high enough to ensure that the transistor M1 is switched off. During the integration, the photocurrent will then discharge this node to create a linear response. However, eventually this node will fall below the gate bias voltage for the transistor M1, which will switch this device on. An equilibrium is eventually achieved when the transistor M1 operates in weak inversion and pixel output voltage becomes proportional to the logarithm of the photocurrent. The pixel output is measured after a predetermined time after the reset pulse. If the input light is low, the pixel remains in its active pixel sensor mode and a linear response is achieved, as shown in Fig. 5.18. However, if the light levels are high, a logarithmic output is obtained on account of the device M1. The pixel response curve undergoes a smooth but quick transition between these two regions of operation [63]. Operating in two regions, however, introduces higher fixed pattern noise, particularly in the transition between linear and logarithmic region. A mixture of double sampling as
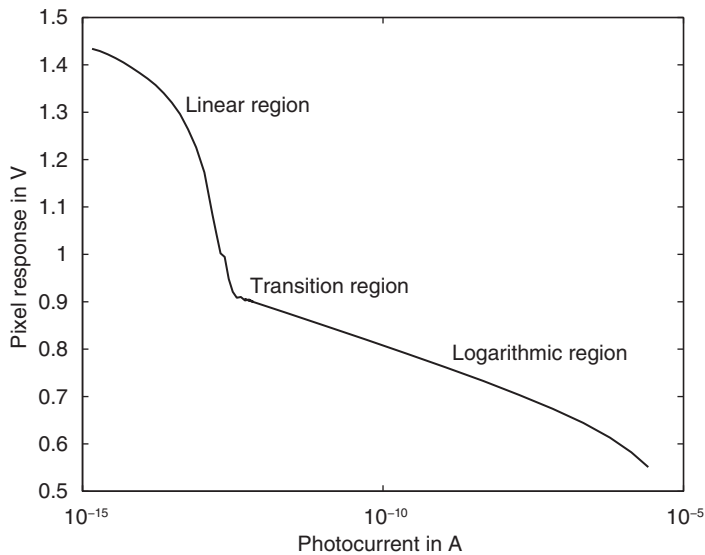
5.18 Response curve of a combined response pixel showing linear response for low light and logarithmic output for high light.

well as calibration can be used to improve the fixed pattern noise of these sensors.

5.6.3 Threshold comparing pixels

It is worth noting that CMOS APSs lose their ability to record wide dynamic range information only because their output is recorded after a fixed integration time and this allows the high intensity photocurrents to be lost at the saturation level. If, however, one can monitor the pixel continuously and automatically, one can record the high dynamic range information. One technique to do so could be to change the pixel output to record the time it takes to reach a certain threshold and use this information to ascertain the light level [64, 65]. High intensity currents will reach this threshold faster than low intensity light levels. To record this time, one can use a high speed clock as shown by the pixel in [66]. Alternatively, one can reset the pixel every time it reaches a threshold and use the number of resets in an integration cycle to record the light levels [67]. The number of reset levels can be recorded in a local memory element inside each pixel. The system, however, suffers as low photocurrents require a long time to reach any meaningful threshold leading to reduction of the frame rate. It is possible, however, to use the linear integrated voltage for low photocurrents and the time to reach a threshold for high photocurrents [64]. Implementation of

such techniques nevertheless requires complex circuits which often have one or more comparators and memory elements inside each pixel, reducing its fill factor and significantly affecting its quantum efficiency.

A modification of threshold comparing pixels are spiking neurons such as image sensors, which have been inspired by biological systems. Pixels have been proposed which generate a spike when a certain voltage threshold is reached [68]. Alternatively, pixels have also been proposed which generate a spike when a threshold is reached and then reset the pixel. The pixel is then allowed to integrate again until the threshold is reached and another spike is generated. The number of spikes generated in a frame is used as a measure of the input light [69].

5.6.4   Integration time control sensors

Controlling the integration time of pixels can provide another approach to increase their dynamic range. There are a number of ways in which this can be achieved. Depending upon the average intensity of a scene, one may change the integration time of all pixels. This does not change the dynamic range of the images captured but allows for a different region of intensities to be captured depending upon the scene [70, 71]. An extension to this principle is to locally adapt the integration time for each pixel or group of pixels [72, 73]. This may be achieved by continuously comparing the pixel output to one or many references and using the output of comparison to generate reset pulses or a few bits of flags to be stored in a memory [73]. Alternatively, one can reset the pixel a number of times in a frame where each individual integration time is different to one another. Use of local memory to record a few bits of information related to pixel output in these smaller integration times can be used to reproduce a wide dynamic range signal. Circuits used to achieve these goals are often similar or identical to an active pixel sensors, however, complex external circuitry often at the end of columns, is used to control the reset signal [72]. Furthermore, use of local memory and need of comparators often leads to higher noise as well as difficulty in the reconstruction of signals.

5.6.5   Threshold comparing as well as integration time control pixels

Combining the two approaches of threshold comparison and integration time control provides another approach to obtain wide dynamic range information [74, 75]. Figure 5.19 shows the signal diagram to reflect the approach. In a linear pixel operation as shown by the dashed lines, high photocurrent, I1, leads to saturation. In this approach, however, the integrated signal is compared to a constantly increasing reference signal
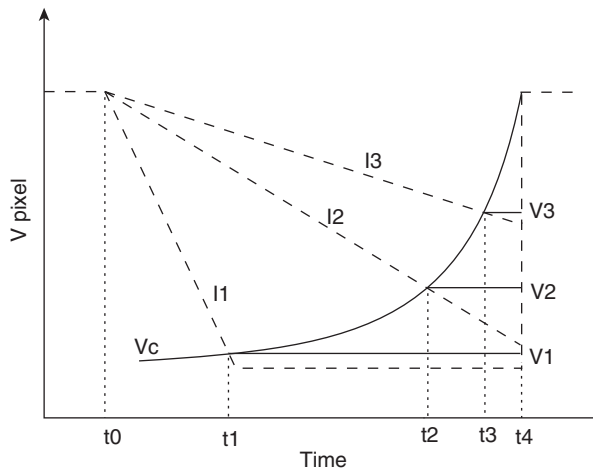
5.19 Signal flow in an integrating pixel with threshold comparison and integration time control.

and the integration is stopped when the two are equal. This means for the high current, I1, the integration will stop at lower voltage V1 than the voltages for lower currents I2 and I3. The pixel output after the integration frame will thus depend on the Vc signal.

Figure 5.20 shows a circuit where a p-type metal-oxide-semiconductor (pMOS) comparator and switch is used to compare the reference value and the integrated value [75]. Alternatively, and preferably, a n-type metal-oxide-semiconductor (nMOS) only pixel is also shown which records the reference value rather than the integrated value [76]. In both of these circuits, the pixel operation starts by resetting it at the start of the frame, by applying a high voltage Rst on the gate of transistor M1, turning it on, thereby placing a high reset voltage on node N1. Once the reset transistor M1 is switched off by lowering its gate voltage, the high voltage placed on the capacitance of node N1 is discharged by the photo-generated charge. Different to an APS, however, an external and monotonically increasing reference signal V_C is simultaneously applied at the drain of transistor M4. At the start of integration, the voltage Vc is lower than VN1 and hence transistor M4 is on. However, discharging VN1 and monotonically increasing Vc leads to a situation when the gate and source voltages of M4 lead to it being switched off. After this time, the gate voltage of M2 is held by its gate capacitance. At the end of the integration time, transistor M3 is switched on by a high RS signal and this held voltage is readout as the pixel's output using the source follower M2.
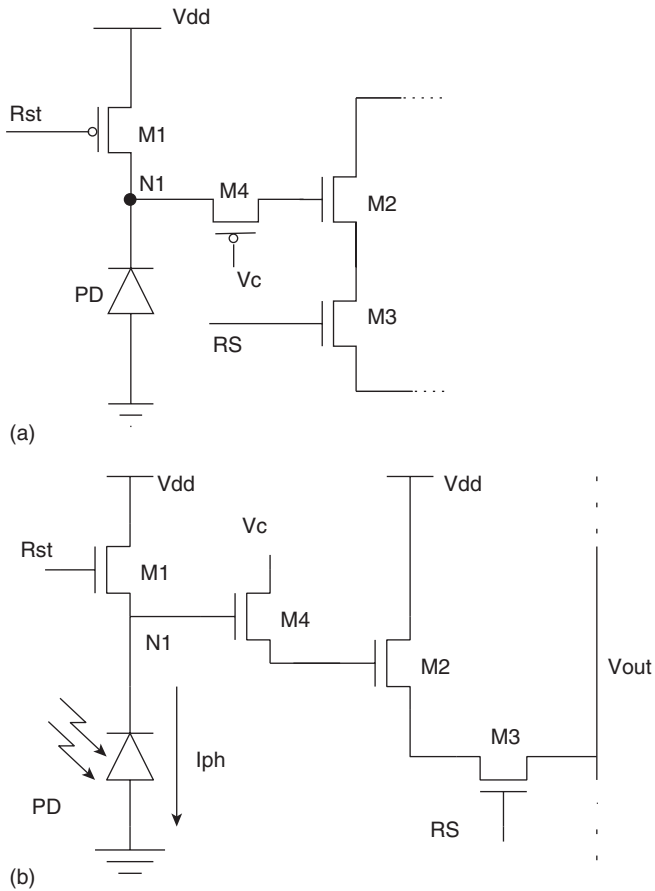
5.20 Four-transistor integrating pixels capable of producing wide dynamic range output of any monotonically increasing transduction function. (a) Uses a pMOS device as switch and (b) uses an nMOS device as switch.

With a suitable control signal Vc, it is possible to record a unique response of the pixel even for higher intensities at which an APS would saturate. More importantly, by changing the reference signal, one can change the transduction function of the pixel. This means that the pixel can produce any monotonically increasing transduction function from the pixel. A simple output can be that of logarithmic or even Steven's power law [77]. However, and more importantly, one can change the transduction function to generate tone mapped outputs which can be directly displayed on to typical screens. Yet another aspect of the pixel rises from the fact that the output is held irrespective of the integration time. This means that all pixels in a large array can be reset at the same time and readout serially after a fixed
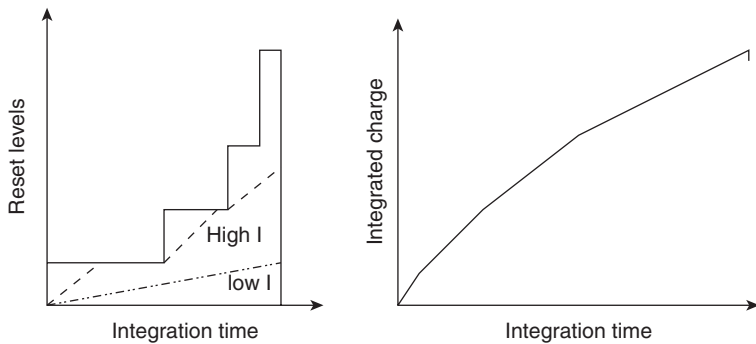
5.21 A stepped reset signal and corresponding collected charge used to enhance the dynamic range of the sensor [78].

integration time. This further means that the pixel provides inherent global shuttering.

5.6.6 Well capacity adjustment

One can also increase the well capacity of the pixel to enhance the amount of charge it can hold and hence its dynamic range. A simple approach would be to increase the operating voltage; however, this will lead to increasing power consumption. With the power supply constant, any charge generated in the pixel after its saturation often spills. This spilling charge has been used first to enhance the dynamic range of pixels in CCDs and later in CMOS [78]. The reset device itself can act to modulate the charge storage in the diode [78]. For example, Fig. 5.21 shows a non-linear stepping of the reset signal in a typical APS which can be used modulate the charge stored in the pixel and increase the dynamic range.

Alternatively, one can independently integrate the spilled charge on a different capacitor and use the integrated photocharge and spilled charge can be used to extract the input light on the pixel [79]. For example, in an active pixel with pinned diode, one can design such that once the diode is saturated, the spilled charge is stored over the floating diffusion. One can also add additional capacitors to further enhance the ability to integrate spilled charge as well as reduce noise [80]. The light to pixel output relationship in such pixels is linear; however, with a sudden change in slope as shown in Fig. 5.22.

5.6.7 Frequency-based wide dynamic range image sensors

Mapping a wide dynamic range of intensities on to a voltage domain for pixel output increases the requirements from the data conversion stage as
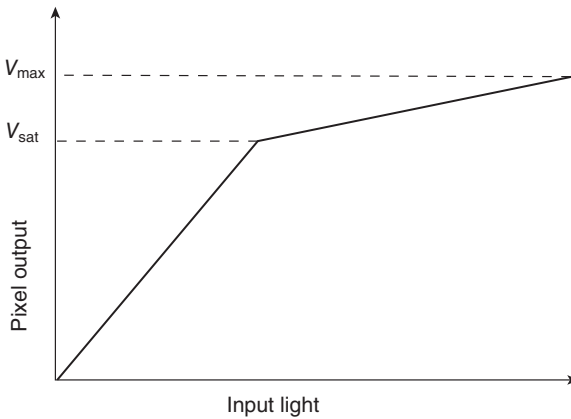
5.22 Pixel response when independently integrating the spilled over charge [79].

an ADC with a large number of bits may be required. One way to limit the need of a high performance ADC is to map the input intensity on to the frequency domain where a large dynamic range can be easily recorded and stored. A number of pixels with such potential have been suggested [81, 52]. However, these approaches often require large circuitry which is incompatible with the requirement of small pixels.

5.6.8 Multiple sampling image sensors

Another approach which has been extensively used to increase the dynamic range of CCDs as well as linear APSs is to capture the same scene two or more times with different integration times. One can capture the high intensity regions of the image with a short integration time. Similarly, a long integration time can be used to capture low intensity regions as shown earlier in Fig. 5.12. An algorithmic fusion of these two images can then be used to produce a wide dynamic range image [82, 83].

An extension of this approach leading to improved image quality is to capture the same scene with a several different integration times, for example at exponentially increasing times [84]. In-pixel memories can be used to enhance the performance of multiple sampling techniques. Predictive algorithms have been used to automatically find an optimal integration time (the longest integration time for which the pixel does not saturate) [85]. This can be achieved by dividing the integration time into slots of different durations and using an iterative procedure to select the longest integration time before the pixel saturates. A pixel similar to an APS with the addition of a reset-select switch can achieve this requirement. Additionally, in-pixel

memory can also be used to store the reset information [40]. Dual and multi-sampling techniques, however, are unable to capture the full dynamic range in a single frame and often require a large number of frames along with computationally complex post-processing, thereby making them difficult to use in video and other fast capture applications.

5.7   Other high performance designs

CMOS image sensors as low cost low-power highly integrable detectors have found applications in a number of diverse fields. This has also prompted development of application specific image sensors. These include very large format image sensors for radiography, very low light detectors for astrophysics, computational sensors to perform focal plane analysis as well as sensors with enhanced spectral response. A number of these are covered in detail in other chapters of this book but will be briefly discussed here.

With reduction in thickness of the polysilicon layer in typical CMOS processes, CMOS image sensors are now directly sensitive to X-rays over a good energy band without the need for a scintillator. This provides an opportunity to develop low cost radiography detectors or even detectors for image guided radiotherapy. However, it is difficult to design lenses for X-rays and therefore one needs an imaging array of the same size as that of the image to be acquired [86]. This makes the design of such detectors fairly challenging. Nevertheless, such detectors do not need very small pixel pitch and often pixels with $100\,\mu m$ or more of pitch are acceptable. However, the readout and interconnect between the pixels require further effort as such image sensors are often built over several chips, as single chips are incapable of providing the necessary imaging area. These are then stitched together to build a large format sensor. Fixed pattern noise as well as temporal noise correction need further attention with this multi-chip stitching. Finally, radiotherapy detectors are required in larger sizes compared to the wafer size itself. This has led to the design of a single sensor per wafer, which can then be stitched to form even larger format image sensors [87].

Low light level image sensors are often desired in commercial imaging; however, they are a fundamental requirement in scientific instrumentation, particularly telescopes and particle detectors. CMOS processes inherently have high leakage which leads to high dark current in the photodiode [13, 88]. This significantly affects the low light performance of the pixels. Figure 5.23 shows various sources of leakage in a typical photodiode [13]. These include physical phenomena of injection-diffusion and thermal generation-recombination followed by drift. In addition, there are parasitic leakage currents due to defects near the isolation regions as well as surface damage. A range of manufacturing techniques have been proposed to reduce the
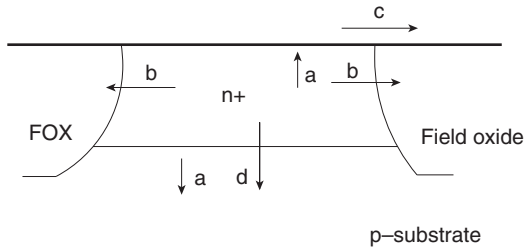
5.23 Sources of dark current in a standard n+ – P- sub diode in CMOS technology (a = thermal drift current, b = lateral leakage current, c = surface leakage current and d = diffusion current).
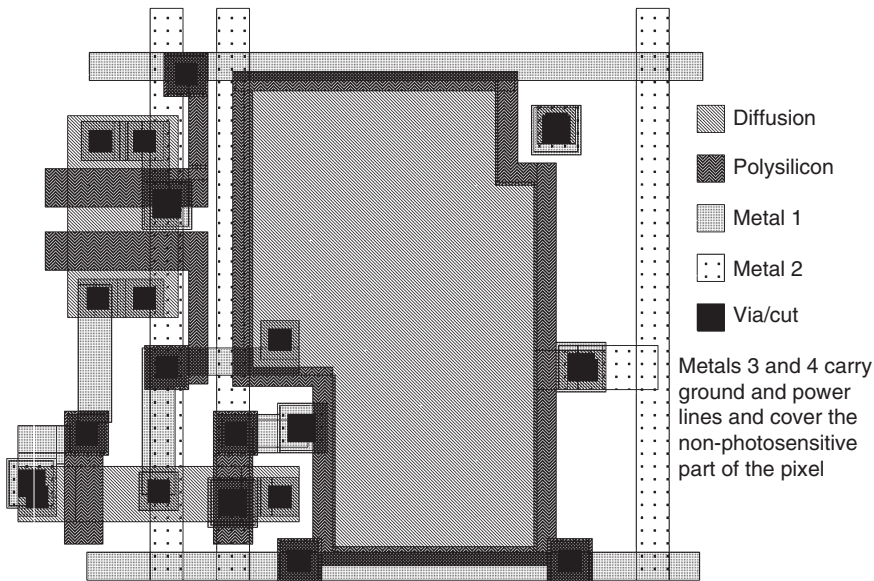
5.24 Layout of a logarithmic pixel with a guard ring around the photodiode to reduce the edge leakage currents [13].

inherent leakage current in the diode. These include enhanced surface cleaning [89] and separation of the diode from stressed areas by using pinned or buried devices [88, 90]. Furthermore, special barrier layers have also been suggested to reduce the effect on diode leakage [91, 92]. A low cost way to imitate the effect of isolation is by surrounding the photodiode with a guard ring to block the doping of the active area diode during fabrication [13, 93, 94] with a guard ring of polysilicon as shown in Fig. 5.24.

However, these approaches can only reduce the leakage current, but cannot enable measurement of single or few photons. For such a device,

image sensors can be built with single photon avalanche detectors (SPADs) [95], which are covered elsewhere in this book. Finally, the noise floor is determined by thermally generated electrons. This can be reduced by lowering the temperature of the image sensor. It is standard practice to cool detectors in telescopes to very low temperatures and this enables a sub-electron noise floor over hours of measurement.

5.8   Conclusion

In this chapter, we have studied approaches to enhance the performance of CMOS image sensors including higher speed, lower noise, lower power, wider dynamic range and lower dark performance. A number of these have required technology improvement as well as co-design of circuits and technologies. Higher speed has been achieved primarily by advancing the signal conversion stage. Circuits of correlated double sampling as well as the ability to manufacture special diodes has enabled reduction of temporal as well as fixed pattern noise. A number of circuit approaches have been presented to improve the dynamic range of the image sensor. Low light performance has primarily been improved by process modifications; however, special circuits such as single photon avalanche detectors are being used for very low light measurement and imaging. Further advancement of the optical properties of the detectors in the standard CMOS process will enhance the abilities of the image sensors. These may include integration of non-standard detectors to enhance spectral sensitivity of image sensors as well as nanophotonic structures to image into the terahertz regime as well.

5.9   References

1. K. Kitamura, T. Watabe, Y. Sadanaga, T. Sawamoto, T. Kosugi, T. Akahori, T. Iida, K. Isobe, T. Watanabe, H. Shimamoto, H. Ohtake, S. Aoyama, S. Kawahito and N. Egami, 'A 33 Mpixel, 120 fps CMOS image sensor for UDTV application with two-stage column-parallel cyclic ADCs,' in *Proc. of Int. Image Sensor Workshop*. Hokkaido, Japan: IISS, pp. 3–6, 2011.
2. S. Huang, T. Yamashita, Y. Wang and K. Ong, 'A 2.5 inch, 33 Mpixel, 60 fps CMOS image sensor for UHDTV application,' in *Proc. of 2009 Int. Image Sensor Workshop*. Bergen, Norway: IISS, pp. 308–311, 2009.
3. G. Holst, CMOS/CCD Sensors and Camera Systems. Bellingham, Washington USA: SPIE Press, 2007.
4. C. Chao, H. Tu, K. Chou, P. Chou and F. Hsueh, 'Crosstalk metrics and the characterization of 1.1 μm-pixel CIS,' in *Proc. of 2011 Int. Image Sensor Workshop*, no. 5. Hokkaido, Japan: IISS, pp. 3–6, 2011.
5. T. Chen, P. Catrysse, A. E. Gamal and B. Wandell, 'How small should pixel size be?' in *Proc. of SPIE*, vol. 3965. SPIE, pp. 451–459, 2000.

6. Y. C. Kim, Y. T. Kim, S. H. Choi, H. K. Kong, S. I. Hwang, J. H. Ko, B. S. Kim, T. Asaba, S. H. L. Hee, J. S. Hahn, J. H. Im, T. S. Oh, D. M. Yi, J. M. Lee, W. P. Yang, J. C. Ahn, E. S. Jung and Y. H. Lee, '½-inch 7.2 MPixel CMOS image sensor with 2.25 μm pixels using 4-shared pixel structure for pixel-level summation,' in *2006 IEEE Int. Solid State Circuits Conference – Dig. of Tech. Papers. IEEE*. pp. 1994–2003, 2006.

7. H. Takahashi, M. Kinoshita, K. Morita, T. Shirai, T. Sato, T. Kimura, H. Yuzurihara, S. Inoue and S. Matsumoto, 'A 3.9-μm pixel pitch VGA format 10-b digital output CMOS image sensor with 1.5 transistor/pixel,' *IEEE Journal of Solid-State Circuits*, vol. 39, no. 12, pp. 2417–2425, Dec. 2004.

8. R. F. Lyon and P. M. Hubel, 'Eyeing the camera : into the next century,' in *The Tenth Color Imaging Conference: Color Sci. and Eng. Syst., Technologies, Applicat*. Scottsdale, Arizona, USA: IST – The Society for Imaging Science and Technology, pp. 349–355, Nov. 2002.

9. H. Tian, B. Fowler and A. Gamal, 'Analysis of temporal noise in CMOS photodiode active pixel sensor,' *IEEE Journal of Solid-State Circuits*, vol. 361, no. 1, pp. 92 –101, Jan. 2001.

10. A. E. Gamal, B. Fowler, H. Min and X. Liu, 'Modeling and estimation of fpn components in CMOS image sensors,' in *Proc. SPIE*, vol. 3301, pp. 168–177, 1998.

11. D. Joseph and S. Collins, 'Modelling, calibration and correction of illumination-dependent fixed pattern noise in logarithmic CMOS image sensor,' *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 5, pp. 996 –1001, October 2001.

12. R. G. Driggers, R. Vollmerhausen and K. Krapels, 'Target identification performance as a function of temporal and fixed pattern noise,' *Optical Engineering*, vol. 40, no. 3, pp. 443–447, March 2001.

13. B. Choubey and S. Collins, 'Wide dynamic range CMOS pixels with reduced dark current,' *Analog Integrated Circuits and Signal Processing*, vol. 56, pp. 53–60, 2008.

14. D. Joseph and S. Collins, 'Temperature dependence of fixed pattern noise in logarithmic CMOS image sensors,' *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 8, pp. 2503 –2511, August 2009.

15. B. Choubey, 'Fixed pattern noise' in R. G. Driggers (ed.) *Encyclopedia of Optical Engineering*. Taylor and Francis, 2012.

16. M. J. M. Pelgrom, A. C. J. Duinmaijer and A. P. G. Welbers, 'Matching properties of MOS transistors,' *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, October 1989.

17. D. L. Lin, C. C. Wang and C. L. Wei, 'Quantified temperature effect in a CMOS image sensor,' *IEEE Transactions on Electron Devices*, vol. 57, no. 2, pp. 422–428, February 2010.

18. M. Loose, K. Meier and J. Schemmel, 'A selfcalibrating single-chip CMOS camera with logarithmic response,' *IEEE Journal of Solid-State Circuits*, vol. 36, no. 4, pp. 586–96, April 2001.

19. N. Ricquer and B. Dierickx, 'Active pixel CMOS sensor with on chip nonuniformity correction,' in *Proceedings of the IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, April 1995.

20. S. K. Mendis, S. E. Kemeny, R. C. Gee, B. Pain, C. Staller, Q. Kim and E. R. Fossum, 'CMOS active pixel image sensors for highly integrated imaging

systems,' *IEEE Journal of Solid-State Circuits*, vol. 32, no. 2, pp. 187–97, February 1997.

21. R. C. Hardie, M. M. Hayat, E. Armstrong and B. Yasuda, 'Scene-based nonuniformity correction with video sequences and registration,' *Applied Optics*, vol. 39, no. 8, pp. 1241–1250, March 2000.

22. S. Lim and A. E. Gamal, 'Gain fixed pattern noise correction via optical flow,' *IEEE Transactions on Circuits and Systems I*, vol. 51, no. 4, pp. 779–786, April 2004.

23. Z. Zhou, B. Pain and E. Fossum, 'CMOS active pixel sensor with on-chip successive approximation analog-to-digital converter,' *IEEE Trans. on Electron Devices*, vol. 44, no. 10, pp. 1759–1763, 1997.

24. A. Krymski, N. Bock, D. Van Blerkom and E. Fossum, 'A high-speed, 240-frames/s, 4.1-mpixel CMOS sensor,' *IEEE Trans. on Electron Devices*, vol. 50, no. 1, pp. 130–135, Jan. 2003.

25. I. Takayanagi, M. Shirakawa, K. Mitani, M. Sugawara, S. Iversen, J. Moholt, J. Nakamura and E. Fossum, 'A 1.25-inch 60-frames/s 8.3-M-pixel digital-output CMOS image sensor,' *IEEE J. of Solid-State Circuits*, vol. 40, no. 11, pp. 2305–2314, Nov. 2005.

26. S. Matsuo, T. J. Bales, M. Shoda, S. Osawa, K. Kawamura, A. Andersson, M. Haque, H. Honda, B. Almond, Y. Mo, J. Gleason, T. Chow and I. Takayanagi, '8.9-Megapixel video image sensor with 14-b column-parallel SA-ADC,' *IEEE Trans. on Electron Devices*, vol. 56, no. 11, pp. 2380–2389, Nov. 2009.

27. X. Wang, J. Bogaerts, G. Vanhorebeek, K. Ruythoren, B. Ceulemans, G. Lepage, P. Willems and G. Meynants, 'A 2.2M CMOS image sensor for high-speed machine vision applications,' in *Proc. Sensors, Cameras, and Syst. for Ind./ Scientific Applicat. XI*, vol. 7536, pp. 75 360M–75 360M–7, 2010.

28. I. Takayanagi and J. Nakamura, 'High-resolution CMOS video image sensors,' *Proc. of the IEEE*, vol. 101, no. 1, pp. 1–13, 2012.

29. M. F. Snoeij, A. J. P. Theuwissen, K. A. A. Makinwa and J. H. Huijsing, 'Multiple-ramp column parallel ADC architectures for CMOS image sensors,' *IEEE Journal of Solid-State Circuits*, vol. 42, no. 12, pp. 2968–2977, December 2007.

30. J. Reyneri, H. Tian, D. Wile and D. Yang, 'A 0.18 m high dynamic range NTSC/ PAL imaging system-on-chip with embedded DRAM frame buffer,' in *IEEE Int. Solid-State Circuits Conference, 2003. Dig. of Tech. Papers*. ISSCC., vol. 1. IEEE, pp. 212–488, 2003.

31. A. Kitchen, A. Bouzerdoum and A. Bermak, 'Time domain analogue to digital conversion in a digital pixel sensor array,' in *Second IEEE Int. Workshop on Electron. Design, Test and Applicat*. IEEE, pp. 108–112, 2004.

32. Z. Ignjatovic, D. Maricic and M. F. Bocko, 'Low power, high dynamic range CMOS image sensor employing pixel-level oversampling sigma-delta analog-to-digital conversion,' *IEEE Sensors J.*, vol. 12, no. 4, pp. 737–746, Apr. 2012.

33. S. Kleinfelder and A. El Gamal, 'A 10000 frames/s CMOS digital pixel sensor,' *IEEE J. of Solid-State Circuits*, vol. 36, no. 12, pp. 2049–2059, 2001.

34. P. Willems, G. Meynants, G. Vanhorebeek and C. Ma, 'High-speed VGA resolution CMOS image sensor with global shutter,' in R. Widenhorn, V. Nguyen, and A. Dupret (eds), *Proc. of SPIE-IS&T Electron. Imaging*, vol. 8298, no. V. SPIE, pp. 829 802–829 802–7, 2012.

35. T. Toyama, K. Mishina, H. Tsuchiya, T. Ichikawa, H. Iwaki, Y. Gendai, H. Murakami, K. Takamiya, H. Shiroshita, Y. Muramatsu and T. Furusawa, 'A

17.7 Mpixel 120 fps CMOS image sensor with 34.8 Gb/s readout,' in *IEEE Int. Solid-State Circuits Conference*. IEEE, pp. 420–422, Feb. 2011.

36. Y. Tochigi, K. Hanzawa, Y. Kato, R. Kuroda, H. Mutoh, R. Hirose, H. Tominaga, K. Takubo, Y. Kondo and S. Sugawa, 'A global-shutter CMOS image sensor with readout speed of 1 Tpixel/s burst and 780 Mpixel/s continuous,' in *2012 IEEE Int. Solid-State Circuits Conference*, vol. 43, no. 4. IEEE, pp. 382–384, Feb. 2012.

37. H. Takeshita, T. Sawada, T. Iida, K. Yasutomi and S. Kawahito, 'High-speed charge transfer pinnedphotodiode for a CMOS time-of-flight range image sensor,' in *Proc. SPIE 7536, Sensors, Cameras, and Syst. for Ind./Scientific Applicat. XI*, vol. 7536. SPIE, pp. 75 360R–75 360R–9, 2010.

38. A. Krymski, 'A high speed 4 Megapixel digital CMOS sensor,' in *Proc. of 2007 Int. Image Sensor Workshop*. Ogunquit, Maine, USA: IISS, pp. 78–81, 2007.

39. A. Spickermann, D. Durini, A. Suss, W. Ulfig, W. Brockherde, B. J. Hosticka, S. Schwope and A. Grabmaier, 'CMOS 3D image sensor based on pulse modulated time-of-flight principle and intrinsic lateral drift-field photodiode pixels,' in *2011 Proc. of the ESSCIRC* (ESSCIRC). Helsinki: IEEE, pp. 111–114, Sep. 2011.

40. A. Belenky, A. Fish, A. Spivak and O. Yadid-Pecht, 'Global shutter CMOS image sensor with wide dynamic range,' *IEEE Transactions on Circuits and Systems-II: Express briefs*, vol. 54, no. 12, pp. 1032–1036, December 2007.

41. E. C. Fox and O. Nixon, 'Pinned photodiode five transistor pixel,' US Patent no. 6566697, January 2000.

42. Y. Tochigi, K. Hanzawa, Y. Kato, R. Kuroda, H. Mutoh, R. Hirose, H. Tominaga, K. Takubo, Y. Kondo and S. Sugawa, 'A global-shutter CMOS image sensor with readout speed of 1-tpixel/s burst and 780-mpixel/s continuous,' *IEEE Journal of Solid-State Circuits*, vol. 48, no. 1, pp. 329–338, 2013.

43. K. Yasutomi, S. Itoh and S. Kawahito, 'A two stage charge transfer active pixel CMOS image sensor with low-noise global shuttering and a dual shuttering mode,' *IEEE Transactions on Electron Devices*, vol. 58, no. 3, pp. 740–747, 2011.

44. K.-B. Cho, A. Krymski and E. Fossum, 'A 1.5- v 550- mu; w 176 times; 144 autonomous CMOS active pixel image sensor,' *IEEE Transactions on Electron Devices*, vol. 50, no. 1, pp. 96–105, 2003.

45. I. Brouk, K. Alameh and Y. Nemirovsky, 'Design and characterization of CMOS/ SOI image sensors,' *IEEE Transactions on Electron Devices*, vol. 54, no. 3, pp. 468–475, 2007.

46. V. Suntharalingam, D. D. Rathman, G. Prigozhin, S. Kissel and M. Bautz, 'Back-illuminated three dimensionally integrated CMOS image sensors for scientific applications,' in *Proc. SPIE 6690, Focal Plane Arrays for Space Telescopes III*, p. 669009, 2007.

47. A. Fish and O. Yadid-Pecht, 'Low-power smart CMOS image sensors,' in *IEEE International Symposium on Circuits and Systems*, pp. 1408–1411, 2008.

48. W. Gao, E. Shen and R. I. Hornsey, 'Low-power realization in main blocks of CMOS APS image sensor,' in *Proceedings of the SPIE*, vol. 5969, pp. 507–515, 2005.

49. A. Spivak, A. Teman, A. Belenky, O. Yadid-Pecht and A. Fish, 'Low-voltage 96 db snapshot CMOS image sensor with 4.5 nw power dissipation per pixel,' *Sensors*, vol. 12, no. 8, pp. 10067–10085, 2012.

50. L. McIlrath, 'A low-power low-noise ultra wide dynamic-range CMOS imager with pixel-parallel a/d conversion,' *IEEE Journal of Solid-State Circuits*, vol. 36, no. 5, pp. 846–853, 2001.

51. M. Perenzoni and L. Gonzo, 'Solar-powered CMOS image sensor,' *Electronics Letters*, vol. 46, no. 1, pp. 77–79, 2010.

52. W. W. X. Wang and R. Hornsey, 'A high dynamic range CMOS image sensor with in pixel light to-frequency conversion,' *IEEE Transactions on Electron Devices*, vol. 53, no. 12, pp. 2988–2992, December 2006.

53. S. Hanson, Z. Foo, D. Blaauw and D. Sylvester, 'A 0.5 v sub-microwatt CMOS image sensor with pulse width modulation read-out,' *IEEE Journal of Solid-State Circuits*, vol. 45, no. 4, pp. 759–767, 2010.

54. O. Yadid-Pecht, 'Wide dynamic range sensors,' *Optical Engineering*, vol. 38, no. 10, pp. 1650–1660, October 1999.

55. B. Choubey, 'CMOS Pixels for Wide dynamic range imaging,' *Advances in Electrical Engineering Research*. Nova Science Publisher. pp. 329–375, 2011.

56. S. Kavadias, B.Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts and J. Bogaerts, 'A logarithmic response CMOS image sensor with on-chip calibration,' *IEEE Journal of Solid-State Circuits*, vol. 35, no. 8, pp. 1146–52, August 2000.

57. E. Labonne, G. Sicard, M. Renaudin and P.-D. Berger, 'A 100 db dynamic range CMOS image sensor with global shutter,' in *Electronics, Circuits and Systems*, 2006. ICECS '06. 13th IEEE International Conference on Electronics, Circuits and Systems, ICES, pp. 1133–1136, 2006.

58. Y. Ni and K. Matao, 'A CMOS log sensor with on chip FPN compensation,' in *Proceedings of the European Solid-State Circuits Conference (ESSCIRC)*, pp. 128–132, 2001.

59. L. W. Lai, C. H. Lai and Y. C. King, 'A novel logarithmic response CMOS image sensor with high output voltage swing and in-pixel fixed-pattern noise reduction,' *IEEE Sensors Journal*, vol. 4, no. 1, pp. 122–126, February 2004.

60. B. Choubey, S. Otim, S. Aoyama, D. Joseph and S. Collins, 'An electronic calibration scheme for logarithmic CMOS image sensor,' *IEEE Journal of Sensors*, vol. 6, no. 4, pp. 950–956, August 2006.

61. G. Strom, R. Henderson, J. Hurwitz, D. Renshaw, K. Findlater and M. Purcell, 'Extended dynamic range from a combined linear-logarithmic CMOS image sensor,' *IEEE Journal of Solid-State Circuits*, vol. 41, pp. 2095–2106, 2006.

62. E. Fox, J. Hynecek and D. Dykaar, 'Wide dynamic range pixel with combined linear and logarithmic response and increased signal swing,' in *IS&T/SPIE 12th International Symposium on Electronic Imaging – Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications, Proceedings of SPIE*, vol. 3965. SPIE, pp. 4–10, May 2000.

63. B. Choubey and S. Collins, 'Models for pixels with wide dynamic range combined linear and logarithmic response,' *IEEE Sensors Journal*, vol. 7, no. 7, pp. 1066–1072, July 2007.

64. D. Stoppa, A. Simoni, L. Gonzo, M. Gottardi and G. D. Betta, 'Novel CMOS image sensor with a 132-db dynamic range,' *IEEE Journal of Solid-State Circuits*, vol. 37, no. 12, pp. 1846–1852, December 2002.

65. M. Ikebe and K. Saito, 'A wide-dynamic-range compression image sensor with negative-feedback resetting,' *IEEE Sensors Journal*, vol. 7, no. 5, pp. 897–904, May 2007.

66. C. H. Lai, Y. C. King and S. Y. Huang, 'A 1.2-v 0.25 m clock output pixel architecture with wide dynamic range and self-offset cancellation,' *IEEE Sensors Journal*, vol. 6, no. 2, pp. 398–405, April 2006.

67. F. Andoh, H. Shimamoto and Y. Fujita, 'A digital pixel image sensor for real time readout,' *IEEE Transactions on Electron Devices*, vol. 47, no. 11, pp. 2123–2127, November 2000.
68. Q. Luo, J. G. Harris and Z. J. Chen, 'A time-to first spike CMOS image sensor with coarse temporal sampling,' *Analog Integrated Circcuits and Signal Processing*, vol. 47, pp. 303–313, 2006.
69. E. Culurciello, R. Etienne-Cummings, and K. A. Boahen, 'A biomorphic digital image sensor,' *IEEE Journal of Solid-State Circuits*, vol. 38, no. 2, pp. 281–294, February 2003.
70. O. Yadid-Pecht, R.Ginosar and Y. Shacham-Diamand, 'A random access photodiode array for intelligent image capture,' *IEEE Transactions on Electron Devices*, vol. 38, no. 8, pp. 1772–1781, August 1991.
71. T. Kinugusa, M. Noda, T. Imaide, I. Aizawa, Y. Todaka and M. Ozama, 'An electronic variable shutter system in video camera use,' *IEEE Transactions on Consumer Electronics*, vol. CE-33, pp. 249–258, August 1987.
72. M. Schanz, C. Nitta, A. Bubmann, B. J. Hostika and R. K. Wertheimer, 'A high dynamic range CMOS image sensor for automotive applications,' *IEEE Journal of Solid-State Circuits*, vol. 35, no. 7, pp. 932–938, July 2000.
73. T. Yasuda, T. Hamamoto and K. Aizawa, 'Adaptive-integration-time image sensor with realtime reconstruction function,' *IEEE Transactions on Electron Devices*, vol. 50, no. 1, pp. 111–120, January 2003.
74. H. Y. Cheng, B. Choubey and S. Collins, 'A high dynamic range integrating pixel with an adaptive logarithmic responses,' *IEEE Photonics Technology Letters*, vol. 19, no. 15, pp. 1169–1171, August 2007.
75. B. Choubey, H. Cheng and S. Collins, 'A wide dynamic range CMOS image sensor with adjustable logarithmic response,' in *Proceedings of the SPIE Electronic Imaging*, vol. 6816 (2008), January 2008.
76. B. Choubey, 'Pixel circuit,' GB Patent 20 110 014 898 20 110 826, 2011.
77. Anon., 'A wide dynamic range CMOS pixel with Steven's power law response,' in *Proceedings of the SPIE Optics and Photonics – Photonic Devices and Applications*, vol. 7780, San Diego, August, 2010.
78. S. Decker, D. McGrath, K. Brehmer and C. Sodini, 'A 256 times; 256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output,' *IEEE Journal of Solid-State Circuits*, vol. 33, no. 12, pp. 2081–2091, December 1998.
79. R. Guidash, 'Variable collection of blooming charge to extend dynamic range,' US Patent no. 6307195, October 2001.
80. N. Akahane, S. Sugawa, S. Adachi, K. Mori, T. Ishiuchi and K. Mizobuchi, 'A sensitivity and linearity improvement of a 100-db dynamic range CMOS image sensor using a lateral overflow integration capacitor,' *IEEE Journal of Solid-State Circuits*, vol. 41, no. 4, pp. 851–8, 2006.
81. P. Seitz and J. M. Raynor, 'A linear array of photodetectors with wide dynamic range and near quantum-noise limit,' *Sensors and Actuators A*, vol. 61, no. 1–3, pp. 327–330, June 1997.
82. O. Yadid-Pecht and E. R. Fossum, 'Image sensor with ultra-high-linear-dynamic range utilizing dual output CMOS active pixel sensors,' *IEEE Transactions on Electron Devices*, vol. 44, no. 10, pp. 1721–1723, October 1997.
83. J. O. Schrey, G. Filimonovic, A. Bussmann, W. Brockherde and B.J. Hosticka, 'A 1 k × 1 k high dynamic range CMOS image sensor with on-chip programmable

region of interest readout,' in *Proceedings of the European Solid-State Circuits Conference*, ESSCIRC, 2001.

84. D. X. D. Yang, A. E. Gammal, B. Fowler and H. Tian, 'A 640 × 512 CMOS image sensor with ultrawide dynamic range floating point pixel level ADC,' *IEEE Journal of Solid-State Circuits*, vol. 34, no. 12, pp. 1821–1999, December 2000.

85. P. M. Acosta-Serafini, I. Masaki, and C. G. Sodini, 'A ⅓″VGA linear wide dynamic range CMOS image sensor implementing a predictive multiple sampling algorithm with overlapping integration intervals,' *IEEE Journal of Solid-State Circuits*, vol. 39, no. 9, pp. 1487–1496, September 2004.

86. R. Turchetta, N. Guerrini and I. Sedgwick, 'Large area CMOS image sensors,' *Journal of Instrumentation*, vol. 6, no. 01, p. C01099, 2011.

87. Y. Yamashita, H. Takahashi, S. Kikuchi, K. Ota, M. Fujita, S. Hirayama, T. Kanou, S. Hashimoto, G. Momma and S. Inoue, 'A 300mm wafer-size CMOS image sensor with in-pixel voltage-gain amplifier and column-level differential readout circuitry,' in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 2011 IEEE International, pp. 408–410, 2011.

88. H. I. Kwon, O. J. Kwon, H. Shin, B. G. Park, S. S.Park and J. D. Lee, 'The effects of deuterium annealing on the reduction of dark currents in the CMOS APS,' *IEEE Transactions on Electron Devices*, vol. 51, no. 8, pp. 1346–1349, August 2004.

89. H. Kwon, I. M. Kang, B. G. Park, J. D. Lee and S. S. Park, 'The analysis of dark signals in the CMOS APS imagers from the characterization of test structures,' *IEEE Transactions on Electron Devices*, vol. 51, no. 2, pp. 178–184, February 2004.

90. I. Inoue, N. Tanaka, H. Yamashita, T. Yagamuchi, H. Ishiwata and H. Ihara, 'Low-leakage-current and low-operating-voltage buried photodiode for a CMOS imager,' *IEEE Transactions on Electron Devices*, vol. 50, no. 1, pp. 43–47, January 2003.

91. A. Okita and S. Suzuki, 'Photoelectric transducer and imaging device,' Japanese Patent publication no. 2002-353430, December 2002.

92. R. Mann, 'Improved semiconductor device for isolating a photodiode to reduce junction leakage and method of formation,' International Patent no. WO 03/019675 A1, March 2003.

93. T. E. Kopley, D. W. Vook and T. Dungan, 'Method of manufacturing a structure for reducing leakage currents by providing isolation between adjacent regions of an integrated circuit,' US Patent no. 6417074, July 2002.

94. H. Y. Cheng and Y. C. King, 'An ultra low dark current CMOS image sensor cell using n+ ring reset,' *IEEE Electron Devices Letters*, vol. 23, no. 9, pp. 538–540, September 2002.

95. C. Niclass, A. Rochas, P.-A. Besse and E. Charbon 'Design and characterization of a CMOS 3-D image sensor based on single photon avalanche diodes,' *IEEE Journal of Solid-State Circuits*, vol. 40, no. 9, pp. 1847–1854, 2005.

6

Smart cameras on a chip: using complementary metal-oxide-semiconductor (CMOS) image sensors to create smart vision chips

D. GINHAC, Université de Bourgogne, France

Abstract: In this chapter, we introduce the fundamental concept of smart cameras on a chip or smart vision chips that simultaneously integrate the same die image capture capability and highly complex image processing. Successive technology scaling has made possible the integration of specific processing elements designed at chip level, at column level or at pixel level. To illustrate this continuous evolution, we survey three different categories of vision chips, exploring first the pioneering works on artificial retinas, then describing the most significant computational chips, and finally presenting the most recent image processing chips able to perform complex algorithms at a high frame rate.

Key words: smart camera on a chip, vision chips, focal-plane image processing.

6.1 Introduction

Today, digital smart cameras are rapidly becoming ubiquitous, due to reduced costs and the increasing demands of multimedia applications. Improvements in the growing digital imaging world continue to be made with two main image sensor technologies: charge coupled devices (CCDs) and complementary metal-oxide-semiconductor (CMOS) sensors. Historically, CCDs have been the dominant image-sensor technology. However, the continuous advances in CMOS technology for processors and memories have made CMOS sensor arrays a viable alternative to the popular CCD sensors. This led to the adoption of CMOS image sensors in several high-volume products, such as webcams, mobile phones or tablets. New technologies provide the potential for integrating a significant amount of very-large scale integration (VLSI) electronics into a single chip, greatly reducing the cost, power consumption, and size of the camera (Fossum, 1993; Seitz, 2000; Litwiller, 2001). By exploiting these advantages, innovative CMOS sensors have been developed (see Fossum, 1997 and Bigas *et al.*, 2006 for two detailed surveys on CMOS image sensors). Numerous works

165

have focused on major parameters such as sensitivity (Krymski and Tu, 2003; Murari *et al.*, 2009), noise (Sumi, 2006), power consumption (Hanson *et al.*, 2010), voltage operation (Xu *et al.*, 2002; Gao and Yadid-Pecht, 2012), high-speed imaging (Dubois *et al.*, 2008; El-Desouki *et al.*, 2009) or dynamic range (Schrey *et al.*, 2002; Fontaine, 2011).

Moreover, the main advantage of CMOS image sensors is the flexibility to integrate signal processing at focal plane down to the pixel level. As CMOS image sensors technologies scale to 0.13 µm processes and under, processing units can be realized at chip level (system-on-chip approach), at column level by dedicating processing elements to one or more columns, or at pixel level by integrating a specific processing unit in each pixel (El Gamal *et al.*, 1999; El Gamal and Eltoukhy, 2005). By exploiting the ability to integrate sensing with analog or digital processing, new types of CMOS imaging systems can be designed for machine vision, surveillance, medical imaging, motion capture and pattern recognition among other applications. This extends the basic concept of electronic camera on a chip proposed by Fossum (1997) to the more sophisticated concept of smart camera on a chip (also called vision chip) including both analog signal processing functions, analog-to-digital conversion, digital signal and image processing as described in Fig. 6.1. In this chapter, we first define the concept of smart vision chips or smart cameras on a chip that integrate both sensing and complex image processing on the same chip. Image processing tasks are wide, spanning from basic image quality enhancements to complex applications managed by analog or digital microprocessors integrated in the sensor. These tasks can be performed globally at chip level (system-on-chip approach), regionally at column level by dedicating processing elements to one or more columns (typically analog to digital converters), or locally at pixel level by integrating a specific unit in each pixel.

The remainder of this chapter is chronologically organized. It first describes the pioneering works on spatial and spatio-temporal image processing vision chips. These vision chips can be viewed as the first smart sensors and came to light during the 1990s under the well-known term of silicon artificial retinas. Secondly, it talks about computational chips that have turned the first generation of vision chips into fully programmable smart vision chips reusable in many fields of application. In this section, we successively address the cellular neural networks paradigm and the software-programmable single instruction multiple data (SIMD) processor arrays.

Thirdly, this chapter deals with high-speed image processing chips. Such chips integrate a processing element within each pixel based on SIMD architecture, enabling massively parallel computations and leading to high frame rates up to thousands of images per second. In this section, we survey the state-of-the-art vision chips, in both the analog and digital domains.

Image
capturing

Analog to digital
conversion

Analog/digital
processing

Clock

Counter

Digital
output

DA

Analog input

Smart camera on a chip
vision chip
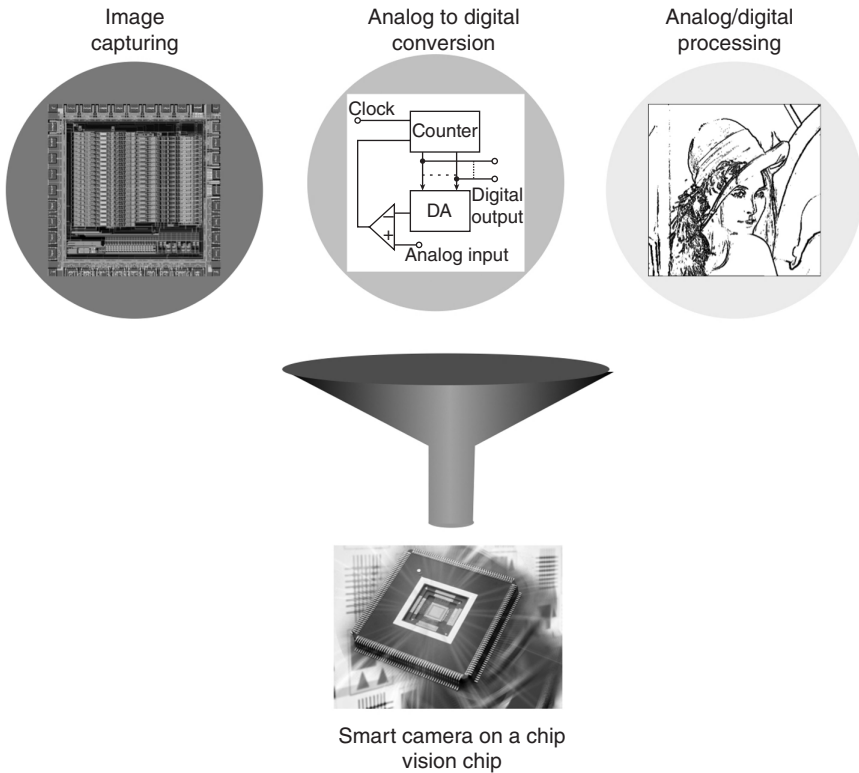
6.1 Conceptual view of a smart camera on a chip.

Finally, we set out recent trends on smart vision chips. From a technological point of view, three-dimensional integrated imagers, based on 3D stacking technology, become an emerging solution to design powerful imaging systems because the sensor, the analog-to-digital converters and the image processors can be designed and optimized in different technologies, improving the global system performance. Combined with a backside-illuminated (BSI) technology, 3D vision chips allow high-speed signal processing and have an optical fill factor of 100%. From a conceptual point of view, electronic imaging aims at detecting individual photons. Single-photon imaging can be seen as the next step to reach in the design of smart imaging systems. Such vision chips are able to detect single photons by combining high sensitivity with excellent photon timing properties in the range of a few tens of picoseconds. So, joint optimizations of the sensor, of the analog to digital converters and the processors offer opportunity to improve the sensor performance and allow the emergence of new applications such as real time 3D imaging with time-of-flight cameras

(delivering simultaneously intensity images (2D) and ranges of the observed scene), medical imaging, molecular biology, astronomy and aerospace applications.

6.2   The concept of a smart camera on a chip

CMOS image sensors have become increasingly mature and now dominate image sensor market shipments. Despite a large variety of applications, imaging systems always embed the same basic functions allowing the formation of a 2D image from a real illuminated scene. These basic functions consist of:

- optical collection of photons (e.g. a lens)
- conversion of photons to electrons (e.g. a photodiode)
- readout of the collected signal and
- logic control for driving the sensors.

Note that readout may include some basic analog processing in order to enhance the image quality by removing temporal noise and fixed pattern noise (FPN). However, embedding such processing functions into a single chip does not turn a standard camera into a smart camera. The fundamental differences between a smart camera and a standard camera is that a smart camera must include a special intelligent image processing unit to run specific algorithms, in which the primary objective is not to improve image quality but to extract information and knowledge from images (Shi and Lichman, 2006). The close colocation of sensing and processing in a smart camera transforms the traditional camera into a smart sensor (Rinner and Wolf, 2008).

In CCD technology, integrating electronics dedicated to specific image processing onto the silicon is by essence impractical (Litwiller, 2001) because analog-to-digital conversion and signal processing functions are performed outside CCD sensors. On the contrary, CMOS image sensors and a smart camera on a chip are intimately closed because CMOS technologies provide the ability to integrate complete imaging systems within the pixel sensor (Aw and Wooley, 1996; Loinaz et al., 1998; Smith et al., 1998). Basically, a smart camera on a chip or a vision chip includes image capturing, analog-to-digital conversion, and analog / digital image processing as seen in Fig. 6.1 on the same die.

The key advantages are:

- to release the host computer of complex pixel processing tasks by integrating the image sensor and the processors into a single chip
- to accelerate processing speed by using parallel processing elements and

- to minimize the data transfer between cameras and the outside world by only outputting extracted feature information (Zhang *et al.*, 2011a).

To summarize, the smart camera on a chip has the advantages of small size, high processing speed, low power consumption, and can be tailor-made for broad applications.

As an illustrative example, the VISoc single chip smart camera designed by Albani *et al.* (2002) integrates a 320×256-pixel CMOS sensor, a 32-bit RISC processor, a neural co-processor, a 10-bit analog-to-digital converter and I/O on to a $6 \times 6\,mm^2$ single chip in a $0.35\,\mu m$ standard CMOS process.

The greatest promise of CMOS smart cameras arises from the ability to flexibly integrate both sensing and complex image processing on the same chip (El Gamal and Eltoukhy, 2005). As CMOS image sensors technologies scale further down, smart vision chips are able to integrate focal-plane image processing tasks early in the signal chain. The range of pixel processing is wide, spanning from simple amplifiers dedicated to SNR enhancements to complete programmable digital or analog microprocessors in each pixel. Processing units can be realized at chip level (system-on-chip approach), at column level by dedicating processing elements to one or more columns (typically analog to digital converters), or at pixel level by integrating a specific unit in each pixel. Historically, most of the research has dealt with chip level and column level (Dickinson *et al.*, 1995; Kemeny *et al.*, 1997; Hong and Hornsey, 2002; Yadid-Pecht and Belenky, 2003; Acosta-Serafini *et al.*, 2004; Kozlowski *et al.*, 2005; Sakakibara *et al.*, 2005). Indeed, pixel-level processing has been generally dismissed for years because pixel sizes are often too large to be of practical use. However, as CMOS scales down, integrating a processing element at each pixel or group of neighboring pixels becomes more feasible since the area occupied by the pixel transistors decreases, leading to an acceptable small pixel size. A fundamental tradeoff must be made between three dependent and correlated variables: pixel size, processing element area, and fill-factor. This implies various points of view (Ginhac *et al.*, 2008):

- For a fixed fill-factor and a given processing element area, the pixel size is reduced with technology improvements. As a consequence, reducing pixel size increases spatial resolution for a fixed sensor die size.
- For a fixed pixel size and a given processing element area, the photodiode area and the fill-factor increase as technology scales since the area occupied by the pixel transistors in each processing element decreases. It results in better sensibility, higher dynamic range and signal-to-noise ratio.
- For a fixed pixel size and a given fill-factor, the processing element can integrate more functionalities since the transistors require less area as

technology scales. Consequently, the image processing capabilities of the sensor increase.

In summary, each new technology process offers

- integrating more processing functions in a given silicon area, or
- integrating the same functionalities in a smaller silicon area.

This can benefit the quality of imaging in terms of resolution or noise for example by integrating specific processing functions such as correlated double sampling (Nixon *et al.*, 1995), anti blooming (Wuu *et al.*, 2001), high dynamic range (Decker *et al.*, 1998), and even all basic camera functions (color processing functions, color correction, white balance adjustment, gamma correction) on to the same camera-on-chip (Yoon *et al.*, 2002). However, shrinking pixels size inevitably reach foreseen physical limits leading to poor performance of small pixels because of the reduced incident light on each pixel. Maintaining reasonable pixel performance (quantum efficiency, crosstalk, pixel capacity, angular signal response) and image quality (color reproduction, high dynamic range, limited chromatic aberration) when shrinking pixel size to small values such as 1.4 μm and smaller (typically used in mobile devices) is a big challenge for image sensor designers (Xiao *et al.*, 2009).

6.3   The development of vision chip technology

From an historical point of view, the pioneering works have concentrated efforts on spatial and spatio-temporal image processing vision chips in the field of machine vision applications. These vision chips can be viewed as the first smart sensors and came to light during the 1990s under the well-known term of silicon artificial retinas. Based on models of the vertebrate retina, they are able to implement some of its characteristics such as adaptation to local and global light intensity, and edge enhancement.

Generally, silicon artificial retinas are arrays of identical pixels including significantly more transistors per pixel than the three or four found in typical active pixel sensor (APS)-based sensors. This additional electronic circuitry performs pixel parallel processing over images immediately after they are captured without time-consuming and power-consuming image transfer. Spatial vision chips mainly implement basic neighborhood functions such as edge detection, smoothing, stereo processing, and contrast enhancement. On the other hand, spatio-temporal image processing vision chips are mainly devoted to motion detection functions requiring the implementation of simultaneous time-space processing. Moini (2000) proposes an exhaustive overview of significant development up to 1997
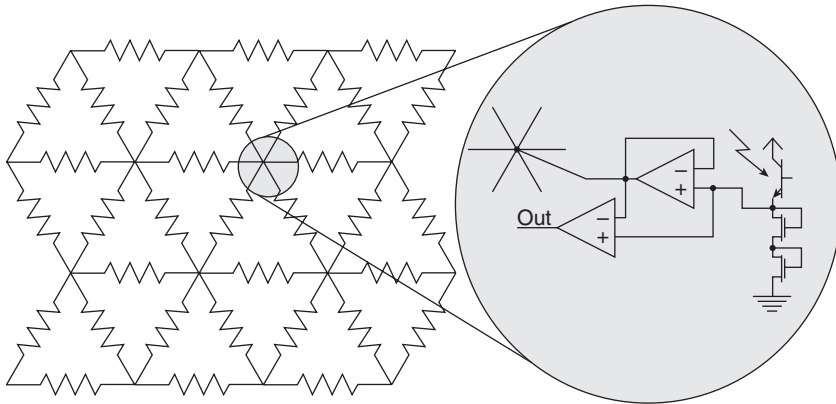
6.2 Architecture of Mead's artificial retina (from Moini, 2000).

about these two kinds of smart sensors while more recent developments are covered by Ohta (2008).

Artificial retinas were pioneered by Carver Mead in the late eighties, when he developed the first silicon retina that implements the first stages of retinal processing on a single silicon chip (Mead and Mahowald, 1988). This retina is based on models of computation of the vertebrate retina including specific structures such as cones, horizontal cells, and bipolar cells. First of all, cones, i.e. the light detectors, have been implemented using phototransistors and MOS-diode logarithmic current to voltage converters. Secondly, the outputs of the cones are then averaged, both spatially and temporally, by the horizontal cells. This averaging step is performed electronically using a hexagonal network of active resistors as seen on Fig. 6.2. Finally, bipolar cells detect the difference between the averaged output of the horizontal cells and the input.

These first works lead to several other research projects dedicated to the design of an analog artificial retina based on CMOS photodetectors combined with CMOS signal processing circuitry. Delbruck (1993) describes a two-dimensional silicon retina that computes a complete set of local direction-selective outputs. The chip motion computation uses unidirectional delay lines as tuned filters for moving edges. As a result, the detectors are sensitive to motion over a wide range of spatial frequencies. Brajovic and Kanade (1996) describe a VLSI computational sensor using both local and global interpixel processing that can perform histogram equalization, scene change detection, and image segmentation in addition to normal image capture. Deutschmann and Koch (1998) present the first working analog VLSI implementation of a one-dimensional velocity sensor that uses the gradient method for spatially resolved velocity computation.

Etienne-Cummings *et al.* (1999) implement a retina for measuring two-dimensional visual motion with two one-dimensional detectors. The pixels are built around a general-purpose analog neural computer and a silicon retina. Motion is extracted in two dimensions by using two one-dimensional detectors with spatial smoothing orthogonal to the direction of motion. In 2000, the same team presented a silicon retina chip with a central foveated region for smooth-pursuit tracking and a peripheral region for saccadic target acquisition (Etienne-Cummings *et al.*, 2000). This chip has been used as a person tracker in a smart surveillance system and a road follower in an autonomous navigation system.

6.4   From special-purpose chips to smart computational chips

One of the main drawbacks of the above-mentioned works is that these vision chips are not general-purpose. In other words, many vision chips are not programmable to perform different vision tasks. They are often built as special-purpose devices, performing specific and dedicated tasks, and not really reusable in another context (Dudek and Hicks, 2001). This inflexibility is particularly restrictive and even unacceptable for such vision systems that aim to flood several consumer markets. So, the main challenge when designing a smart vision system is to design a compact but versatile and fully programmable processing element, known as computational chip.

For this purpose, the processing function can be based on the paradigm of cellular neural networks (CNN), introduced by Chua and Yang in 1988 (Chua and Yang, 1988a; 1988b). CNN can be viewed as a very suitable framework for systematic design of image processing chips (Roska and Rodriguez-Vazquez, 2000). The complete programmability of the interconnection strengths, its internal image-memories, and other additional features make this paradigm a powerful front-end for the realization of simple and medium-complexity artificial vision tasks (Espejo *et al.*, 1996; Morfu *et al.*, 2008). Some proof-of-concept chips operating on preloaded images have been designed (Rekeczky *et al.*, 1999; Czuni and Sziranyi, 2000). Only a small amount of researches have integrated CNN on real vision chips. As an example, Espejo *et al.* (1998) report a 64 × 64 pixel programmable computational sensor based on a CNN. This chip is the first fully operational CNN vision-chip reported in literature which combines the capabilities of image-transduction, programmable image-processing and algorithmic control on a common silicon substrate. It has successfully demonstrated operations such as low-pass image filtering, corner and border extraction, and motion detection.

More recently, Galan *et al.* (2003) have focused on the development of CNN-based sensors with a chip including 1024 processing units arranged

into a 32×32 grid corresponding to approximately 500 000 transistors in a standard $0.5\,\mu m$ CMOS technology. An enhanced 128×128 version was also described in Rodriguez-Vazquez *et al.* (2004). The chip designed in a $0.35\,\mu m$ standard CMOS technology, contains about 3.75 million transistors and exhibits peak computing figure of 330 GOPS. Each processing element in the array contains a reconfigurable computing kernel capable of calculating linear convolutions on 3×3 neighborhoods in less than $1.5\,\mu s$, Boolean combinations in less than 200 ns, arithmetic operations in about $5\,\mu s$, and CNN-like temporal evolutions with a time constant of about $0.5\,\mu s$. Successive evolutions of these chips are presented on the historical roadmap depicted in Fig. 6.3.

However, hardware realization of such chips has turned out to be difficult because they suffer from large area and high power consumption. In the above-mentioned vision chips, the pixel size is often over $100\,\mu m \times 100\,\mu m$. Obviously, these dimensions cannot be considered as realistic dimensions for a real vision chip and numerous pioneering works have been abandoned to date. However, a major part of this crucial problem should be resolved in future years by using the new emergent CMOS technologies. Indeed, CMOS image sensors directly benefit from technology scaling by reducing pixel size, increasing resolution and integrating more analog and digital functionalities on the same chip with the sensor.

Other architectures in this category are the SCAMP family – SIMD current-mode analog matrix processor (Dudek, 2005; Dudek and Carey, 2006) of software-programmable SIMD (single instruction, multiple data) processor arrays implementing a variety of low-level image processing tasks. A SIMD processor array is built around multiple processing elements that simultaneously perform the same operation on different data. In the field of image sensors, the key idea is the introduction of an analog processing element (APE) per pixel, operating on the pixel value. The APE executes software instructions in a similar manner to a digital processor, but it operates on analog samples of data. The SCAMP-3 chip, described in Fig. 6.4, fabricated in a $0.35\,\mu m$ CMOS technology contains a 128×128 processor array and achieves cell density of 410 processors/mm^2 (a single cell measures under $50\,\mu m \times 50\,\mu m$).

The same team worked also on other complementary vision chips called ACLA (asynchronous cellular logic array) and ASPA (asynchronous/synchronous processor array). ACLA (Dudek, 2006; Lopich and Dudek, 2011) is an asynchronous cellular processor array that facilitates binary trigger-wave propagations, extensively used in various image-processing algorithms. A proof-of-concept array of 2460 cells has been fabricated in a $0.35\,\mu m$ CMOS process.

The ASPA family (Lopich and Dudek, 2008) includes vision chips embedding fine-grain processor arrays based on novel control schemes,

ACE400

20 × 22 cells
Black and white
Pixel memory: 4 B&W
Program memory: 8 templates
27.5 cells/mm^2
15.8 GOPS
I/O: 22 line binary

ACE4k

48 × 48 cells
Gray scale (7 bits)
Pixel memory: 4 B&W + 4 Gray
Prog. memory : 32 templates +
+ 64 Digital instructions
82 cells/mm^2
40 GOPS
I/O: 16 bit binary bus +
+ 16 line analog bus

ACE16k

128 × 128 cells
Gray scale (7 bits)
Pixel memory: 2 B&W + 8 Gray
Prog. memory : 32 templates +
+ 4096 Digital instructions
180 cells/mm^2
0,19 TOPS
I/O: 32 bit digital data bus
Embedded data ADCs
Multimode optical sensor

CACE1k

32 × 32 cells
Gray scale (7 bits)
Pixel memory: 4 B&W + 4 Gray
Prog. memory : 32 templates +
+ 64 Digital instructions
0,47 TOPS
I/O: 16 bit binary bus +
+ 16 lines analog bus
Intercoupled 2-layer architecture
Retina-like emulation capability

Vision system on-chip

Retina emulation

1999   2003

6.3 Historical roadmap of CNN-based vision chips (from Rodriguez-Vazquez *et al.*, 2004).
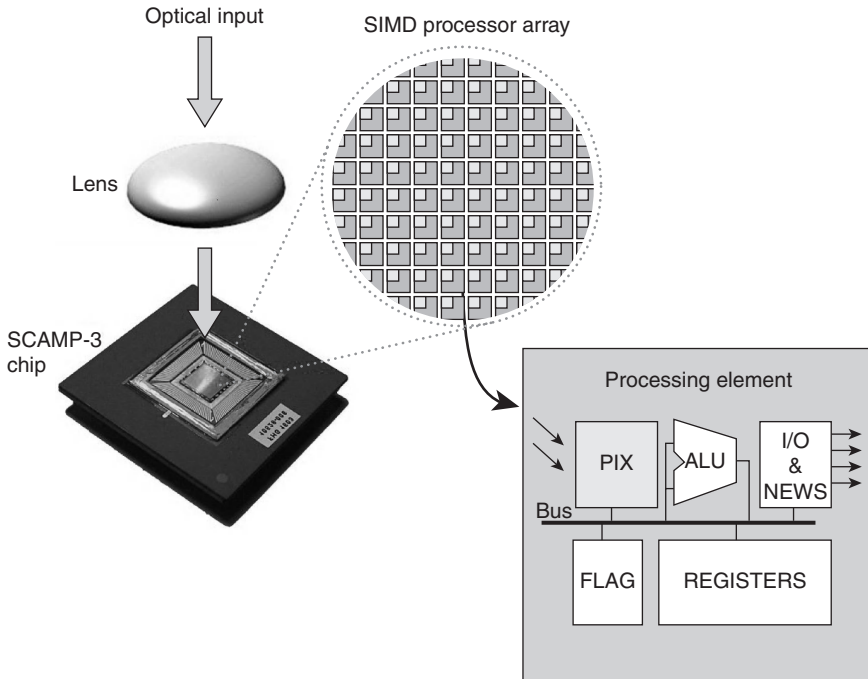
6.4 Overview of the SCAMP architecture (from Dudek *et al.*, 2006).

where individual processors are triggered, as data are available at their neighbors, optimizing speed and power consumption of the devices. The aim is to provide image processing engines suitable for both low-level, pixel-based operations (filtering, feature detection, etc.) as well as more global, object-based algorithms, such as object reconstruction, skeletonization, watershed transform, distance transform, etc. The latest chip in this family (ASPA-3) has a 160×80 processor array fabricated in a 180 nm CMOS technology with a chip area of 50 mm^2.

Another approach, which is potentially more programmable, is the PVLSAR (programmable versatile large scale artificial retina) retina chip (Paillet *et al.*, 1998; 1999). The PVLSAR is a highly integrated CMOS smart sensor device comprising an SIMD array of 128×128 pixel processors. Each pixel processor contains a photodiode as the optical sensor and a logical unit. The retina chip is a fine grain massively parallel SIMD processing unit with optical input. It is fully programmable and very powerful especially on low-level image processing. The PVLSAR can perform a plethora of retinotopic operations including early vision functions, image segmentation, and pattern recognition.

Table 6.1 Comparison of computational chips with on-pixel programmable analog processor

Chip	SCAMP-3 (Dudek and Carey, 2006)	MIMD IP Chip (Etienne-Cummings *et al.*, 2001)	ACE16k (Rodriguez-Vazquez *et al.*, 2004)
Technology	0.35 µm	1.2 µm	0.35 µm
Resolution	128 × 128	80 × 78	128 × 128
Die size	50 mm²	16 mm²	145 mm²
Pixel pitch	49.35 µm × 49.35 µm	45.6 µm × 45 µm	60 µm × 60 µm
Fill factor	5.6%	33%	n/a
Transistors / PE	128 tr.	9 tr.	198 tr.
Performance	20 GOPS	48 GOPS	330 GOPS
Power per chip	240 mW	74 mW	2.9 W
Image processing	Low-level image processing	Spatial convolutions	Spatial convolutions

Table 6.2 Comparison of computational chips with on-pixel programmable digital processor

Chip	ASPA (Lopich and Dudek, 2008)	VCS-IV (Komuro *et al.*, 2004)	PVLSAR2.2 (Paillet *et al.*, 1999)
Technology	0.35 µm	0.35 µm	0.8 µm
Resolution	128 × 128	64 × 64	128 × 128
Die size	213 mm²	49 mm²	76 mm²
Pixel pitch	100 µm × 117 µm	67.4 µm × 67.4 µm	60 µm × 60 µm
Fill factor	n/a	10%	30%
Transistors / PE	460 tr.	84 tr.	50 tr.
Performance	157 GOPS	n/a	49 GOPS
Power per chip	5.4 W	4 mW	1 W
Image processing	Low to mid-level convolutions	Low-level image processing	Low to mid-level image processing

To summarize, Tables 6.1 and 6.2, respectively show an overview of some representative analog and digital computational chips based on different alternatives for implementing vision processing at focal plane.

6.5 From video rate applications to high-speed image processing chips

The random access readout of CMOS image sensors provides the potential for high-speed readout and window-of-interest operations at low power consumption (El Gamal and Eltoukhy, 2005), especially when dealing with

low-level image processing algorithms. Indeed, such low-level image processing tasks are inherently pixel-parallel in nature. So, integrating a processing element within each pixel based on a SIMD architecture is a natural candidate to cope with the temporal processing constraints (Cembrano *et al.*, 2004). This approach is quite interesting for several reasons.

First, SIMD image-processing capabilities at the focal plane have not been fully exploited because the silicon area available for the processing elements is very limited. Nevertheless, this enables massively parallel computations allowing high frame rates up to thousands of images per second. The parallel evaluation of the pixels by the SIMD operators leads to processing times, independent of the resolution of the sensor. In a standard system, in which low-level image processing is externally implemented after digitization, processing times are proportional to the resolution leading to lower frame rates as resolution increases. Several papers have demonstrated the potentially outstanding performance of CMOS image sensors (Krymski *et al.*, 1999; Stevanovic *et al.*, 2000; Kleinfelder *et al.*, 2001). Krymski *et al.* (1999) describe a high-speed (500 frames/s) large format 1024×1024 APS with 1024 analog to digital converters (ADCs). Stevanovic *et al.* (2000) describe a 256×256 APS which achieves more than 1000 frames/s with variable integration times. Kleinfelder *et al.* (2001) describe a 352×288 digital pixel sensor (DPS) in which analog-to-digital (A/D) conversion is performed locally at each pixel, and digital data is read out from the pixel array in a manner similar to a random access digital memory, achieving 10 000 digital frames/s capturing and 1 giga-pixels/s for readout.

Secondly, the high-speed imaging capability of CMOS image sensors can benefit the implementation of new complex applications at standard rates and improve the performance of existing video applications such as motion vector estimation (Handoko *et al.*, 2000; Lim and El Gamal, 2001; Liu and El Gamal, 2001a), multiple capture with dynamic range (Yang *et al.*, 1999; Yadid-Pecht and Belenky, 2001; Stoppa *et al.*, 2002), motion capture (Liu and El Gamal, 2001b), and pattern recognition (Wu and Chiang, 2004). Indeed, standard digital systems are unable to operate at high frame rates, because of the high output data rate requirements for the sensor, the memory, and the processing elements. Integrating the memory and processing with the sensor on the same chip removes the classical input/output bottleneck between the sensor and the external processors in charge of processing the pixel values. Indeed, the bandwidth of the communication between the sensor and the external processors is known as a crucial aspect, especially with high-resolution sensors. In such cases, the sensor output data flow can be very high, and needs a lot of hardware resources to convert, process and transmit a lot of data. So, integrating image processing at the

pixel-level can alleviate the high data rate problem because the pixel values are pre-processed on-chip by the SIMD operators before sending them to the external world via the communication channels. This results in data reduction, which allows sending the data at lower data-rates, and reduces the effect of the computational-load bottleneck.

Thirdly, one of the main drawbacks to design specific circuits integrating sensing and processing on the same chip is that these vision chips are often built as special-purpose devices, performing specific and dedicated tasks, and not reusable in another context (Dudek and Hicks, 2001). So, it can be widely beneficial to integrate a versatile device, whose functionality can be easily modified. Moreover, except for the basic operations such as convolutions with small masks, the majority of computer vision algorithms require the sequential execution of different successive low-level image processing on the same data. So, each processing element must be built around a programmable execution unit, communication channels, and local memories dedicated to intermediate results. Because of the very limited silicon area, the processing units are necessarily very simple, providing the best compromise between various factors such as versatility, complexity, parallelism, processing speeds and resolution.

To sum up, the flexibility to integrate processing down to the pixel level allows us to rearchitect the entire imaging system to achieve much higher performance (El Gamal and Eltoukhy, 2005). The key idea is

1. to capture images at a very high framerate
2. to process the data on each pixel with a SIMD programmable architecture exploiting the high on-chip bandwidth between the sensor, the memory and the elementary processors and
3. to provide results at the best frame rate depending on the complexity of the image processing.

To illustrate this concept, we designed a massively parallel, SIMD vision chip implementing low-level image processing based on local masks (Dubois et al., 2008; Ginhac et al., 2010). The core includes a two-dimensional array of 64×64 identical processing elements (PEs). Each PE is able to convolve the pixel value issued from the photodiode by applying a set of mask coefficients to the image pixel values located in a small neighborhood. The key idea is that a global control unit can dynamically reconfigure the convolution kernel masks and then implement the most part of low-level image processing algorithms. This confers the functionality of programmable processing devices to the PEs embedded in the circuit. As seen in Fig. 6.5, each individual PE includes the following elements:

• a photodiode dedicated to the optical acquisition of the visual information and the light-to-voltage transduction
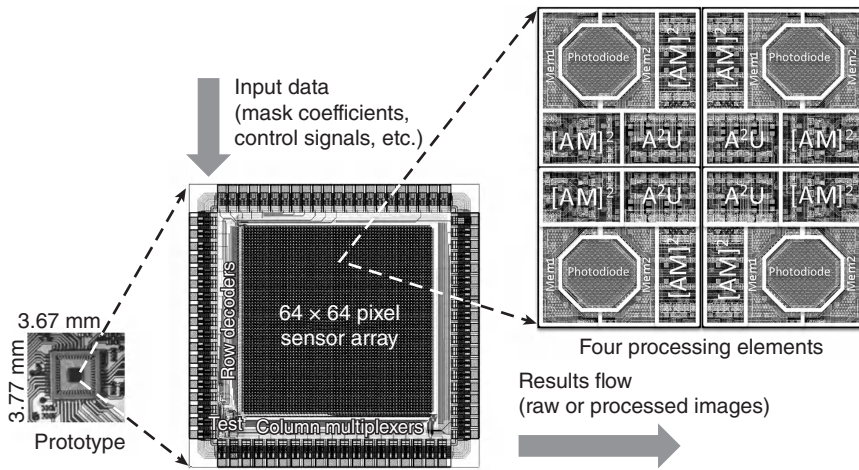
6.5 Overview of the high-speed vision chip with an analog processor-per-pixel array (from Ginhac *et al.*, 2008).

- a set of two analog memory, amplifier and multiplexer structures called $[AM]^2$, which serve as intelligent pixel memories and are able to dissociate the acquisition of the current frame in the first memory and the processing of the previous frames in the second memory, and
- an analog arithmetic unit named A^2U based on four analog multipliers, which performs the linear combination of the four adjacent pixels using a 2×2 convolution kernel.

In brief, each PE includes 38 transistors integrating all the analog circuitry dedicated to the image processing algorithms. The global size of the PE is $35\,\mu m \times 35\,\mu m$ ($1225\,\mu m^2$). The active area of the photodiode is $300\,\mu m^2$, giving a fill-factor of 25%. In terms of pixel size and fill-factor, this chip shares similar characteristics with the vision chips previously described in Tables 6.1 and 6.2. The chip can capture raw images up to 10 000 frames per second and runs low-level programmable image processing at a frame rate of 2000 to 5000 frames per second.

Based on the same principle, a large number of equivalent pixel-level image sensors have been designed during the past ten years, taking advantage of pixel-level processing elements to achieve massively parallel computations and thus, to exploit the high-speed imaging capability of CMOS image sensors. In an increasingly digital world, we can imagine that the most part of state-of-the-art imaging systems has become almost entirely digital, including analog to digital conversion and digital processing in the chip. But, for low-level image processing, an analog or a mixed-approach can offer superior performance leading to a smaller, faster, and lower power solution than digital processors (Martin *et al.*, 1998). Indeed, low-level

image processing usually involves basic operations using local masks. These local operations are spatially dependent on other pixels around the processed pixel. Since the same type of operation is applied to a very large data set, these low-level tasks are computationally intensive and require a high bandwidth between the image memory and the digital processor. Following this idea, analog vision chips such as those described in Brea *et al.* (2004), Rodriguez-Vazquez *et al.* (2004), Dudek and Hicks (2005), Chi *et al.* (2007), Massari and Gottardi (2007), or Kim *et al.* (2008) are characterized by a very compact area, optimized dedicated processing, high processing speed and impressive performance but suffers from low flexibility.

On the other hand, as integrated circuits keep scaling down following Moore's Law, recent trends show a significant number of papers discussing the design of digital imaging systems that take advantage of the increasing number of available transistors integrated in each pixel in order to perform analog to digital conversion, data storage and sophisticated digital imaging processing. Following the first digital pixel sensors designed by Kleinfelder *et al.* (2001), numerous works have been conducted to optimize this new design concept. While on-pixel conversion provides a number of advantages, there are still many challenges and issues to be solved and more particularly the dynamic range limitation due to the number of bits used for the conversion (Kitchen *et al.*, 2005), the optimization of the silicon area of the memory (Zhang *et al.*, 2011b), and the compression of the data (Zhang and Bermak, 2007). Then, digital data can be processed by specific digital computational elements such as those described in Komuro *et al.* (2004), Leon-Salas *et al.* (2007), Miao *et al.* (2008), Komuro *et al.* (2009), or Lin *et al.* (2009). Such vision chips offer more versatility and flexibility, more programmability and perform more complex algorithms from low- to mid level image processing.

6.6   Future trends

CMOS active-pixel image sensors have become increasingly mature because of continuous technological advances. They now dominate image sensor market shipments, both in volume and in revenue. Successive improvements have resulted in shrinkage of pixel size. State-of-the-art CMOS image sensors used typically in mobile devices are built with 1.4 μm generation pixels. However, with such pixel pitches, the number of incident photons is limited and sensor optical response can be blocked or interfered by metal layers in traditional front-side illumination (FSI) sensor structure. So, there is a growing interest in the mass production of BSI devices. BSI CMOS sensors have the metal wiring layer positioned below the photodiode layer which means light is not reflected by pixel wiring and then lost. Due to this

design, the photodiodes receive more light and the sensor is able to produce higher quality images in dark or low light scenes.

As photo-detector size shrinks, approaches that decouple sensing from readout and processing by employing separate stacked structures for photo-detection and processing is also of growing interest. The main idea is to exploit the features of the 3D technologies for the fabrication of a stack of very thin and precisely aligned CMOS APS layers, each of these stacks being optimized for a given function. 3D chips are obtained by segmenting 2D chips into functional blocks, stacking these blocks, and interconnecting them with short signal paths. As an example, a CMOS digital smart vision system consisting of a CMOS sensor, analog to digital converters and digital programmable processing elements is particularly suitable for a 3D integration. Combined with BSI technology, image sensor devices equipped with 3D technologies allow high-speed signal processing and have an optical fill factor of 100%. The optimization of the sensor, the ADC and the processors fabricated in different technologies offer the opportunity to improve the sensor performance and decrease the cost (Suntharalingam *et al.*, 2009; Motoyoshi and Koyanagi, 2009; Yeh *et al.*, 2011).

The ability to integrate dedicated signal processing functions within the pixel site is finally the most significant difference between CMOS and CCD sensors. In CMOS sensors, data processing can take place concurrently with image acquisition and can contribute to acquire complementary data, such as 3D information. Real time 3D imaging is a rapidly emerging field, due to the fact that 3D image acquisition systems can be used in a variety of applications such as automobile, robot vision systems, security and so on. Time-of-flight (TOF) image sensors using high frequency modulation of near-infrared light have emerged as a viable alternative to stereo and structured light imaging for capturing range information (Lee *et al.*, 2011). However, effort will have to be made to improve performance and precision that remain critical to mass adoption in consumer electronics applications.

Finally, the ultimate sensitivity in electronic imaging is the detection of individual photons. Single-photon imaging can be seen as the next step to reach in the design of smart imaging systems (Seitz and Theuwissen, 2011). Single-photon avalanche diodes (SPADs) are, as their name suggests, highly sensitive optical detectors capable of distinguishing single photons. They combine this high sensitivity with excellent photon timing properties in the range of a few tens of picoseconds. In the past, SPADs were designed using custom processes. Recent works have dealt about integration of SPADs on standard CMOS processes. One of the major challenges still remaining is the creation of large arrays of SPADs with reduced pitch and capable of single photon sensitivity and precise photon timing with sub-nanosecond resolution. Applications of such innovative sensors are evident in many fields: 3D scanning, medical imaging, molecular biology, astronomy and

aerospace applications, etc., and more generally in any situation requiring brilliant pictures captured under extreme low-light conditions.

6.7   Conclusion

In this chapter, we have introduced the fundamental concept of smart cameras on a chip or smart vision chips that simultaneously integrate on the same die image capture capability and highly complex image processing. Successive technology scaling has made possible the integration of analog or digital processing elements into single pixels.

To illustrate this continuous evolution, we chronologically surveyed three different categories of vision chips, exploring first the pioneering works on artificial retinas, then describing the most significant contributions to computational chips, and finally presenting the most recent state-of-the-art high-speed image processing chips able to perform complex algorithms at a high framerate. During this detailed survey, we have outlined the challenges of implementing complex image processing applications at focal-plane and the underlying complexity of resolving the fundamental tradeoffs between high-performance tailor-made chips and less powerful but more reusable programmable chips.

This chapter ends with particular focus on trending topics, mentioning recent technological innovations such as BSI and three-dimensional stacking. The democratization of these recent technological developments may lead to the design of very innovative smart image processing chips. We also mention the new challenges of designing efficient single-photon imaging chips able to sense every individual photon, both spatially and temporally with high precision.

6.8   References

Acosta-Serafini, P., Ichiro, M. and Sodini, C. (2004), A ⅓ VGA linear wide dynamic range CMOS image sensor implementing a predictive multiple sampling algorithm with overlapping integration intervals, *IEEE Journal of Solid-State Circuits*, 39, 9, 1487–1496.

Albani, L., Chiesa, P., Covi, D., Pedegani, G., Sartori, A. and Vatteroni, M. (2002), Visoc: A smart camera soc. In *Proceedings of the 28th European Solid-State Circuits Conference*, ESSCIRC 2002, 367–370.

Aw, C. H. and Wooley, B. (1996), A 128 × 128-pixel standard-CMOS image sensor with electronic shutter, *IEEE Journal of Solid State Circuits*, 31, 12, 1922–1930.

Bigas, M., Cabruja, E., Forest, J. and Salvi, J. (2006), Review of CMOS image sensors, *Microelectronics Journal*, 37, 5, 433–451.

Brajovic, V. and Kanade, T. (1996), A sorting image sensor: an example of massively parallel intensity-to- time processing for low-latency computational sensors. In *Proceedings of IEEE International Conference on Robotics and Automation, Vol. 2*. 1638–1643.

Brea, V. M., Vilariño, D. L., Paasio, A. and Cabello, D. (2004), Design of the processing core of a mixed-signal CMOS DTCNN chip for pixel-level snakes, *IEEE Trans. Circuits Syst. I, Reg. Papers*, 51, 5, 997–1013.

Cembrano, G., Rodriguez-Vazquez, A., Galan, R., Jimenez-Garrido, F., Espejo, S. and Domniguez-Castro, R. (2004), A 1000 fps 128 × 128 vision processor with 8-bit digitized I/O, *IEEE Journal of Solid State Circuits*, 39, 7, 1044–1055.

Chi, Y., Mallik, U., Clapp, M. A., Choi, E., Cauwenberghs, G. and Etienne-Cummings, R. (2007), CMOS camera with in-pixel temporal change detection and ADC, *IEEE J. Solid-State Circuits* 42, 10, 2187–2196.

Chua, L. and Yang L. (1988a), Cellular neural networks: theory, *IEEE Transactions on Circuits and Systems*, 35, 10, 1257–1272.

Chua, L. and Yang L. (1988b), Cellular neural networks: applications, *IEEE Transactions on Circuits and Systems*, 35, 10, 1273–1290.

Czuni, L. and Sziranyi, T. (2000), Motion segmentation and tracking optimization with edge relaxation in the cellular nonlinear network architecture. In *Proceedings of the 6th IEEE International Workshop on Cellular Neural Networks and Their Applications*, CNNA 2000, 51–56.

Decker, S., McGrath, D., Brehmer, K. and Sodini, C. (1998), A 256 × 256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output, *IEEE Journal of Solid-State Circuits*, 33, 12, 2081–2091.

Delbruck, T. (1993), Silicon retina with correlation-based, velocity-tuned pixels, *IEEE Transactions on Neural Networks*, 4, 3, 529–541.

Deutschmann, R. and Koch, C. (1998), An analog VSLS velocity sensor using the gradient method. In *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, ISCAS '98, Vol. 6, 649–652.

Dickinson, A., Mendis, S., Inglis, D., Azadet, K. and Fossum, E.R. (1995), CMOS digital camera with parallel analog to digital conversion architecture. In *Program of 1995 IEEE Workshop on CCDs and Advanced Image Sensors*, Dana Point, CA, 20–22, April.

Dubois, J., Ginhac, D., Paindavoine, M. and Heyrman, B. (2008), A 10 000 fps CMOS sensor with massively parallel image processing, *IEEE Journal of Solid-State Circuits*, 43, 3, 706–717.

Dudek, P. (2005), Implementation of SIMD vision chip with 128 × 128 array of analogue processing elements. In *IEEE International Symposium on Circuits and Systems*, ISCAS, Vol. 6, 5806–5809.

Dudek P. (2006), An asynchronous cellular logic network for trigger-wave image processing on fine-grain massively parallel arrays, *IEEE Transactions on Circuits and Systems II: Express Briefs*, 53, 5, 354–358.

Dudek P. and Hicks, P. J. (2001), An analogue SIMD focal-plane processor array. In *Proceedings of the International Symposium on Circuits and Systems*, ISCAS, vol. 4, 6–9, 490–493.

Dudek P. and Hicks, P. J. (2005), A general-purpose processor-per-pixel analog SIMD vision chip, *IEEE Trans. Circuits Syst. I*, 52, 1, 13–20.

Dudek P. and Carey, S. (2006), General-purpose 128 × 128 SIMD processor array with integrated image sensor, *Electronics Letters*, 42, 12, 678–679.

Dudek, P., Barr, D., Lopich, A. and Carey, S. (2006), Demonstration of real-time image processing on the SCAMP-3 vision system, *10th International Workshop on Cellular Neural Networks and Their Applications*, CNNA '06.

El-Desouki, M., Jamal Deen, M., Fang, Q., Liu, L., Tse, F. and Armstrong, D. (2009), CMOS image sensors for high speed applications, *Sensors*, 9, 1, 430–444.

El Gamal, A. and Eltoukhy, H. (2005), CMOS image sensors, *IEEE Circuits and Devices Magazine*, 21, 3, 6–20.

El Gamal, A., Yang, D. and Fowler, B. (1999), Pixel level processing – why, what and how? In *Proceedings of the SPIE Electronic Imaging '99 Conference*. Vol. 3650. 2–13.

Etienne-Cummings, R., Van Der Spiegel, J. and Mueller, P. (1999), Hardware implementation of a visual-motion pixel using oriented spatiotemporal neural filters, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 46, 9, 1121–1136.

Etienne-Cummings, R., Van Der Spiegel, J., Mueller, P. and Zhang, M. (2000), A foveated silicon retina for two-dimensional tracking, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 47, 6, 504–517.

Etienne-Cummings, R., Kalayjian, Z. and Cai, D. (2001), A programmable focal-plane MIMD image processor chip, *IEEE Journal of Solid-State Circuits*, 36, 1, 64–73.

Espejo, S., Carmona, R., Dominguez-Castro, R. and Roriguez-Vazquez, A. (1996), A 0.8 μm CMOS programmable analog-array-processing vision-chip with local logic and image-memory. In *Proceedings of the 22nd European Solid- State Circuits Conference*, ESSCIRC '96, 280–283.

Espejo, S., Dominguez-Castro, R., Linan, G. and Roriguez-Vazquez, A. (1998), A 64 × 64 CNN universal chip with analog and digital I/O. In *IEEE International Conference on Electronics, Circuits and Systems*, Vol. 1, 203–206.

Fontaine, R. (2011), A review of the 1.4 μm pixel generation. In *International Image Sensor Workshop*, Hokkaido, Japan, 1–4.

Fossum, E. (1993), Active pixel sensors: Are CCDs dinosaurs? *International Society for Optical Engineering* (SPIE), 1900, 2–14.

Fossum, E. (1997), CMOS image sensor: Electronic camera on a chip, *IEEE Transactions on Electron Devices*, 44, 10, 1689–1698.

Galan, R., Jimenez-Garrido, F., Dominguez-Castro, R., Espejo, S., Roska, T., Rekeczky, C., Petras, I. and Rodriguez-Vazquez, A. (2003), A bio-inspired two-layer mixed-signal flexible programmable chip for early vision, *IEEE Transactions on Neural Networks*, 14, 5, 1313–1336.

Gao, Q. and Yadid-Pecht, O. (2012), A low-power block-based CMOS image sensor with dual vdd. *IEEE Sensors Journal*, 12, 4, 747–755.

Ginhac, D., Dubois, J., Paindavoine, M. and Heyrman, B. (2008), A SIMD programmable vision chip with high speed focal plane image processing, *EURASIP Journal of Embedded Systems*, vol. 2008, Article ID 961315.

Ginhac, D., Dubois, J., Heyrman, B. and Paindavoine, M. (2010), A high-speed programmable focal- plane SIMD vision chip, *Analog Integrated Circuits and Signal Processing*, 65, 389–398.

Handoko, D., Takokoro, K., Kumahara, M. and Matsuzawa, A. (2000), A CMOS image sensor for local-plane motion vector estimation, in *Symposium of VLSI Circuits*, vol. 3650, pp. 28–29.

Hanson, S., Foo, Z., Blaauw, D. and Sylsverter, D. (2010), A 0.5 v sub-microwatt CMOS image sensor with pulse-width modulation read-out. *IEEE Journal of Solid-State Circuits*, 45, 4, 759–767.

Hong, C. and Hornsey, R. (2002), On-chip binary image processing with CMOS image sensors, *Proc. SPIE*, 4669, 125.

Kemeny, S.E., Panicacci, R., Pain, B., Matthies, L. and Fossum, E.R. (1997), Multiresolution image sensor, *IEEE Trans. Circuits and Systems for Video Technology*, 7, 4, 575–583 .

Kim, D., Cho, J., Lim, S., Lee, D. and Han, G. (2008), A 5000 S/s single-chip smart eye-tracking sensor. in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, 46–594.

Kitchen, A., Bermak, A. and Bouzerdoum, A. (2005), A digital pixel sensor array with programmable dynamic range, *IEEE Transactions on Electron Devices*, 52, 12, 2591–2601.

Kleinfelder, S., Lim, S., Liu, X. and El Gamal, A. (2001), A 10 000 frames/s CMOS digital pixel sensor, *IEEE Journal of Solid-State Circuits*, 36, 12, 2049–2059.

Komuro, T., Kagami, S. and Ishikawa, M. (2004), A dynamically reconfigurable SIMD processor for a vision chip, *IEEE J. Solid-State Circuits*, 39, 1, 265–268.

Komuro, T., Iwashita, A. and Ishikawa, M. (2009), A QVGA-size pixel-parallel image processor for 1000-fps vision, *IEEE Micro*, 29, 6, 58–67.

Kozlowski, L., Rossi, G., Blanquart, L., Marchesini, R., Huang, Y., Chow, G., Richardson, J. and Standley, D. (2005), Pixel noise suppression via SoC management of target reset in a 1920×1080 CMOS image sensor, *IEEE Journal of Solid-State Circuits*, 40, 12, 2766–2776.

Krymski, A. and Tu, N. (2003), A 9-v/lux-s 5000-frames/s 512×512 CMOS sensor, *IEEE Transactions on Electron Devices*, 50, 1, 136–143.

Krymski, A., Van Blerkom, D., Andersson, A., Bock, N., Mansoorian, B. and Fossum, E. (1999), A high speed, 500 frames/s, 1024×1024 CMOS active pixel sensor, *Digest of Technical Papers of Symposium on VLSI Circuits*, pp. 137–138.

Lee, T.Y., Lee, Y.J., Min, D.K., Lee, S.H., Kim, W.H., Kim, S.H., Jung, J.K., Ovsiannikov, I., Jin, Y.G., Park, Y.D., Fossum, E.R. and Chung, C.H. (2011), A 192×108 pixel ToF-3D image sensor with single-tap concentric-gate demodulation pixels in $0.13\,\mu$m technology, *IEEE International Electron Devices Meeting* (IEDM), 8.7.1–8.7.4.

Leon-Salas, W. D., Balkir, S., Sayood, K., Schemm, N. and Hoffman, M. W. (2007), A CMOS imager with focal plane compression using predictive coding, *IEEE J. Solid-State Circuits*, 42, 11, 2555–2572.

Lim, S. and El Gamal, A. (2001), Integrating image capture and processing – beyond single chip digital camera. In *Proceedings of the SPIE Electronic Imaging '2001 Conference*, Vol. 4306, San Jose, CA.

Lin, Q.-Y., Miao, W., Zhang, W.-C., Fu, Q.-Y. and Wu, N.-J. (2009), A 1000 frames/s programmable vision chip with variable resolution and row-pixel-mixed parallel image processors, *Sensors*, 9, 5933–5951.

Litwiller, D. (2001), CCD vs. CMOS: Facts and fiction, *Photonics Spectra*, 35, 154–158.

Liu, X. and El Gamal, A. (2001a), Photocurrent estimation from multiple non-destructive samples in a CMOS image sensor. In *Proceedings of the SPIE Electronic Imaging '2001 Conference*, Vol. 4306, San Jose, CA.

Liu, X. and El Gamal, A. (2001b), Simultaneous image formation and motion blur restoration via multiple capture. In *IEEE International Conference on Acoustics, Speech and Signal Processing*. Vol. 3, 1841–1844.

Loinaz, M., Singh, K., Blanksby, A., Inglis, D., Azadet, K. and Ackland, B. (1998), A 200mv 3.3v CMOS color camera IC producing 352 × 288 24-b video at 30 frames/s, *IEEE Journal of Solid-State Circuits*, 33, 12, 2092–2103.

Lopich, A. and Dudek, P. (2008), ASPA: Focal plane digital processor array with asynchronous processing capabilities. In *IEEE International Symposium on Circuits and Systems*, ISCAS 2008, 1592–1595.

Lopich, A. and Dudek, P. (2011), Asynchronous cellular logic network as a co-processor for a general purpose massively parallel array. *International Journal of Circuit Theory Applications*, 39, 9, 963–972.

Martin, D., Lee, H.-S. and Masaki, I. (1998), A mixed-signal array processor with early vision applications, *IEEE Journal of Solid-State Circuits*, 33, 3, 497–502.

Massari, N. and Gottardi, M. (2007), A 100dB dynamic-range CMOS vision sensor with programmable image processing and global feature extraction, *IEEE J. Solid-State Circuits*, 42, 3, 647–657.

Mead, C. and Mahowald, M. (1988), A silicon model of early visual processing, *Neural Networks*, 1, 1, 91–97.

Miao, W., Lin, Q., Zhang, W. and Wu, N.-J. (2008), A programmable SIMD vision chip for real-time vision applications, *IEEE J. Solid-State Circuits*, 43, 6, 1470–1479.

Moini, A. (2000), *Vision Chips*. Kluwer Academic, Boston; London.

Morfu, S., Marquié, P., Nofiélé, B. and Ginhac, D. (2008), Nonlinear systems for image processing. In *Advances in Imaging and Electron Physics*, P. W. Hawkes (ed.), Vol. 152, 79–151, Elsevier.

Motoyoshi, M. and Koyanagi, M. (2009), 3D-LSI technology for image sensor, *Journal of Instrumentation*, 4, 03, P03009.

Murari, K., Etienne-Cummings, R., Thakor, N. and Cauwenberghs, G (2009), Which photodiode to use: A comparison of CMOS-compatible structures, *IEEE Sensors Journal*, 9, 7, 752 –760.

Nixon, R. H., Kemeny, S. E., Staller, C. O. and Fossum, E. R. (1995), 128 × 128 CMOS photodiode-type active pixel sensor with on-chip timing, control, and signal chain electronics. In *Proceedings of the SPIE Charge-coupled Devices and Solid State Optical Sensors V*, M. M. Blouke (ed.), (Vol. 2415, 117–123).

Ohta, J. (2008), *Smart CMOS images sensors and applications*. CRC Press, Boca Raton

Paillet, F., Mercier, D. and Bernard, T. (1998), Making the most of 15k-lambda-2 silicon area for a digital retina. In *Proceedings of SPIE*, Vol. 3410, 158–167.

Paillet, F., Mercier, D. and Bernard, T. (1999), Second generation programmable artificial retina. In *Proceedings of the Twelfth Annual IEEE International ASIC/ SOC Conference*, 304–309.

Rekeczky, C., Tahy, A., Végh, Z. and Roska, T. (1999), CNN-based spatio-temporal nonlinear filtering and endocardial boundary detection in echocardiography. *International Journal of Circuit Theory and Applications*, 27, 1, 171–207.

Rinner, B. and Wolf, W. (2008), An introduction to distributed smart cameras, *Proceedings of the IEEE*, 96, 10, 1565–1575.

Rodriguez-Vazquez, A., Linan-Cembranon, G., Carranza, L., Roca-Moreno, E., Carmona-Galan, R., Jimenez-Garrido, F., Dominguez-Castro, R. and Meana, S. (2004), Ace16k: the third generation of mixed-signal SIMD-CNN ACE chips toward VSOCs. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 51, 5, 851–863.

Roska, T. and Rodriguez-Vazquez, A. (2000), Review of CMOS implementations of the CNN universal machine-type visual microprocessors. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, ISCAS 2000 Geneva, Vol. 2. 120–123.

Sakakibara, M., Kawahito, S., Handoko, D., Nakamura, N.,Higashi, M., Mabuchi, K. and Sumi, H. (2005), A high-sensitivity CMOS image sensor with gain-adaptive column amplifiers, *IEEE Journal of Solid-State Circuits*, 40, 5, 1147–1156.

Schrey, O., Huppertz, J., Filimonovic, G., Bussmann, A., Brockherde, W. and Hosticka, B. (2002), A 1 k × 1 k high dynamic range CMOS image sensor with on-chip programmable region-of-interest readout, *IEEE Journal of Solid-State Circuits*, 37, 7, 911–915.

Seitz, P. (2000), Solid-state image sensing, in Jähne, B., Haußecker, H. and Geißler, P. (eds) *Handbook of Computer Vision and Applications*, Vol. 1, Academic Press, San Diego, pp. 165–222.

Seitz, P. and Theuwissen, A. (eds) (2011), *Single Photon Imaging*, Springer Series in Optical Sciences, Springer-Verlag, Heidelberg.

Shi, Y. and Lichman, S. (2006), Smart cameras: A review, *CCTV Focus*, 37, 38–45.

Smith, S., Hurwitz, J., Torrie, M., Baxter, D., Holmes, A., Panaghiston, M., Henderson, R., Murrayn, A., Anderson, S. and Denyer, P. (1998), A single-chip 306 × 244-pixel CMOS NTSC video camera. In *ISSCC Digest of Technical Papers*, San Fransisco, CA, 170–171.

Stevanovic, N., Hillebrand, M., Hosticka, B. and Teuner, A. (2000), A CMOS image sensor for high-speed imaging, *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, pp. 104–105.

Stoppa, D., Somoni, A., Gonzo, L., Gottardi, M. and Dalla Betta, G.F. (2002), Novel CMOS image sensor with a 132-dB dynamic range, *IEEE Journal of Solid-State Circuits*, 37, 12, 1846–1852.

Sumi, H. (2006), Low-noise imaging system with CMOS sensor for high-quality imaging. In *IEDM '06 International Electron Devices Meeting*, 1–4.

Suntharalingam, V., Berger, R., Clark, S., Knecht, J., Messier, A., Newcomb, K., Rathman, D., Slattery, R., Soares, A., Stevenson, C., Warner, K., Young, D., Lin, Ping Ang, Mansoorian, B. and Shaver, D., (2009), A 4-side tileable back illuminated 3D-integrated Mpixel CMOS image sensor, *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, ISSCC 2009, 38–39,39a.

Wu, C.-Y. and Chiang, C.-T. (2004), A low-photocurrent CMOS retinal focal-plane sensor with a pseudo-bit smoothing network and an adaptive current schmitt trigger for scanner applications, *IEEE Sensors Journal*, 4, 4, 510–518.

Wuu, S., Chien, H., Yaung, D., Tseng, C., Wang, C., Chang, C. and Hsaio, Y. (2001), A high performance active pixel sensor with 0.18 lm CMOS color imager technology. In *Electron Devices Meeting, 2001. IEDM Technical Digest International*, pp. 555–558.

Xiao, F., Farrell, J., Catrysse, P. and Wandell, B. (2009), Mobile imaging: The big challenge of the small pixel. In *Digital Photography V, Proceedings of the SPIE Electronic Imaging '2009 Conference*, Vol. 7250, January 2009.

Xu, C., Zhang, W., Ki, W.-H. and Chan, M. (2002), A 1.0-v vdd CMOS active-pixel sensor with complementary pixel architecture and pulsewidth modulation

fabricated with a 0.25-µm cmos process, *IEEE Journal of Solid-State Circuits*, 37, 12, 1853–1859.

Yadid-Pecht, O. and Belenky, A. (2001), Autoscaling CMOS aps with customized increase of dynamic range. In *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, ISSCC 2001, 100–101.

Yadid-Pecht, O. and Belenky, A. (2003), In-pixel autoexposure CMOS APS, *IEEE Journal of Solid-State Circuits*, 38, 8, 1425–1428.

Yang, D., El Gamal, A., Fowler, B. and Tian, H. (1999), A 640 × 512 CMOS image sensor with ultra wide dynamix range floating-point pixel-level ADC, *IEEE Journal of Solid-State Circuits*, 34, 1821–1834.

Yeh, S., Lin, J., Hsieh, C., Yeh, K. and Li, C. (2011), A new CMOS image sensor readout structure for 3D integrated imagers. In *IEEE Custom Integrated Circuits Conference* (CICC), 1–4.

Yoon, K., Kim, C., Lee, B. and Lee, D. (2002), Single-chip CMOS image sensor for mobile applications, *IEEE Journal of Solid-State Circuits*, 37, 1839.

Zhang, M. and Bermak, A. (2007), Compressive acquisition CMOS image sensor: From the algorithm to hardware implementation, *IEEE Trans. Very Large Scale Integr. (VLSI) Syst*, 18, 490–500.

Zhang, W., Fu, Q. and Wu, N.J. (2011a), A programmable vision chip based on multiple levels of parallel processors, *IEEE Journal of Solid-State Circuits*, 46, 9, 2132–2147.

Zhang X., Leomant S., Lau K.L. and Bermak A. (2011b), A compact digital pixel sensor (DPS) using 2T-DRAM, *Journal of Low Power Electronics and Applications*, 1, 1, 77–96.

7
Complementary metal-oxide-semiconductor (CMOS) image sensors for mobile devices

R. J. GOVE, Aptina Imaging, USA

Abstract: This chapter discusses the application of CMOS image sensors in the mobile market. The chapter first reviews the history and how innovation has overcome the limitations of core technologies to enable mobile imaging. A review of the CMOS image sensor architectures and product considerations provides insight into how the basic image sensor evolves to a smart camera imaging solution, featuring key functions beyond picture taking into full 'imaging for information.' The chapter then discusses emerging technologies that will transform mobile imaging to enable unique and useful applications, which will fuel market growth and broaden the adoption of imaging technology across the globe.

Key words: CMOS image sensors, camera phones, computational imaging, imaging for information, mobile imaging, smartphone.

7.1   Introduction

Imaging for mobile phones has become the primary showcase example of rapid adoption of a new imaging technology, complementary metal-oxide-semiconductor (CMOS) imaging, into an existing market. Image sensors for mobiles became the fastest growing product category of semiconductor wafers in history, ramping to nearly a million silicon wafers or the equivalent capacity of five wafer fabrications (fabs) over the period of 2003 to 2005. This level has been sustained and has grown for at least seven years.

This section describes the technology enablers that moved image capture to the second most-used feature (after voice) and a 'billion-unit-per-year' market, the ability to capture and share pictures using one's mobile phone by almost anyone in the world at any time. Later applications of CMOS mobile imaging, including notebooks, tablets and handheld gaming devices, drove new capabilities like ultra thin/slim form factors and HD video capabilities. Going forward the image sensor enables improved control of these mobile devices with easy-to-use complex gesture control of the mobile phones. The primary scene-facing camera of the mobile phone typically features a very high-resolution camera for general photography or videography needs. However, the secondary, user-facing video camera found on most smartphones use CMOS image sensors similar to those used

191

in notebooks, tablets and gaming devices, with the purpose of capturing self portrait images with friends, or video conferencing. We will explore both sensor types.

7.1.1 Evolution of camera phones with complementary metal-oxide-semiconductor (CMOS) image sensors

Camera phones emerged when semiconductor, optics and packaging technology advanced to enable creation of miniature cameras. Simultaneously, mobile phones proliferated to a point where 100s of millions of users kept their mobile phones within an arm's length and virtually continuous usage. Around 1999, the combination of these miniature cameras with the mobile phone created what has become the largest and continuously growing imaging market ever, with many billions of camera phones in use producing trillions of pictures and billions of video clips. Huge portions of the global population first experienced photography with a camera phone. The scale of the mobile market demands production of over six million sensors per day, two billion per year, or over two hundred thousand 300-mm wafers per month. A key sector, the smartphone segment, where camera performance leads, will increase dramatically in the coming years. Imagine the number of processor chips or display devices in the world and the growth rate of those. Each will be connected to at least one image sensor. Clearly, we anticipate the CMOS image sensor market will continue at double digit compound growth per year. Figure 7.1 shows the predicted growth of the
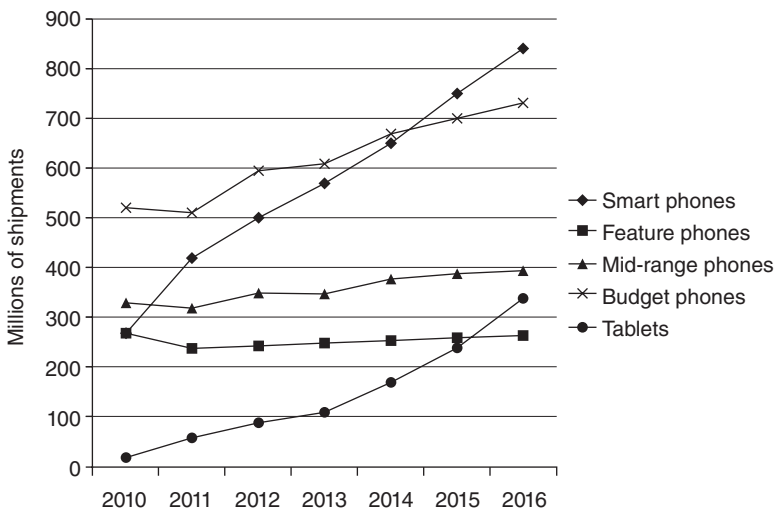
7.1 Market growth of mobile phones and tablets growing to 2.6 billion shipments in 2016. Source: Forward Concepts, June 2012.

mobile phone and tablet market, most of which will be camera-enabled with one or in many cases with two image sensors (Forward Concepts, 2012).

With camera phones, more than any other camera device, simple and convenient photography result due to their small size, flatness, low power, low-cost, and wireless connectivity. For mobile cameras to achieve the size and cost requirements, designers initially compromised image and/or video quality to gain the other benefits. However, the technology continues to improve generation by generation such that camera phones have moved from 'good-enough' image quality, to 'very good' image quality today.

Perhaps the most significant aspect of camera phones originates from their simplicity of use and range of possible applications, opening seemingly limitless new uses and markets. As well, we project the technology will advance to a point where mobile imaging with camera phones will exceed all other camera devices in their ability to capture compelling images, especially when taking full advantage of the interconnected, mobile nature of the cameras, and accessibility of massive amounts of memory and compute cycles emerging in mobile-phones and cloud-based computing. Ultimately, the challenge of creating a camera that captures compelling images by virtually any user in difficult and varying capture situations, like dark scenes or with rapidly moving photographers, remains an unsolved challenge of mobile imaging.

Here we discuss the limitations of the technology, the solutions that overcame technology limitations, and trends for the future.

7.1.2 Evolution of an application-rich camera phone

Mobile imaging, using smartphones, notebooks and tablets featuring embedded image sensors, has become the newest 'platform of innovation' in the industry. The innovation platform derives from a coupling of small high performance mobile image sensors with advanced image processors and fast communication networks, all readily under software control, inspiring hundreds or thousands of developers to create interesting new imaging applications. With an ideal combination of attributes, mobile imaging offers application developers quick access to a large installed base of cameras, and customers for their applications. These camera/sensor attributes include: small/thin form factor, low power using digital CMOS technology, easy software access to powerful image and video processors inside the mobile device, and ubiquitous interconnection to others using social media services and cloud computing with wireless connectivity. Analogous to the PCs of the 1980s and 1990s, these new applications of the mobile imaging platform will explode to billions of users. This section explores the technology and sensor architecture implications as new applications emerge. Examples include:

- mobile scanning at the molecular level with a sensor attached to mobile computing devices like mobile phones and tablets, and
- computational imaging for high dynamic range photography, where the final image or information is computed from multiple images or sensor arrays.

Uniquely, the experience of imaging with camera-enabled mobile phones can offer immediate and convenient capture of pictures or video clips, by almost anyone, anywhere. Smartphones create a unique imaging application platform. They integrate a camera with a super-computer, multi-core processor with giga-floating-point-operations-per-second (GFLOPS) computational speeds and optionally wirelessly connect the captured or processed information to the cloud for further processing. This creates perhaps the most powerful and versatile consumer imaging device created to date.

With camera-enabled mobile phones, an industry of application developers can easily deliver a wide range of imaging innovative applications. Figure 7.2 highlights many such applications. Examples include:

- automatic face detection for immediate cataloging or tagging of photos (Klik™ by Face.com)
- panoramic stitching of a sequence of images
- automatic removal of 'moving' unwanted objects in a single scene by using a short sequence of captured pictures and 'selective' usage of portions of the scene for cases where the object moved away (Clear-View™ from Scalado®), and
- creating the perfect picture of a group with a similar 'selective' usage of portions of the scene, but selecting the best picture in a sequence for each member with a perfect composition resulting (Rewind™ from Scalado®).

Augmented reality applications fuse scenes seen by the camera-enabled phone with 3D generated images by geometrically aligning the two image streams, one synthetic and one captured image (Henrysson, 2007). Phone-embedded global positioning system (GPS) chips, accelerometers and image recognition algorithms enable the phone to align or fuse the real and synthetic images, followed by selective overlay of useful data, such as interactive visual directions through a subway station, or in the case of ReFlow™ game, overlays flowing objects around your body. As well, camera phones can be programmed to identify the owner of the phone and essentially automatically upload a photo to an online account when the user portrays a particular gesture. Another social camera application can, with image recognition or vision algorithms, automatically access your friend's or contact's social networking applications, tagging pictures taken with the mobile phone with text tags streamed from their social networking sites

Camera-enabled smartphone	Picture taking 1. Voice actuated shutter 2. Stabilization (Luma Cam app) 3. Slow shutter for motion blur (slow shutter app) 4. High-dynamic range and tone mapping (Instagram.com) 5. Beam picture anywhere using quick response(QR) code (Scalado^R Photo Beamer^{TM})
	Video/picture editing 1. Automatic face detect for cataloging photos 2. ClearView^{TM} app to remove moving objects (Scalado.com)
	Information capture for application control 1. Snap picture, then look up media content (SnapTell.com) 2. Augmented reality word recognition and translation (WordLens) 3. Room measurement for remodeling (Metaio.com) 4. Camera-based gesture control (Eyesight.com) 5. Information enhanced picture searching (Pixolu.com) 6. Augmented reality navigation (AccessAir.com, Wikitude.com)
Mobile hand gesture detection	Multi-shot picture fusion 1. Panoramic image stitching (Cloudburst Autostitch Panorama) 2. Stitching multi-shots into collage (Boinx.com, Photosynth) 3. Perfect group shot composite of best poses (Scalado's Rewind)
	Computational imaging and vision apps 1. User recognition for social camera (Viewdl.com) 2. Augmented reality game with virtual liquid flow by user orientation (ReFlow)

7.2 Mobile applications for camera-enabled smartphones (top photo by istockphoto.com file#000019780418; bottom photo by Eyesight at eyesight-tech.com, used with permission).

(the SocialCamera^{TM} app from Viewdl.com can tag your photos and upload them to Facebook on the fly). These applications are emerging for the reasons above, easy image or information capture, and abundant processing readily available, all in a rich application development environment.

7.1.3 Image quality performance race

The market inflection point for mobile imaging occurred in the early 2000s when the low-light sensing quality and resolution of low-cost CMOS sensors became 'good-enough' for most users. Charge completed device (CCD) sensors were also employed in early camera phones, but the cost and power associated with CCDs could not survive in a scale and cost-driven market. As well, with CMOS, complete integrated camera solutions or systems-on-a-chip (SOCs) were viable with a single chip camera. In that case, a fully tuned CMOS camera that automatically optimized exposure, white balance,

color balance and others, created an impressive camera. With those cameras, good pictures or video were nicely captured in dimly lit situations. Another inflection point occurred in 2005 when camera phones, such as the distinctive Nokia N95™, were created to virtually achieve in many cases the same experience as a consumer digital still camera (DSC) product, with exceptional quality at sufficient resolution with 5 megapixels so the user could edit the images after capture. Today, camera phone manufacturers typically use 8, 10 and now 13 megapixel image sensors, with higher resolutions predicted.

Ultimately the camera became the second or third most-used feature of mobile phones, behind voice and texting. Mobile phone developers found new ways to differentiate their products by skillful design and integration of the camera and display components to create compelling images. The choice of lens, sensor, memory, display and processors defined a competitive camera-phone platform. That coupled with hundreds of application developers creating new ways to capture, manipulate, and share those pictures over the Internet, offered convenience and value.

In the future, we expect combining these approaches with a mobile camera-phone platform will yield picture-taking experiences that rival the best consumer DSLR (digital single-lens reflex) cameras available, yet be smaller and easier to use, without regard to how the picture was taken. Yet most camera users will choose camera-phones over other cameras simply because their camera-phone will nearly always be within an arm's reach, and ready to use.

7.1.4 CMOS sensor drafting charge coupled device (CCD) and mainstream CMOS technology

As engineers, we solve problems. As innovators we innovate on a continuum of building upon the accomplishments of others before us. CMOS image sensors evolution was exactly that, building on the successes and innovations from the CCD image sensor generation, with key enhancements. From the pixels, to the analog-to-digital converters (ADCs), to today's back-side illumination (BSI) technology, most CMOS solutions leverage the ideas implemented with totally different generations of semiconductor technology decades ago. Our ability to rapidly apply those advances and 'draft' fundamental industry-standard CMOS semiconductor technology defined the creation of this world's largest application of image sensors, the mobile camera phone.

On the other hand, CCD technology did not successfully draft the semiconductor industry, setting and leveraging a unique path, and as such has been overcome by CMOS. Figure 7.3 shows how CMOS has overcome CCD technology, particularly in pixel size, critical to mobile applications.

Pixel size (μm)

10

Moore's Law for pixels

▲ 5.2

4.2

CMOS

CCD s

3.2

2.8

2.2

1.75

1.4

Light
guide

BSI

?

1

1989 1992 1995 1998 2001 2004 2007 2010
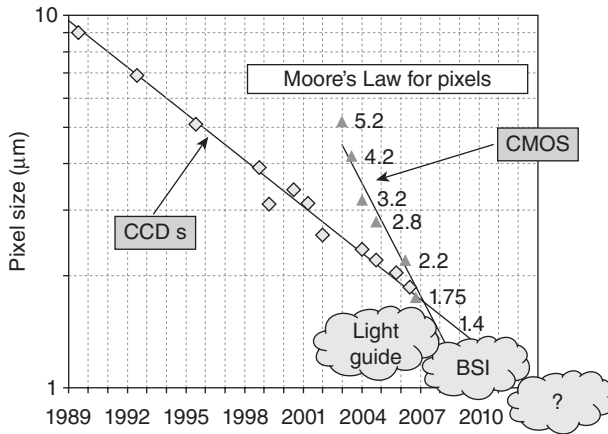
7.3 CMOS imaging Moore's Law trend.

Success was defined by the proper timing of leveraging high volume, low-cost, stable semiconductor processes, pragmatically optimizing the size, cost and performance of miniature cameras. As well, Fig. 7.3 shows the progression of pixel size and, since the smallest CMOS pixels in products today are larger than 1 um in size, Fig. 7.3 also highlights the recent slowing of this pixel size advancement. Conventional pixel architectures moved to the use of special light guides per pixel, to BSI technology, and to a somewhat uncertain future of solutions that we discuss in Section 7.5.

7.2   Core image/video capture technology requirements and advances in mobile applications

From the first digital camera in the 1970s, new applications have emerged due to the ever-shrinking size of the cameras, increasing picture taking performance, and ability to operate longer with lower power consumption. These performance, cost, size and power trends continue to challenge the 'Moore's Law of Imaging' (analogous to the 'Moore's Law' of semiconductors), which defines advancement and scaling of technology by 18-month generations, fundamentally powered by core silicon technology advancements in density and power, yet historically limited by optical advancements. This section reviews the aspects of sensor power consumption, silicon area usage, image capture performance, pixel design and others, as they were advanced by the industry to solve the critical challenges of mobile applications.

The widespread deployment of image sensing technology across billions of cameras happened by advancement of a broad spectrum of technologies,

Categories for mobile CMOS imaging core technologies

Process	Pixel	Sensor	Image computing and system	Optics and packaging

• Dark current
• Hot pixels
• Defects
• Dzero
• Die size
• Light guides
• BSI
	• 2.2 µm, 1.75 µm
	  1.4 µm, 1.1 µm
	  0.9 µm pixels
	• FSI and BSI
	• Low light
	  sensitivity
	• Dark current
	• High-dynamic
	  range (HDR)
	• Light guide
	• Color-filter-array
	  (CFA)
	• Micro-lenses
		• Row or column
		  noise
		• Defect replace
		• ADC speed/
		  power/watt
		• High-dynamic
		  range (HDR)
			• Color processing
			• Shading
			  correction
			• Fast auto focus
			• Image
			  stabilization
			• Zero shutter lag
			• User tracking
			• Co-processor
			  ASIC
			• 2-chip stacking
			• Software support
			• Fast interface
			  (HiSPi,MIPI, etc.)
				• High chief-ray
				  angle (CRA)
				  tolerance
				• Fast voice coil
				  auto focus
				• Image
				  stabilization
				  support
				• Low-cost
				  manufacturing
				• Re-flowable
				  module
				• Wafer-level
				  module
				• Chip-scale
				  package

Technology constraints: image capture performance, cost, size, and feature requirements

7.4 Categories of camera phone core technology advancements.

rather than simply one or two areas. A simple model of a typical camera for mobile applications and the areas of critical technology advancements is the basis for Fig. 7.4. Core technology areas include semiconductor process, pixel, sensor, image processing, optics and packaging to optimize the system performance. The following sections describe some of the key advancements in these areas.

7.2.1   The pixel for mobile applications

Perhaps one overarching element that determines the image capture performance of an image sensor is the array of pixels. A simple model for each pixel is shown in Fig. 7.5, whereby each pixel contains an optical pathway to a photodiode, the photodiode which provides a site or region for photon recombination, and a method to transfer charge out of the pixel (usually via a transistor and transfer gate). While the pixel mostly determines image capture performance, three items mostly determine the performance of the pixel:

1. the underlying optimization of the fab process to create successful operation (photon collection, storage and data transfer) of the pixel
2. the architecture of the pixel, and
3. the analog circuits for readout of the pixels.

Incident light

Microlens and planarization
Color filter
Passivation
Metal
Oxide layers
Photo-diode
Silicon
e^- e^- e^- e^- e^- e^- e^-

Conventional pixel architecture

Incident light

Microlens and planarization
Color filter
Passivation
Light guide
Light guide
Metal
Oxide
Anti-reflective coating
Deep photo diode
e^- e^- e^- e^- e^- e^-
Silicon

FSI Aptina A-pix™ architecture

Incident light

Microlens and planarization
Color filter
Anti-reflective coating
Deep photo diode
e^- e^- e^- e^- e^- e^- e^- e^-
Silicon
Metal
Oxide layers

BSI pixel architecture

7.5 Mobile pixel architectures and technology advancement.

For mobile applications, the limitations were plentiful 10 years ago. Initial mobile sensor pixels were 5.6 μm on a side, producing a VGA image with a ¼″ optical format array. By far, the ¼″ optical format had dominated the image sensor markets due to an effective compromise of performance, cost and size. The industry evolved to a performance leading node at ⅓″ or larger formats, a mainstream node at ¼″, and a value node at ⅕″ or smaller. Some technology areas which impeded success in this period included color performance, crosstalk between pixels, uniformity across the sensor, color shading, and eclipse (Murakami *et al.*, 2005).

Today, the challenge centers on maintaining 'scaled performance' as pixels shrink toward the diffraction limit of light. We have been able to retain scaled performance at 1.4 μm pixels, yet 1.1 μm and smaller pixels remain a challenge. Techniques employed to achieve scaled performance include low-crosstalk color filter array (CFA) methods. Nano-semiconductor technologies and organic films could create significant improvement in the sensitivity of image sensors, key to continuation of pixel shrinks to 0.9 μm and beyond.

7.2.2 The CMOS camera semiconductor process

We have discussed the advantages of 'drafting' CMOS process technology, producing image sensors by leveraging industry standards and processes already running in fab facilities. The term 'drafting' means we leverage the learning from previous developments. For imaging, the foremost limitations of this 'drafting' of a baseline CMOS technology relates to achieving

- lower dark (ld) current and
- the minimization of hot pixels.

Dark current creates a fundamental noise signal, which builds or accumulates over time within the pixel's photo diode. As well, it worsens with temperature. This noise signal adds to the primary, photon-created electrons, creating a noisy captured picture. Hot or white pixels emerge by either defects or other sources in the silicon to saturate the pixel independent of incident light. A clean fab, with a low defect density process, as specified by the Semiconductor Industry Association's International Technology Roadmap for Semiconductor (ITRS) zero defects (Dzero), coupled with attention to process and pixel optimization can minimize Id and hot pixels. Defects originate from a variety of sources, but mostly from particulate contamination, processing variations or misalignments. In process, failures can result from implantation, lithography, etching, planarization, etc. Stress induced defects (like wafer warp at back-side oxide removal or shear stress at wet oxidation), process steps (like laser annealing), and limitations in the photodiode can create defects as well. The use of gettering in some cases can

'repair' and reduce dark current. The use of a pinned-photodiode in the pixel dramatically reduces dark current.

CMOS process geometries used to produce mobile sensors range from 180 nm to 65 nm today, depending on the performance, cost, size, and power needed for the sensor. The smallest, highest-density pixels, logic and memory are possible with 65 nm process geometries. As pixel performance remains a key performance criterion for success in mobile applications, pixel designers usually push the limits of the technology node to achieve good performance. Similar to other semiconductor devices, the manufacturing yield (the percentage of defect-free die per wafer) and die size determines the cost of the sensor at a particular process node, with each node scaled by the cost to operate the fab. As well, a larger die with smaller dimensions will not yield as well as a smaller die. Many video graphics array (VGA) resolution CMOS sensors with larger pixels have been fabricated with nearly 100 percent yield per wafer, whereas larger sensors may only achieve 80 percent yield in some cases. Clearly, attention to the process optimization can greatly influence the resultant yield and cost.

After silicon front-end fabrication, CMOS image sensors are taken to another wafer manufacturing line that can include spin-coat/bake for creating a color filter array (CFA), microlens array, and finally packaging with a cover glass. The CFA and microlens arrays selectively steer red, green or blue photons across the sensor's optical plane on to a defined grid of photosensors or pixels so the sensor can measure color intensity across the image plane. A Bayer color filter pattern uses twice as many green as red or blue pixels to mimic the performance of the human visual system with more sensitivity in the green. For mobile imaging, the Bayer pattern has been the best trade-off of sensitivity, cost and color performance. To manufacture the sensor, typically green CFA is applied first, followed by red and blue filter processing. Years ago, a key CFA performance challenge was resolved by using a recessed CFA. In practice, an image sensor uses dark reference pixels as a dark ring, and especially to the left and right of the pixel arrays. This creates a black reference for measuring the pixel signals with respect to these dark pixels. In practice a metal layer as well blocks light to the dark photodiodes. However, this metal layer essentially pushes the nominal pixel's CFA away from the silicon surface, reducing quantum efficiency (QE) and increasing crosstalk. As a result, a recessed CFA was developed to improve performance. In the end, a challenging two level planar CFA was necessary to boost performance. Finally, micro-lenses are created on top of the CFA with a coat, bake and expose process.

Two technology developments enabled the evolution of high performance 1.4 μm pixels:
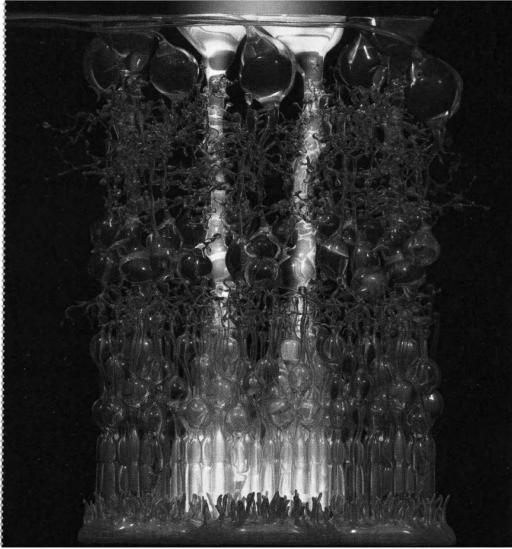
7.6 Light guides, as found in nature, optical fibers in the retina (Muller cells, used with permission).

1. the light guide (LG) process technology illustrated in Fig. 7.5 as A-Pix™, and as found in nature in Fig. 7.6 (Franze *et al*., 2012), and
2. wafer-level BSI process technology (Iwabuchi *et al*., 2006).

BSI is necessary for acceptable performance at pixels sizes of 1.1 um and smaller. LGs more efficiently move photons into the photodiodes, rather than lose the photons to reflection or absorption above the photodiode. These reflected photons can lead to lower quantum efficiency or crosstalk. BSI flips the sensor over so that the photodiodes are then on the surface of the sensor. Pixels designed with BSI process have substantial benefit with regards to the pixel's fill factor (or percentage of pixel area dedicated to light collection). Essentially the entire pixel has usable storage capacity, as there are no transistors in the optical path to disrupt the filling of the photodiode with photons. With BSI, pixel readout can create more pull and less lag on the photodiode. The pixel reset structure can use lower resistance metal for pixel reset, boosting performance (Wakabayashi *et al*., 2012).

The challenges of wafer-level BSI range from complex processes operating on ultra-thin wafers following grinding of the backside of the wafer to essentially bring the pixels to the back surface for higher fill factor and closeness to the micro-optics. As seen in Fig. 7.7, after thinning, the wafers are less than the thickness of paper allowing one to literally see through them. Essentially the wafer is only a few microns thick vs a typical front-side imaging (FSI) wafer that's typically thinned to 100 microns in thickness. As
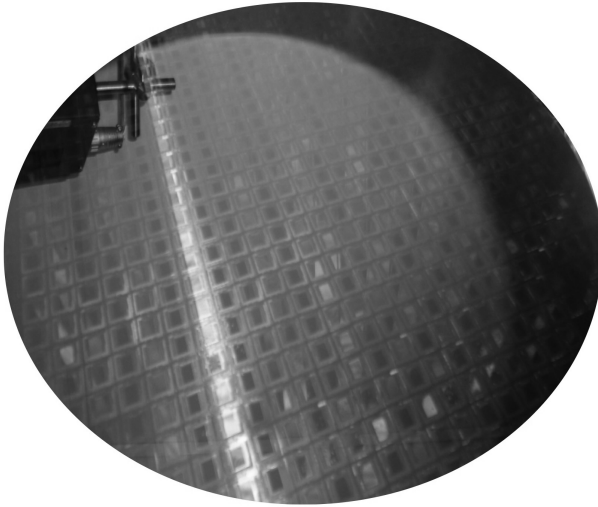
7.7 Photograph of a thinned BSI wafer, only several microns thick.

a result the BSI wafer is attached to a thin carrier wafer, introducing stretching and warping, and ultimately causing problems as CFA and microlenses are attached to the wafer. Complexity and cost remain a challenge for industry-wide adoption of the wafer-level BSI technology.

7.2.3   The CMOS sensor's pixels and pixel array

Most limitations in image capture performance of a camera phone result from design of the sensor's pixel and the optics. Figure 7.8 and Table 7.1 illustrate typical ranges of possible sensors from a pixel size, resolution and optical format perspective for mobile products. The challenge for the industry has been continuing the advance to smaller pixels and higher resolution, such as the future product depicted in Fig. 7.8 as 28 megapixel, 0.9 μm pixel, $\frac{1}{2.5}''$ format. As discussed in Section 7.2.1, the design of the pixel includes a photodiode, circuitry for storage and readout, and an optical pathway that can include light guides, color filters and micro-lenses.

Perhaps the fundamental technology barrier that was overcome to enable CMOS imaging was the pinned photodiode architecture and methods to eliminate noise sources, such as KTC noise. Early virtual-phase CCD advancements by Hynecek (1979) and the pinned photodiode by Teranishi *et al*. (1982) used in CCDs for low dark current, created a foundation for CMOS imaging. White *et al*. (1974) introduced the original correlated-double sampling (CDS) for fixed pattern noise reduction with CCDs and improved analog CDS by Hynecek (1988) and Tanaka *et al*. (1989). Yang *et al*. (1999) then introduced the digital CDS method. Later development

Image sensor area diagrams
(pixel and total die)
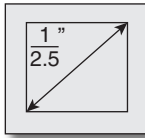
$\dfrac{1}{2.5}$ "   28 megapixels, 0.9 µm^2

Die size = 48 mm^2
Array size = 24 mm^2

$\dfrac{1}{3}$ "   12 megapixels, 1.1 µm^2

Die size = 36 mm^2
Array size = 15 mm^2

$\dfrac{1}{4}$ "   5 megapixels, 1.4 µm^2

Die size = 26 mm^2
Array size = 10 mm^2

$\dfrac{1}{6}$ "

2 megapixels, 1.75 µm^2

*Die size = 16 mm^2
Array size = 6 mm^2

* Full SOC camera chip

7.8 Mobile image sensor array size, die size and resolution.

Table 7.1 Relation of pixel size and resolution to various optical formats

| Number of pixels (resolution) | Optical format | | | | | |
	$\frac{1}{2.5}$ "	$\frac{1}{3}$ "	¼ "	$\frac{1}{5}$ "	$\frac{1}{6}$ "	$\frac{1}{12}$ "
VGA						
1.3 Mpix				2.2 µm^2	1.75 µm^2	1.4 µm^2
2 Mpix			2.2 µm^2	1.75 µm^2	1.4 µm^2	1.1 µm^2
3 Mpix		2.2 µm^2	1.75 µm^2	1.4 µm^2	1.1 µm^2	0.9 µm^2
5 Mpix	2.2 µm^2	1.75 µm^2	1.4 µm^2	1.1 µm^2	0.9 µm^2	
8 Mpix	1.75 µm^2	1.4 µm^2	1.1 µm^2	0.9 µm^2		
12 Mpix	1.4 µm^2	1.1 µm^2	0.9 µm^2			
18 Mpix	1.1 µm^2	0.9 µm^2				
28 Mpix	0.9 µm^2					
		Pixel size (µm^2)				

7.9 Preserving SNR performance with an arbitrary area as pixels shrink.

of the active pixel sensor (APS) technology by Lee *et al.* (1995), combining a pinned photodiode and a transfer gate achieving non-destructive readout and CDS leading to sufficient signal-to-noise ratios (SNRs) for good performance image capture with underlying CMOS process technology. The APS sensor permitted charge transfer to an in-pixel amplifier using correlated double sampling, with integrated circuits for fixed-pattern noise removal (Guidash *et al.*, 1997 and Inoue *et al.*, 2003).

The highly competitive smartphone camera market demands 'state-of-the-art' technology camera performance from each product model, driving sensors to an ever increasing resolution, which in term has driven the pixel size to shrink. The CMOS image sensor industry has been in a resolution race, where within a given diagonal optical format, and especially at ⅓″ and ¼″ optical formats, sensor manufacturers must produce sensors that achieve a 'scaled' performance. Figure 7.9 shows the concept of 'scaled' performance. As the area of the pixel decreases, the fundamental challenge centers on maintaining a 'scaled' sensitivity, such that for any given area of pixels, say $\frac{1}{10}$th of an inch square, the sensor always achieves the same SNR. Figure 7.10 shows an example of the evolution of pixel SNR-10 values over time. Essentially the largest pixel 5.6 μm was developed first, then the next smallest, and so on. However, the improvements from the development of the smaller pixel are usually re-applied to the larger pixels to create subsequent generations of each pixel, over time. Figure 7.11 shows that for a particular optical format (⅓″), the SNR-10 improves over time as solutions emerge to the challenge of scaled performance with ever smaller pixels.

The industry challenge remains achieving the next pixel reduction in a timely manner. Historically, pixel advances have been achieved every 18 months, typical of Moore's Law of CMOS transistor density doubling every 18 months. This Moore's Law of CMOS imaging created a well predictable industry, in terms of performance and product delivery. However, the trend has slowed recently as the limitation of CMOS and optics increase the

Normalized scene light level for luminance SNR-10
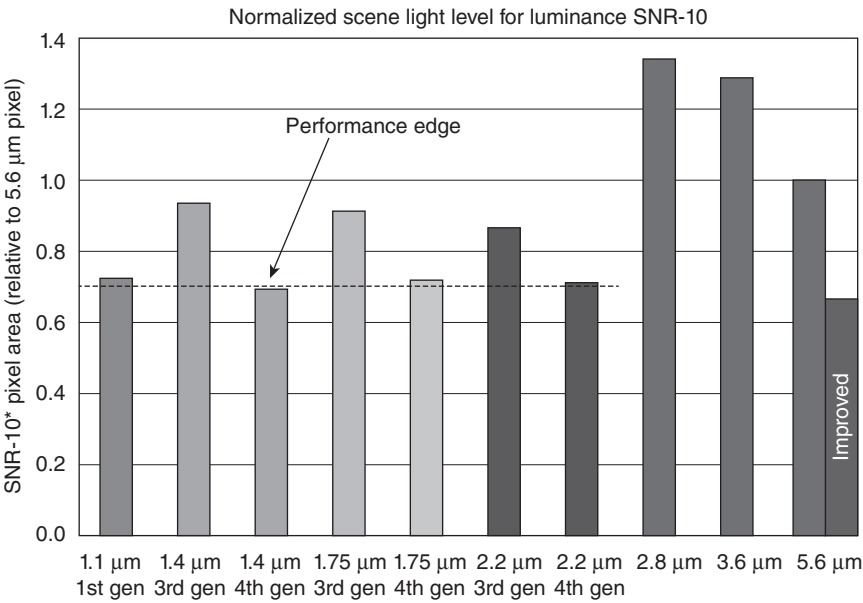
7.10 Example scaled pixel SNR performance per pixel generation.
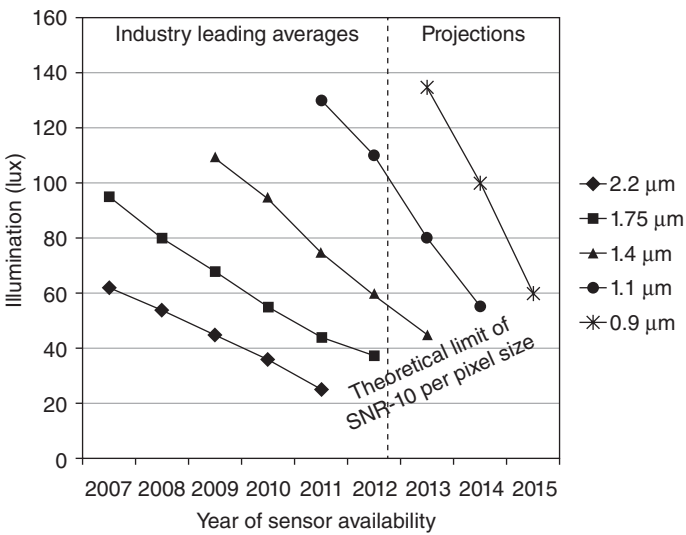
7.11 Industry SNR-10 performance trends for a ⅓″ sensor.

challenges. 'More-than-Moore' technologies have been developed to overcome some of these challenges (Arden *et al.*, 2010). Further camera compaction and performance can result from technologies like 3D silicon stacking structures, including through-silicon-via (TSV) technology. As well, multi-array architecture approaches can extend the camera solutions beyond the limitations of Moore's Law.

Well capacity, determined by the photodiode size and reduced as pixel area reduces in area, represents another parameter paramount to achieving a high performance pixel; the photodiode's well capacity will limit the number of photon re-combinations possible. At least 4,000 electron capacity is needed for mobile. Mobile sensor designers strive to achieve between 6,000 to 10,000 electron capacity. The sensor designer must adjust the sensor's conversion gain, or the output voltage per electron in the photodiode, to accommodate the well capacity or size of the photodiode and the gain of the readout circuitry and ADC.

In 2004 several researchers published a fundamental method of sustaining scaled performance with pixel or transistor sharing (Takahashi *et al.*, 2004; Mori *et al.*, 2004; Mabuchi *et al.*, 2004). Essentially as the pixel becomes smaller, the pixel's transistors tend to become a larger percentage of the overall pixel in area, thereby decreasing the relative size of the photodiode. By sharing of the transistors between neighboring pixels, the problem was overcome. This sharing reduced the number of transistors per pixel from four with a 4T pixel, to as low as 1.5 transistors per pixel.

Shading or non-uniformity of the response across the sensor and crosstalk represent two significant areas historically limiting CMOS sensor's performance as pixels and optics shrank. Both shading and pixel crosstalk can result from difficulties steering the light into the ever smaller pixels. If the individual photons land in the wrong pixel, both problems can occur. As the camera designers push for smaller lenses, the sensor must accept a larger and larger chief ray angle, as the light bends to a higher degree (typically 28 degrees). Attention to the symmetry of the pixel and the tuning of the optical pathway either by use of special materials or optical light guides into the photodiode results in proper performance as pixels and optics shrink. This detailed pixel engineering continues as the new approaches evolve to shrink the pixel further.

7.2.4   The CMOS sensor for mobile devices

Two primary functions of the CMOS sensor are readout of the pixel array and conversion of the analog pixel signals to an output digital data stream (sometimes compressed). The image sensor has limitations of speed, power, and size. Typical CMOS sensors for mobile consume between 500 and 900 milliwatts. The choice of ADC architecture and data rate (which increases

for higher resolution), critically determines the resultant sensor power dissipation (Panicacci *et al.*, 1996). Choices include single-slope ADC (Sugiki *et al.*, 2000 and Nitta *et al.*, 2006), successive approximation ADC (Takayanagi *et al.*, 2005; Zhou *et al.*, 1997), and cyclic ADC (Kawahito *et al.*, 2008). The key parameters are size, power and both fixed and random noise reduction (Matsuo *et al.*, 2008). A benchmark for state-of-the-art mobile sensors is 1 milliwatt per megapixel per fps (frames-per-second).

Mobile sensor ADCs are typically designed with 12 to 14 bits of dynamic range per pixel, preserving at least 10 bits of finite precision at the output of the sensor. Today's 8 Mpixel sensors operating at 30 frames per second (in some cases at 1080p video and in others at native resolution) need several hundred megapixels per second data rates, depending if the sensor outputs Bayer scaled video or YUV via an SOC sensor. The sensor's serial digital interface speed then becomes a limitation, especially as resolutions and frame rates increase in the future. Aptina's High Speed Serial Pixel Interface (HiSPi™) can realize speeds up to 2.8 Gbits/sec per lane. The MIPI™ alliance serial interface can realize up to 1Gbits/sec per lane, with 1.5 Gbits/sec in development. In practice, a mobile sensor may use multiple lanes for multiples of those data rates. However, as sensors increase from 12 to 22 megapixels, and at 30 fps or even 60 fps, it becomes clear serial data rates and number of lanes will increase dramatically in the future.

Many mobile sensors contain an image processing function that performs automatic camera functions like white balance and color processing. This type of sensor is called a SOC and will be discussed in detail in Section 7.3. Figure 7.12 shows a typical CMOS mobile SOC sensor, highlighting the layout of a sensor, with the pixel array, ADCs, image processing logic gates, and memories shown in physical relation to one another.

A critical function of the sensor electronics includes reducing random and fixed pattern noise. Both noise types have limited performance for mobile phones. Typically fixed pattern noise should be reduced to 16 db under random noise (Nobukazu, 2012). Since the sensor reads out sequentially, a typical sensor includes dark reference pixels vertically where no light strikes the pixel, creating a column dark reference per row. This dark reference is used to clamp the dark level, for use by the differential readout circuits, thereby reducing any impact of drifting of the dark reference during the readout process. As well, variations in ground across the sensor due to insufficient metal lines can create troublesome shading across the sensor as a result of potential increases in the dark reference.

7.2.5   The CMOS sensor image computing function

In some sensors the image computing function 'lightly' processes the sensor's digital image with image processing for defect removal or noise
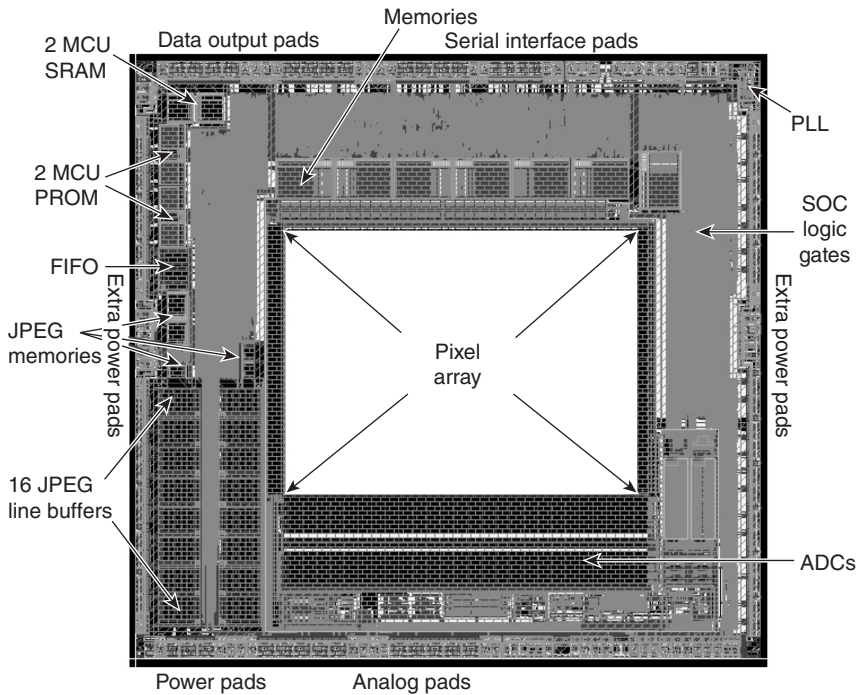
7.12 Die layer geometry showing physical layout for a typical CMOS image sensor SOC.

filtering. In other cases, a full camera image signal processing (ISP) essentially performs the image processing functions of a digital still and video camera, which we call the SOC. However, complexity increases dramatically when implementing a secondary digital processing chip, essentially a two-chip approach, greatly increasing the size of the logic and memory in close proximity to the sensor. Digital image processing performs many tasks, but the fundamental color processing of the sub-sampled Bayer pattern color matrix critically determines image quality in dynamic lighting situations. For example, sunlight, incandescent, fluorescence or others cause large shifts in color performance. The color processing includes color correction for illumination sources and the on-chip color filter array, coupled with interpolation for the spatially sampled colors.

Ultimately, a camera's performance is measured in ISO increments. At nominal illumination and camera sensitivity an ISO of 100 should perform well at a $\frac{1}{30}$th second exposure interval. A camera using ISO 400 at the same $\frac{1}{30}$th of a second exposure time, would perform better, with much less visual noise. Larger, more sensitive pixels, such as those with a $\frac{1}{2.5}''$ sensor would create a higher ISO camera than a $\frac{1}{3}''$ sensor.

7.2.6 Optics and packaging for the CMOS sensor-based camera

For widespread usage of image sensors in camera phones to occur, challenges in manufacturing costs, size, solder re-flow, and particulate contamination in the optical path were each overcome. Traditional semiconductor packaging methods fell short of the demands. Initially, camera modules were devised with a chip-on-board (COB) method where singulated sensor die were bonded to a flex-cable. In order to enable high volume automatic manufacturing methods, chip-scale-packaging (CSP) was developed (Bartek *et al.*, 2004), especially featuring the yield-enhancing benefit of keeping the image sensor die encapsulated as the device was shipped to locations across the world for low-cost manufacturing. However, CSP requires about a five percent addition to the die area to enable connection to the backside of the die for packaging. Another challenge relates to the cost and yield impact of large die. Since CSP is a wafer construction method, the larger the die, the more cost becomes applied to each die. Beyond ¼″, the cost can become prohibitive.

Optics advancements were in the area of reducing cost and size, without significant performance degradation. Introduction of reduced element and/ or plastic lens enabled much of the reduction. High precision machining of optical molding and advancement in plastic molding technology has allowed lens performance to be maintained while reducing lens size and costs.

The mobile camera module developed simultaneously with the sensor silicon. In addition to the image sensor, the camera module contains the lens assembly with two or more lenses, an auto-focus voice coil, and other passive components. The parameters of minimizing cost, increasing yield was challenging with particulate matter frequently sticking to the sensor array and optically obscuring the pixels. Some sensors are susceptible to electromagnetic interference (EMI) and power spikes, causing reset of the sensors. These system level issues have been resolved.

7.3   Emerging complementary metal-oxide-semiconductor (CMOS) 'sensor-embedded' technologies

As pixel arrays tend to commoditize so that each image sensor vendor achieves similar performance and cost, sensor-embedded technologies can be devised to offer huge performance differentiated advantages at the product level. Essentially embedding new technology into the sensor solution without increasing cost, can offer value to the camera developers. While the industry has developed many camera applications like high-dynamic range in isolation from the sensor, such as with the mobile phone's

Quality images in all lighting conditions	Fast, responsive camera operation	Small devices to capture everywhere	Seamless video capture	Power reduction
• Low light sensitivity enhancements • HDR capabilities • Advanced image processing	• Context mode switching • Fast AF algorithms • Reduced startup latency • User presence and tracking	• Pixel shrink for smaller optical formats • Wafer level module technologies • High CRA tolerance	• HD and UDTV video modes • Image stabilization • Fast interface	• Design and process • "Always-on" • User detection • Power down modes • Automatic light sensing (ALS)

7.13 Key technologies to achieve superior image and video capture experiences.

application processor, we can improve many applications with specific technologies 'embedded' directly in the CMOS sensor. Total power, speed of operation, and image quality can all be enhanced with such an integrated solution. This section discusses such 'sensor-embedded' technologies and how the limitations of the technologies have or are being overcome with advancements.

Fundamentally we seek to improve the image and video capture experience for camera phones. Figure 7.13 illustrates the key technologies developed or in development to create superior image and video capture experiences in camera phones. In addition, Fig. 7.14 shows important sensor-embedded technologies that we will discuss in this section. High-dynamic range (HDR), global shutter, and 3D depth capture technologies are embedded in the sensor in the form of special pixels or advanced processing logic. Smart sensor cameras essentially embed advanced image computing functions like metadata calculation, object detection with tracking, and compression into the sensor using silicon chip stacking technologies.

7.3.1 Limitations of HDR image capture

HDR for mobile imaging offers enhanced picture taking in the form of simultaneous sensing of high brightness and low brightness scenes. Camera

High dynamic range (HDR)	Global shutter	3D and depth	Smart sensor cameras

- **High dynamic range (HDR)**
  - Elegantly capturing those difficult high contrast scenes

- **Global shutter**
  - Eliminating rolling shutter motion artifacts

- **3D and depth**
  - Viewing and interacting in the third dimension
  - Object recognition
  - Structured lighting for depth ranging

- **Smart sensor cameras**
  - Co-processor
  - Sensor
  - Sensor with co-processor
  - Object detection and tracking, etc.
  - Noise reduction, compression, etc.

7.14 Sensor-embedded technologies.

phones provide automatic exposure capabilities, adapting to the scene illumination by typically changing the cameras exposure time from tens of milliseconds to hundreds of milliseconds as a function of the scene brightness, and then reading out the sensor to a limited-dynamic range linear display medium (such as LCD displays). Essentially, certain scenes or portions of scenes are then easily visualized, yet others can appear overexposed or underexposed to the point of loss of contrast details in the displayed image. Cameras with HDR capability capture sufficient scene intensity variation to visualize subtle variations in high brightness, mid-level brightness and very low brightness, all equally. In practice with a mobile phone, the resultant image wanders automatically from overexposed to underexposed depending on the scene. For the scene captured indoors, yet with the subject standing in front of a window on a bright sunny day the required range of dark to light variation in a single scene exceeds the capability of typical cameras.

The most common method used by photographers for decades to achieve higher dynamic range images requires taking multiple pictures, each with different exposures. This method of capturing of multiple exposures of the scene in sequence (called exposure bracketing), with at least one long exposure and one short exposure, followed by a combination or overlaying of the resultant images, has promising results. However, alignment during overlaying of the multiple exposed images, and subject motion artifacts (where objects are in different locations in each of those scene due to motion) minimize the effectiveness of that approach. Many camera phones using this multi-shot method leverage the memory and image processors in

the phone to minimize the motion artifacts by minimizing those exposure differences. However, this in turn minimizes the benefit as well. In many cases, only a 10 percent gain in dynamic range results, which leads to infrequent use of this capture mode.

Sensor-based HDR: in-pixel method

The elegant solution to HDR imaging is to devise image sensors that can capture higher dynamic range. Typical mobile image sensors only achieve 12 to 14-bits of dynamic range, limited by the storage capacity of the image sensor's pixels and the noise in the sensor. A typical scene with sunlight and room light in different portions of the scene needs more than 20 bits of dynamic range to portray a picture with good detail in each portion. Solutions have been created by using special pixels in the sensor, like 'in-pixel storage' or 'lateral overflow' approaches and a pulsed transfer gate method (Yasutomi *et al.*, 2009), yet with limited success due to the cost/size inefficiency. The lateral overflow methods can restrict pixel size and sensitivity, or resolution per unit area. The size/cost competitiveness of mobile sensors minimizes its use. Other interesting new solutions include the multi-array image sensors or cameras where each array has independent exposure control. With this method, some of the multi-shot problems are relieved since exposures can be 'time-aligned' to the same microsecond of the scene. The disadvantage is that resolution per unit sensor area (or cost) is sacrificed to compose multiple arrays.

Sensor-based HDR: in-sensor method

A practical in-sensor method features a standard pixel array and the use of different exposure controls for different portions of the sensor, with the combination of each output to sample very high dynamic ranges. The ultimate HDR solution for camera phones centers around an 'always-on' HDR mode, where both still and video sequences are effectively corrected or enhanced in any situation, yet without (or with little) compromise to cost, resolution or frame rate. To this goal, the industry has recently introduced HDR sensors using specialized line-based architectures using variable exposure times, coupled with on-sensor 'fusing' of the information, and sometimes including the tone-mapping function, all operating at real-time video rates. These methods include:

- sensors with multiple-lines outputs each with different exposure controls and line storage, essentially operating multiple exposures simultaneously (a 3-line sensor has been implemented by Aptina to achieve 120 db HDR), and

- sensors with a simplification of that structure whereby each alternate line exposure varies in a line-interlaced exposure control architecture (called Aptina MobileHDR™).

These line-based approaches appear to be a good alternative for sensors in mobile camera phones where a combination of high resolution, high sensitivity, and low cost (small silicon size) constrain the solution. Many mobile image sensors designed today include this HDR feature, combining the long and short integration of alternate lines of video with simultaneous readout possible to create 80 to 100 db HDR. A slight compromise in vertical resolution occurs, yet for video image sensors at 8 megapixels are well beyond the display resolution of video today (1080p). This interlaced line architecture balances motion artifacts and the magnitude of the exposure difference.

Applications which demand very large dynamic range can be addressed with sensors using more and larger variation in exposure control. For example, Aptina's three-row multi-exposure sensor offers extended dynamic ranges with two knee points to approximate an exponential capture transfer curve. The sensor can be designed with any combination of exposure ranges by spatially separating the row readouts then combining the three rows. In use, each row then has a different exposure time. In effect an image can then be created with up to 40 bits of exposure dynamic range. While this may be extreme for mobile applications, automotive applications demand this for mission-critical, accurate sensing when lane-following from bright sunlight roads through dark tunnels.

The final step in many HDR capture applications includes tone mapping of the captured image to a unique color and contrast mapping. The resultant images range from pictures with better clarity to extreme effects showing dramatic enhancements of detail. The use of local tone-mapping or histogram equalization resolves these effects. As well, for mobile sensors the complete HDR solution with tone mapping can fit within an image sensor's digital logic section. Many examples showing the visual effects of tone mapping can be found on Internet-based photo sharing services, such as Flickr.com and Instagram.com, for applications like real estate listing.

7.3.2 Limitations of rolling shutter and global shutter image capture methods

For mobile applications, in order to reduce the camera cost, size and easily support fast video capture, camera designers do not use mechanical aperture shutters found in high-end digital cameras. As well, most CMOS sensors used in mobile products only support a rolling shutter readout method (or in some cases a global reset as well). In operation, the sensor integrates over

a frame capture time, and simultaneously reads out the sensor. This allows continuous readout and limited data buffering in the system, for a lower-cost, smaller solution. Unfortunately the rolling shutter method will create image geometric distortion artifacts with fast motion of the camera's gaze, such as when the user shakes or moves the camera quickly. While most situations enable the capture of good images or video, unfortunately rolling shutter artifacts are seen when capturing scenes with substantial motion of the camera or within the scene. Taking a video from a bullet train can show stretched or tilted pictures as a function of the readout and speed of the train. Rapid and random motion of a camera, such as when riding a car through Mexico's Baja desert at high speeds creates an artifact called Jell-O™, where the video looks as if one's looking through a clear wobbling solid. This distortion becomes magnified as the function of the rate of motion or length of capture time increases. For example, refer to the sampling error created as one observes a propeller blade in motion in Fig. 7.15.

This rolling shutter artifact is a result of the simultaneous integration and readout inherent in CMOS x-y addressable sensors, where readout is overlapped with integration line-by-line to create a continuous integration situation. If not overlapped, the sensor would readout for a portion of the frame time and integrate for only a portion of the frame time, resulting in a reduction of the sensitivity. For that usage mode to be effective, a mechanical shutter would be necessary in the optical path to freeze each pixel's contents as readout progresses sequentially. Methods have been devised to detect and correct such artifacts, such as cleaning YouTube™

7.15 Captured photograph using CMOS rolling shutter sensor showing interesting distortion of actual image captured of a rotating aircraft propeller.

videos taken with jerky and random motion of the handheld camera (Schuler *et al.*, 2012). However, creating a 'clean' capture or video in all situations is desirable within the camera.

CMOS sensors using global shutter (GS) methods, where each pixel captures and stores the illumination signal in a pixel before starting the sequential line-by-line readout, solve many of the image distortion issues with a rolling shutter sensor. These global shutter sensors are typically used in specialized applications like machine vision or scanning where the image must be captured with spatial accuracy. With CCDs, frame or inter-line transfer methods solve the problem, but at a severe area cost (between 50 percent and 100 percent of the integrated sensing area).

The CMOS global shutter requires at least another transistor per pixel and/or storage area for temporary storage of the pixel values while readout progresses, which increases sensor size and cost. Typically for a given sensor area, global shutter reduces the sensitivity by about 6–10 percent as the required area increases with an additional transistor in each pixel. As the size of the pixel shrinks, this percent increases. Another approach to overcome the need for global shutter pixels could include using a faster and more sensitive sensor. Mobile CMOS image sensors began at 10 to 15 fps speeds, with 30 and up to 60 fps today, and with 120 fps predicted in the future. As the readout time is minimized and sensitivity is increased the distorting effects of overlapped readout and integration is minimized. This trend will continue over global shutter until stacking emerges to enable more transistors per pixel.

7.3.3 Limitations of 3D and depth capture

Capture of depth information represents one of the largest growth opportunities for mobile imaging. Determination of the location and gestures of the user for the user-facing camera, and the distance to and between objects for the scene-facing camera will essentially provide Microsoft X-Box^R Kinect^{TM}-like capabilities in a mobile product form. The depth map capture can be accomplished with several methods, as shown in Table 7.2. However, for mobile the constraints of size and power will limit the practical choices for a solution. At a minimum, a stereo-based parallax capturing solution should include modes in the sensor to synchronize the readout of the sensors.

7.4   Mobile image sensor architecture and product considerations

By understanding the advantages and disadvantages of each sensor architecture, the camera designers or users can optimize the selection,

Table 7.2 Depth capture methods

Approach	Method	Limitations
Stereo parallax	Measure positional shift between images from two image sensors; usually in the visible spectrum	Separation of camera limits distance of depth measurement; requires two sensors
Structured light	Multi-frame pattern recognition of a projected structured illumination with an image sensor, usually in the infrared spectrum	Processing complexity; resolution; accuracy; computational complexity; occlusion
Depth of focus	Capture multiple images at narrow field of view with varying focus, then extract depth	Accuracy; motion artifacts
Time-of-flight	High speed pulse with special pixels measuring the return time of the reflected light; processing of the delay as a function of depth per pixel	Power necessary to generate the pulses; sensitivity of pixels required to capture 'pulsed' light drives either large, costly die size or low resolution

specification, and application of the sensor. Parameters that characterize a CMOS image sensor architecture for mobile imaging include:

- the optical format which largely determines size and cost
- pixel and readout methods
- high-speed data interfaces
- analog-to-digital conversion methods
- electrical parameters
- optics partitioning
- z-height
- embedded vs off-chip ISP
- data compression, and
- meta-data creation.

Meta-data creation reflects a trend in image sensors whereby the sensor computes information within the sensor for output, rather than the conventional use of a camera to simply capture pictures or video for output. Examples include light detection (such as automatic light sensing – ALS), user presence detect, hand gesture detection, tracking finger positioning or others. This section reviews the mobile sensor architectures used in the industry.

7.4.1 Rear-facing (RF) and front-facing (FF) sensors

While most mobile phones feature one camera, smartphones typical use two sensors, a RF camera which faces the distant subject or scene when the user faces the device's display, and a FF camera which faces the user, situated at arms length. We expect even more cameras per smartphone in the future as well, but usually they will fit into these categories (RF and FF). The requirements of each camera have diverged, since one fundamentally sees the user in close proximity (with a narrow focal range) and the other sees the scene ahead and in the distance (with a wider focal range). The rear-facing camera has an image and video performance mandate driving high resolution and larger lenses, while the front-facing has the challenge of even smaller size since it is secondary, yet with image and video performance limited to enable user identification and tracking or video conferencing.

7.4.2 RF camera for mobile products

From the basic sensing technology perspective the following parameters are typical of a good quality RF camera-phone today:

- spatial resolution of 8 to 13 megapixels
- 100 lux illumination level at signal-to-noise ratio of 10 (Alakarhu, 2007)
- quantum efficiency approaching 90 percent
- pixel full well capacity of 6,000 electrons, and
- operation at 15 to 30 frames-per-second.

These performance requirements are for a ⅓″ optical format camera module that is typically 8.5 by 8.5 millimeters in area and 5.5 millimeters in height. Digital still camera (DSC) level performance is attained in some products using larger optical formats with $\frac{1}{2.5}$″ or $\frac{1}{1.8}$″ sensors and optics. At the extreme of resolution and size in camera phones, recently Nokia Corporation introduced a 41 megapixel camera phone product, called the Pure View 808™, where the camera uses oversampling algorithms to combine pixels from a large $\frac{1}{1.2}$″ optical format sensor to form a very high quality 5 or 8 megapixel (7,728 × 5,368) image. Another benefit of this approach is the lossless digital zoom that results for video or cases where the target resolution is less than the sensor. In addition, the artifacts and distortions of optical zoom are eliminated. Finally the oversampling approach helps correct errors created in the Bayer color sampling (Alakarhu *et al.*, 2012).

7.4.3 FF camera for mobile products

Mobile phone makers restrict the FF camera in size and cost, with image quality performance traditionally a distant concern. Over time, the FF

image sensor will evolve to higher performance, with less noise and better low light sensitivity due to the increased adoption of new applications including video conferencing, such as with Microsoft™ Skype™ technology. Applications to track the user's eyes and remove background from the scene lead to more efficient video compression and lower latency with today's poor network bandwidth and Internet connections. As suggested, applications like ALS, user presence detection and gesture detection, will utilize the FF camera for many mobile products.

In 2005, wafer-level camera module (WLM) technology was developed to facilitate the FF camera application, partly because of its ultra-small size (less than 2.5 mm height due to constraints in the phone), accepted performance compromise, and low-cost. WLM enabled simple assembly and integration into the mobile phone without concern for the high temperatures of soldering or rework, should that become necessary. WLM combines optics elements, passive components (capacitors), and the image sensor. In this case, the sensor architecture typically includes the ISP, creating a complete performance and cost optimized camera solution.

The original concept was to construct the WLM at the wafer level (hence the name), yet technology limitations resulted in higher cost, rather than lower cost (due to yield issues). However, technology to construct a wafer-level optics (WLO) array for subsequent bonding to the silicon sensor wafer was developed to reduce the cost of conventional camera modules. Certain small modules were constructed with pick-and-place wafer-level optics, yet challenges in alignment of each sensor and lens were limiting. To complete the wafer-level integrated solution, the traditional bond pads found on the top of the sensor were replaced with conventional CMOS through-silicon-via (TSV) technology for component interconnect from the bottom of the sensor, rather than the traditional top of the sensor (Han *et al.*, 2010).

Lenses for WLM are an area for innovation. The lenses are typically attached to the sensor package using a mechanical fit. For WLM, the traditional camera focusing-step is not present. Forgoing this degree-of-freedom for the focus function results in critical tolerances for the lens back focal length, package and assembly processes.

For the camera phone designer, the WLM camera is simple for the integrator at every level. As new applications emerge for the FF camera, image and video quality become more important, such as for video conferencing. However, the limitations of cost and yield for wafer-level methods typically restrict the area to small sizes ($\frac{1}{6}''$ to $\frac{1}{15}''$). As an example, the wafer-level TSV and optics layers adds hundreds of dollars of cost to each combined wafer. For example, a $1,000 CMOS sensor wafer could then double to $2,000. For a 200 mm wafer with 28,000 die candidates per mm^2, this translates into $0.57 for a 16 mm^2 $\frac{1}{6}''$ sensor, yet a much larger cost of $0.92 for a 26 mm^2 ¼$''$ sensor, ignoring typical yield decreases for the bigger die.

7.4.4 A camera-phone image sensor roadmap

Historically, the roadmap of image sensor products was a simple, easy to predict, diagram showing a constant area-scaled performance, generation after generation, as the pixels shrink and the resolution increases (Moore's Law of pixels). Each product generation, defined by the pixel size, such as 2.2 μm, 1.75 μm, 1.4 μm, was advancing on an 18-month per generation pace, per Moore's Law. Within a year or so virtually all products at each optical format of a particular pixel size were introduced into the market, as shown in Fig. 7.16.

Figure 7.16 shows each year's product portfolio enabled by new generations of pixels (2.2 μm, 1.75 μm, etc.), vertically aligned in the diagram. However, the range of products actually tails out over many months and sometimes a year or more for introduction. As well, to generate volume and scale in the fabs, many new and emerging applications utilize these high volume mobile products. For example Aptina's five megapixel, ⅓″ sensor, produced over seven years ago, still ships in high volume in non-mobile markets.

In recent generations of pixel advancements, both at 1.4 μm and 1.1 μm, the industry has delayed the generation cycle to employ more advanced

Sensor optical format	2.2 μm FSI	1.75 μm FSI	1.4 μm BSI	1.1 μm BSI	0.9 μm BSI
$\frac{1}{2.5}$ "	5MP	8MP	12MP	18MP	28MP
$\frac{1}{3}$ "	3MP	5MP	8MP	12MP	18MP
$\frac{1}{4}$ "	2MP	3MP	5MP	8MP	12MP
$\frac{1}{5}$ "	1.3MP	2MP	3MP	5MP	8MP
$\frac{1}{6}$ "		1.3MP	2MP	3MP	5MP
$\frac{1}{12}$ "	VGA		1.3MP	2MP	3MP
		VGA			2MP
			VGA		
	2005	2007	2010	2012	2014+

The small pixel resolution race!

Sensor delivery into the market
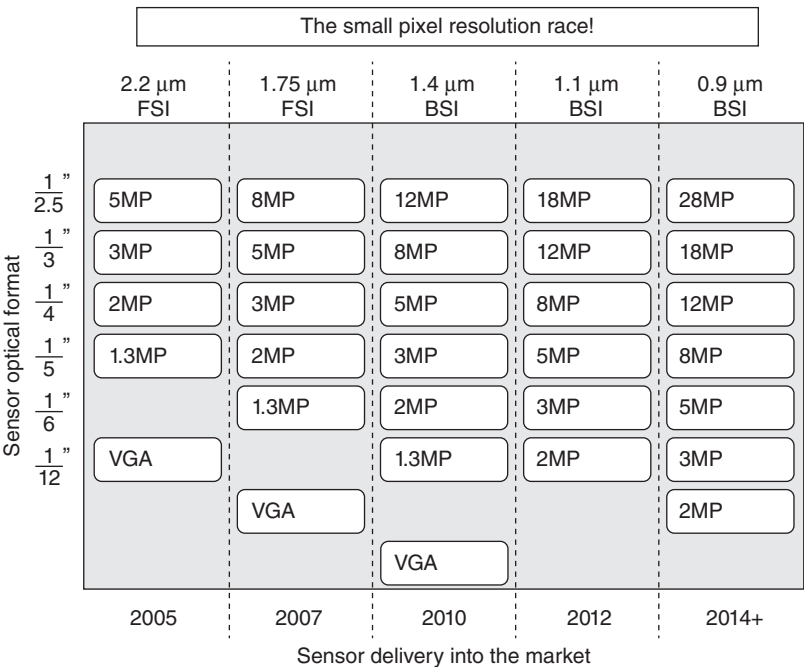
7.16 Mobile sensor product evolution and industry trend.
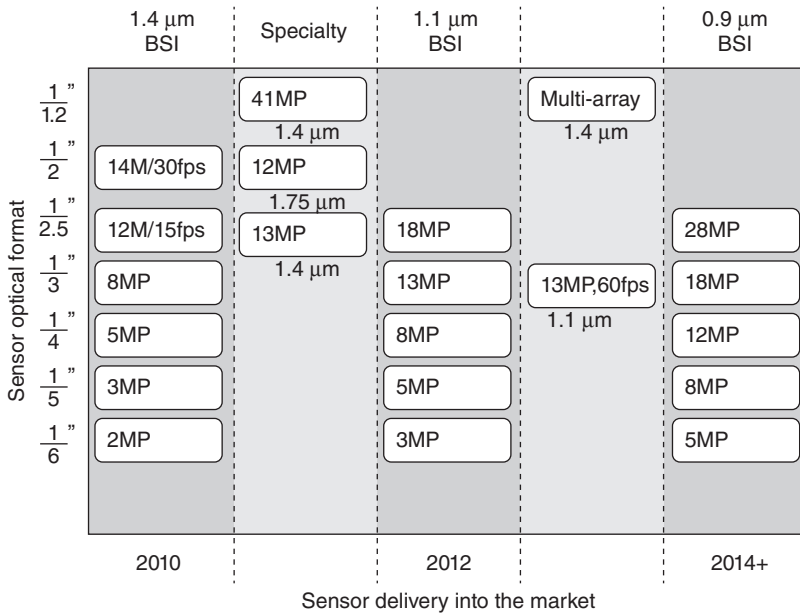
7.17 Mobile sensor specialty product differentiation grows, as pixel race slows.

semiconductor and optical technologies. This delay is shown in Figs. 7.16 and 7.17. Essentially delay results from the innovation and process investment needed to sustain 'scaled' performance beyond conventional CMOS technology. Figure 7.17 shows a few examples of specialty sensors that emerged or will emerge to sustain competitive differentiation between these base pixel node generations. Designers will leverage unique technologies like very large optical formats, higher speeds, and multi-array solutions. As well, sensor designers vary pixel size and optical formats from the values shown in Fig. 7.17 to create product differentiation. As the pixel race slows, sensor and camera differentiation grows, with the use of new methods and solutions.

7.4.5   The mobile image sensor pipeline architecture

The value of utilizing core CMOS technology to build image sensors extends from the economic value of drafting a huge industry of depreciated, low-cost fabs, with circuits and tools to enable high speed and high yield volume production, to the elegance of creating a complete camera on a chip, or system-on-a-chip (SOC) solution. Other approaches to sensing, like CCD sensors, require separate processing chips or chip sets, while CMOS permits full integration into one chip. However, with gate density only doubling

every 18 months, this approach eventually hits the limitations of either cost or the size of the die, which becomes constrained by the size of the camera module. The alternative solution today utilizes die stacking to increase the density of gates or memory within a camera module. Each mobile image sensor pipeline of processing contains:

- the pixel array
- analog processing
- digital processing and
- high-speed interface to the end-using system, as follows and shown in Fig. 7.18.

7.4.6 Analog processing for mobile applications

The image sensor analog functions include amplification of the pixel signal, timing and reset of the pixel readout, noise reduction, and ADC as shown in Fig. 7.18. The limitations of the technology are noise, power, die size and readout speed. By far, the noise in the form of pixel noise, row noise and column noise, both fixed and random (shot) have been key barriers to good images. Dark current and warm or hot pixels have limited the yield of sensors. Fabrication process improvements with respect to cleanliness of the fab equipment and order of process steps greatly influence this yield-impacting parameter.

7.4.7 Sensor-embedded image processing for mobile applications

The image processing pipeline, called the image signal processor, contains digital processing functions like digital filters for Bayer color correction processing (mapping to the eye's response), correction for lens and pixel distortions, elimination or correction of defective pixels, and interpolation of pixels to interpolate Bayer pixels to create full resolution outputs. The Type B SOC sensor architecture in Fig. 7.18 offers a fully tuned solution for the camera phone maker. The SOC became a popular solution in the mainstream and value segment of mobile. These image processing functions are typically implemented with pipelined and line-based methods with simplified algorithms to limit the complexity of the solution. Increases in resolution and frame rate become the primary limitation of these functions in practical sensors. For example, a 3 megapixel sensor operating at 15 fps may only contain 500 K logic gates, yet an 8 megapixel sensor operating at 30 fps may contain 4 million logic gates.

Moving beyond the basic ISP, the 2-chip Type C solution of Fig. 7.18 extends the range of possible functions to features like an advanced
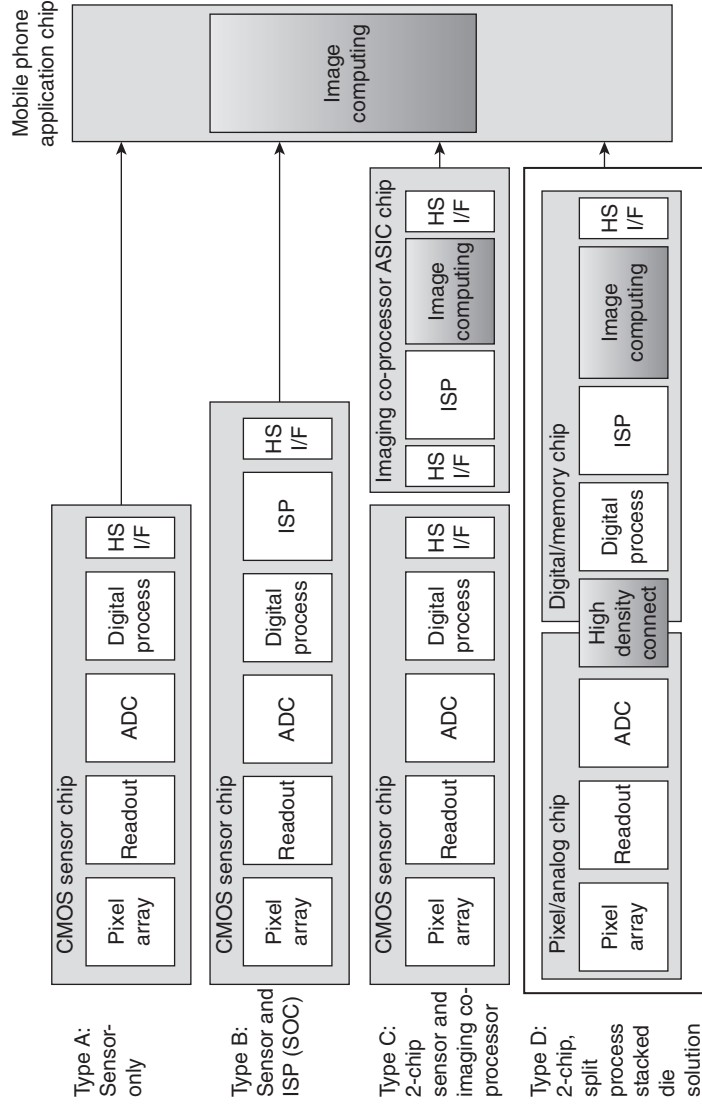
7.18 Four CMOS image sensor solution architectures.

Final 2-die stacked package

Imaging co-processor die (top side)

Through silicon via (TSV) and spacer layer

Imaging sensor with cavity (bottom side)

Redistribution layers and ball-grid array

Image sensor with co-processor (bottom side visible)
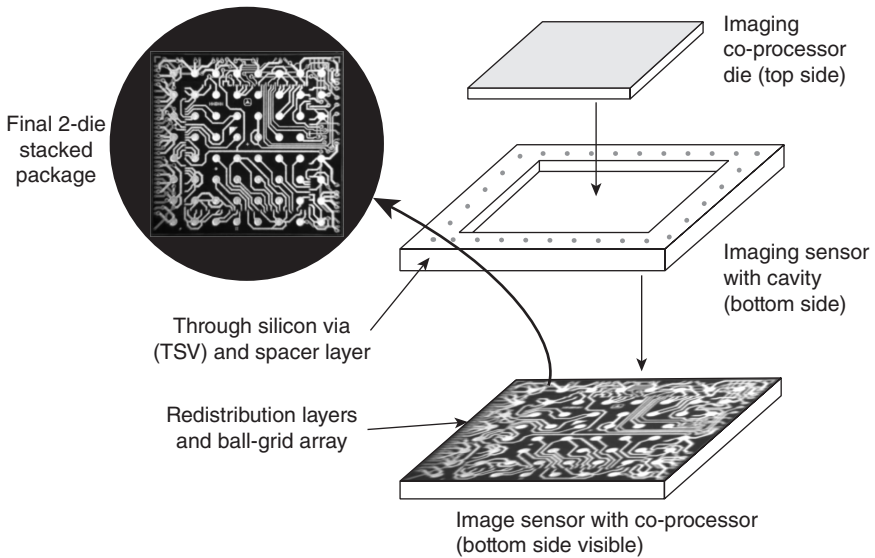
7.19 Die stacking enables sensor and co-processor integration, creating a smart camera.

high-quality ISP, HDR, face tracking, sensor fusion and computer vision algorithms, all of which can offer competitive advantage to the market. Aptina has developed an imaging co-processor that offers the computational power to perform these functions in a small application specific integrated circuit (ASIC) chip. The co-processor can be stacked with the CMOS sensor as shown in Fig. 7.19, or packaged separately. This creates a 'smart' camera, with the ability to compute metadata and image for information, rather than only producing great pictures and video. Type B and C solutions typically use a 'streaming or line-based' architecture, where we constrain algorithms for image processing or vision to operate on a limited set of lines of video, such as nine lines of the image at a time.

Finally, the Type D sensor will address further feature enhancement. The trend to higher resolution with higher frame rate will continue, especially for smartphones, somewhat outpacing the capability of the CMOS sensor to achieve even higher data rates, especially when compared to the multi-core giga floating point operation per second (GFLOPS) capability of application processors. As well, the capability of stacking together three separate layers, such as the sensor, processor and memory wafers/dies may emerge. Ultimately, with die/wafer stacking and power reduction, sensor-embedded image processing for camera-phones solution may transition to 'frame-based' solutions, supporting full-frame memory-based algorithms for image processing or computer vision.

7.4.8 Sensor high-speed interface for mobile applications

The electrical interface for image sensors in mobile applications began as simple parallel (8 or 10-bit) simple handshake interfaces from VGA to 2 megapixel sensors. As resolution and speed increased to 12 megapixels at up to 60 frames-per-second, several new interfaces were adopted to forego bond-pad area limitations. Those include high-speed serial interfaces like HiSpi™, MIPI, and SMIA™. Certainly resolution and speed will increase eventually up to 22 megapixels and 60 fps (and beyond), driving the need for new interface methods. In the future, small fiber-optic interconnects may be used.

7.5   Future trends

In this section we explore new areas of technology evolution critical for advancement of mobile products and applications. In some cases, these technologies can be used to solve the limitations of Moore's Law scaling challenges of imaging. Figure 7.20 shows particular categories of innovation where we expect camera phone core technology advancements.

7.5.1 Pixels: new exotic materials, new color patterns and new non-planar, 3D pixels

Future sensors will emerge to leverage new pixel architectures mostly to overcome the limitations of shrinking the pixels to the diffraction limit of light (about 0.7 μm). Methods in development include the use of pixels with

New pixels	Multi-array sensors	High density stacking	Computational camera
• Increasing sensitivity and reducing noise • Exotic surface materials • New color patterns • High-speed and stacking	• Multiple arrays and multiple cameras • Computational imaging • Super-resolution • 3D depth mapping	Sensor wafer Digital wafer Sensor die Digital die • Through-silicon-via interconnect • Rapid product differentitation	• Software programmable camera • Environmental information capture • Computational imaging engine

7.20 Emerging mobile imaging technologies.

Color matrix pattern

Spectral sensitivity

Bayer pattern

RGBW pattern

Clear (or white) pixel

Quantum efficiency vs wavelength

Quantum efficiency vs wavelength
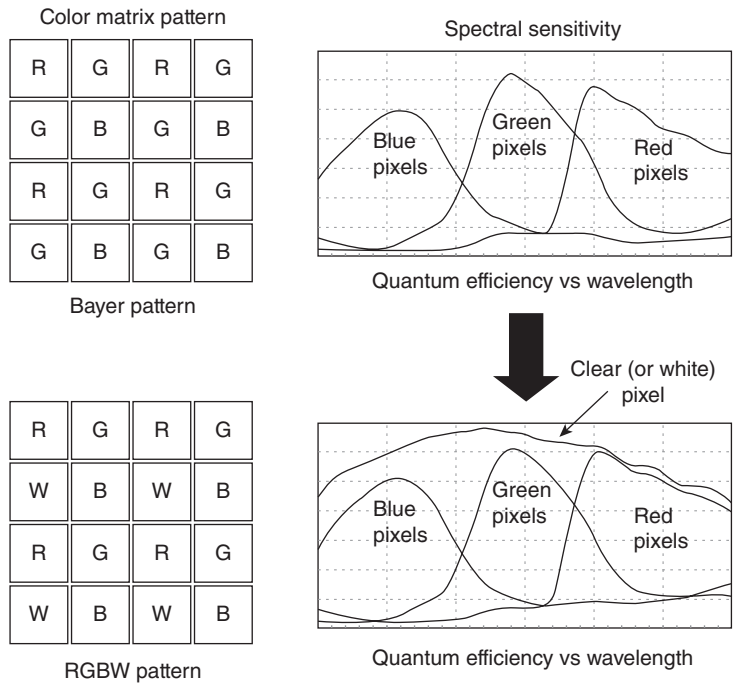
7.21 Higher sensitivity color matrix pattern.

exotic materials for high sensitivity pixels, such as organic films (Ihama *et al.*, 2010) or quantum films (Greenemeier, 2010). Many in the industry also are developing new color patterns for increased sensitivity, by introducing more white light into the sensor with full white (W) pixels in addition to the RGB pixels (Sony, 2010), as shown in Fig. 7.21. Note that a typical Bayer RGBG pattern diminishes luminance by over 50 percent. Finally, pixels which leverage vertical structures within the pixels (also called 3D structure pixels), versus today's planar pixels, are in development. Unfortunately at this time most suffer from the use of more complex and potentially more costly process steps. The expectation is that these methods can help us attain Moore's Law scaled performance with ever smaller pixels.

7.5.2 New computational imaging multi-array sensors

We expect the camera phone industry to soon adopt a new image capture paradigm, called computational imaging (Cossairt, 2011). The conventional image capture paradigm centers on sampling a scene by mimicking the human visual system. In that case, light irradiates and reflects off objects, with the camera simultaneously capturing a pixel-sampled representation of the scene. The camera contains an optical system to focus the light on

that sampling array of pixels. With computational imaging, a new model emerges to capture 'scene information' with a variety of methods. Examples include a plenoptic system or multi-array sensor, each with unique optical systems. For the plenoptic system, a wide variety of depths of focus are presented to a pixel array simultaneously. For multi-array, designers either implement cameras as multiple cameras in a camera system, or a multi-array sensor with small optics array assemblies in compact form for mobile applications. In each case, the captured dataset becomes coded in some way, but in most cases, the conventional camera's color and intensity radiance array is not readily stored. Essentially in all cases of computational imaging, the final resultant radiance image that's presented to the viewer involves some sort of computation to select, measure, or extract the image or information. Without the computation, the data doesn't appear as a normal image. For a plenoptic system, with a many-to-one mapping of the scene to the pixel coordinate requires the final user to extract the final image by selecting one of the mappings. A benefit of the approach includes an increased energy captured by the sensor per pixel. However, in practice this can come at the cost of decreased resolution.

Utilizing multi-arrays and computational imaging, camera performance and features can extend substantially to include features such as re-focus, motion blur removal, multi-spectral measurements, and others. Table 7.3 highlights some of those advantages.

Figure 7.22 shows the concept of a multi-array sensor and its replacement of a conventional rear-facing, single-array camera in a mobile phone. A multi-array sensor, providing multiple views of a scene enables a wide-variety of performance and application enhancements with a highly integrated solution. Perhaps equally interesting, multi-array sensors can enable very low height camera modules, and thereby thin camera phones. Other benefits of multiple arrays sampling a scene are described by Wernersson (2012). These include taking pictures to look past objects such as taking a picture of a building as seen from the other side of a chain-linked fence that may be a meter from the camera. In this case, images from multiple apertures taken from different poses are combined at the pixel level to create a final picture without the obscuring fence.

Super-resolution becomes particularly important for a multi-array camera when compared to a single array with a similar number of pixels, or resolution. However, the difference in optics and the combination of pixels from multi-apertures can reduce the resolution. Super-resolution works to recover that original resolution. Many new methods will emerge to accomplish that goal, including hallucination (Sun and Hays, 2012).

As mobile phones embrace 3D capture technologies like those used in Microsoft's Kinect™ and Nintendo's DS products, each of which use multiple sensors to extract human gesture or track interesting points, like

Table 7.3 Multi-array sensor capabilities

Multi-array (MA) sensor feature options	Conventional sensor baseline
Low *Z* height – reduced physical height from shorter focal length lenses	No option to reduce *Z* height
Depth map – from parallax between arrays	Must use two camera modules
Better signal-to-noise ratio (SNR) – primarily via more silicon area, also milder color correction matrix (CCM)	Less silicon area, less cost than MA
High dynamic range (HDR) – some arrays with neutral attenuation, flicker/motion-free	Lower performance HDR, with artifacts
Depth of field control – via light field or by depth map-based blurring	No depth of field control
High-speed video – stagger arrays exposures times and interleave	Higher cost, higher speed
Multi-spectral capture – add near-infrared, or extra visible bands, etc.	Use multiple cameras; costly
Color fidelity – no cross-channel cross-talk due to Bayer device proximity	Less color performance
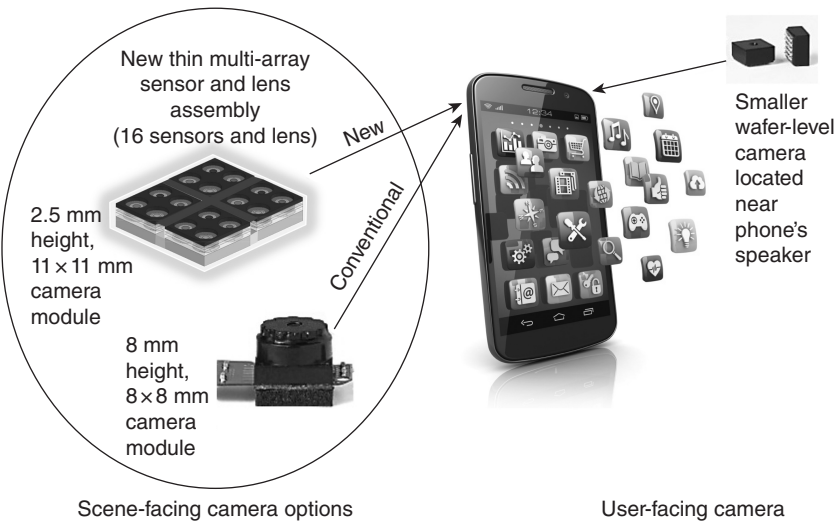
7.22 Multi-array sensor concept for mobile phones.

fingers, arms, legs, etc., we envision new smartphones emerging with similar yet smaller solutions. Either in the form of multiple sensors, or integrated multi-array solutions, we predict the smartphone will adopt this technology.

7.5.3 Software camera-phone and capturing the moment

Ultimately, as computational power evolves at the handset mobile phone device (doubling every 18 months by Moore's Law), coupled with virtually limitless computational power becoming accessible in the cloud for wireless Internet connected devices like mobile phones, the camera architecture will change to rely more on a 'heavy' computation camera model. This will essentially enable a 'software camera' where product value and differentiation comes from software and applications rather than solely the underlying hardware. The benefits of the explosive growth of applications results and are all similar to today's personal computer or tablet products.

This computational power enables new methods of imaging for mobile products. Today's methods employ simple single frame capture of pixel arrays followed by image processing to clean or enhance the pictures. Emerging methods will capture more 'environmental' information about the scene, or dynamic light-field cameras that effectively capture the important representations of light that can then be combined by a 'computation engine' to create a variety of images for the user. An early commercialization of a light-field camera is a plenoptic camera, introduced by Lytro[R]. Ultimately, computational imaging cameras will capture a moment of time, at numerous focal points, at varying exposure times and sample times which, depending on the scene, will allow the user to change focus, gamma, dynamic range and resolution of the resultant picture at any time in the future.

Rather than today's singular picture capture, computational imaging (Nayar, 2012) can utilize the capture of a 'moment' of scene information, recording radiance values or light-fields over some instance of time or variation in depth of field (Ng, 2006). This approach enables later 'reconstruction' of a preferred image, as well the removal of unwanted objects or information. One simple example includes an application created to remove uninteresting objects from a video clip using Scalado[R] ClearView[TM]. Essentially, objects that pass in front of a preferred façade can be easily removed by simply selecting each portion of the image from a frame that was not occluded. Dynamic capture would alter the frame capture for select portions of the scene, dictated by the need to sample at high rates where substantial motion is present in the scene to minimize blur. While application dependent, the complexity of the data sets with multi-spectral and high-dynamic-range values, and dynamic sampling of scene to create complex multi-sample radiance images when coupled with heavy

compute cycles can move us closer to the reality of high performance image capture by anyone, anywhere.

This 'computational camera' promises to create optimal cameras revealing good images and video at exceptionally low light, in virtually any situation, and with renditions of a scene free of artifacts. In that case, the limitations move more to the availability of data bandwidth in the form of either the time required to move the data to the cloud, or the latency of the experience (how long it takes to see the picture or video after capture, which could be minutes). So ultimately, the limitations of imaging will be related to the industry's ability to move and store tremendous quantities of data associated with the captured 'moments.'

7.5.4 High-density 3D stacked sensor

Leveraging 3D silicon die stacking technology from memory applications (Ku 2011), the industry is developing 3D image sensor stacking technology. In this case, the image sensor and co-processor dies are physically matched at certain points for interconnection (Sony Corporation, 2012). TSV technology enables a high-density interconnection of the dies or wafers.

We can also leverage BSI technology with wafer or die stacking. Perhaps an n-type metal-oxide-semiconductor (NMOS) BSI pixel array wafer can be thinned and bonded directly to a CMOS wafer with analog, logic, and memory. This hybrid solution can enable quick redesign of the digital section, generating rapid product differentiation (Sony Corporation, 2012).

7.5.5 New mobile imaging applications

Leveraging the computational model of the camera phone, countless applications will emerge like face recognition for camera control, imaging for gesture control, and augmented reality. As well, when the smartphone is enhanced with a physical plug-in camera module, such as at the top portion of the device, we create a new imaging functionality for the camera phone. In this case, the camera phone can provide a control device for a wide range of markets. Examples include:

- an opto-fluidic microscope and
- a smart remote for controlling TVs or other devices.

For the opto-fluidic application, a conventional image sensor can act as a contact scanner to create images such as in Fig. 7.23. In this case, a fluid like blood can be transported over the image sensor, creating a sequence of scanned images, without the use of a lens. Computational imaging methods are then used to 'compose' images by looking at the superposition of several instances as the fluid passes over the pixels. With this imaging
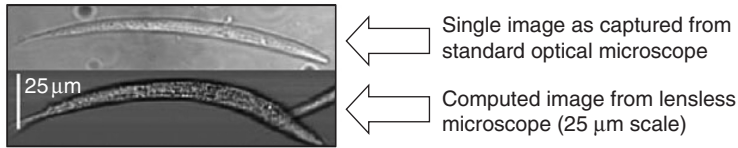
7.23 Computed image from sequence of images with an opto-fluidic microscope camera.

technology we can magnify particles about 60 times. As well, the plug-in module concept can support higher performance (larger) lens solutions and image or video capture at new perspectives. For example, an attached camera module could capture an image projected onto a wall by the mobile phone's integrated pico-projector. We in effect create a closed loop system with the camera capturing images projected by the mobile phone's pico-projector, then automatically aligning and correcting the displayed information as a function of the projection surface. For example, a mobile phone projecting pictures on a non-white surface or at a perspective angle (keystone effect), could be 'automatically corrected' with color correction processing and image warping correction inside the mobile phone.

7.6   Conclusion

Abundant innovation and rich technology advancement over a broad spectrum continues to provide strong growth of mobile imaging applications. We have seen key developments in the areas of semiconductor process, pixel design, sensor design, optics design, packaging design and system level design to enable this market.

Clearly, the camera phone represents one of the most technology-rich areas today, requiring the miniaturization and optimization of size, cost and performance over an entire system level, including analog electronics, mixed-signal transistor/pixel/memory-bit-cell design, digital electronics, memory, ultra high-speed interface, image processing, computer vision, micro-optics, color science, miniature opto-mechanics and semiconductor packaging. In this chapter, we explained some technology limitations for each over time. As the semiconductor methods mature, developments in new system areas like computational imaging and multi-array sensor can offer new methods of differentiation.

A decade ago, the image capture usage model for mobile phones began similar to conventional photography and videography, as a means to capture the moment with good quality pictures for sharing and memorabilia. For billions of new photographers the mobile phone was their first picture-taking experience, and quality was not critical. More recently with so many users having their mobile phones readily available for use, the requirements

for image sensors have changed dramatically. Phone users now demand better quality pictures and features like face detection for use with social networking. In smartphones, face detection has emerged to detect, identify and quickly move clips of one's friends over the Internet, but not necessarily become memorabilia, creating a social camera. In the future, the purpose and scope of mobile imaging will extend again as innovation continues to leverage technologies like computational imaging to further extract information from images rather than just pictures or video to enable new use cases. Gesture tracking and detecting the presence of the user are examples. Finally, camera phone developers will leverage increasing computational imaging capabilities to extend the performance of picture and video capture, establishing even higher levels of performance in the compact consumer mobile imaging device.

Mobile imaging growth not only occurs in market volume and core technologies, but also in breadth of applications used in new and interesting situations. More than any other imaging application, mobile imaging will enable ubiquitous usage of digital cameras by most individuals on the planet. The future will be remarkable with a combination of miniaturized mobile imaging technologies and supercomputer class computation power in a hand-sized consumer platform.

7.7   References

Alakarhu J, 2007, 'Image sensors and image quality in mobile phones', in *Proc. 2007 International Image Sensor Workshop*.

Alakarhu J, Dinning D, Salmelin E, 2012, 'Nokia pure view imaging technology'. Available from:http://i.nokia.com/blob/view/-/1486928/data/2/-/PureView-imaging-technology-whitepaper.pdf. (June 2012).

Arden W, Brillouet M, Cogez P, Graef M, Huizing B, Mahnkopf R, 2010, 'More-than-Moore'. Available from: http://www.itrs.net/Links/2010ITRS/IRC-ITRS-MtM-v2%203.pdf (June 2012).

Bartek M, Zilmer G, Teomin D, Polyakov A, Sinaga S M, Mendes P M, Burghartz J N, 2004, 'Wafer-level chip-scale packaging for low-end FR products', in *IEEE Dig. Tech. Papers from Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, 41–44.

Cossairt O, 'Tradeoffs and limits in computational imaging', PhD thesis, Columbia University, 2011.

Forward Concepts, 'Cellular handset and tablet core chip trends 2012', Report No. 1210, June, 2012. Available from: http://www.fwdconcepts.com/celltab/index.htm. (June 2012)

Franze K, Grosche J, Skatchikov S N, Schinkinger S, Foja C, Schild D, Uckermann O, Travis K, Reichenbach A, Guck J, 2012, 'Muller cells are living optical fibers in the vertegrate retina'. Available from: http://www.vision-research.eu/index.php?id=160 (June 2012).

Greenemeier L, 'Light improvements: could quantum dots boost the quality of cell phone pix?', *Scientific American*, 22 March, 2010.

Guidash R M, *et al.*, 1997, 'A 0.6 mm CMOS pinned photodiode color imager technology', *IEDM Tech. Dig.*, 927–929.

Han H, Kriman M, Boomgarden M, 2010, 'Wafer level camera technology – from wafer level packaging to wafer level integration', in *Proc. 2010 11th International Conference on Electronic Packaging Technology and High Density Packaging* (ICEPT-HDP), 16-19 August.

Henrysson A, 2007, 'Bringing augmented reality to mobile phones', PhD thesis No. 1145, Linkopings University.

Hynecek J, 1979, 'Virtual phase CCD technology', *IEDM Tech. Dig.*, December, 611–614.

Hynecek J, 1988, 'A new device architecture suitable for high resolution and high-performance image sensors', *IEEE Trans. Electron Devices*, 35(5), May, 646–652.

Ihama M, Mitsui T, Nomura K, Maehara Y, Inomata H, Gotou T, Takeuchi Y, 2010, 'Proposal of new organic CMOS imaging sensor for reduction in pixel size', published in Fujifilm Research and Development Report No. 55-2010, 14–17. Available from: http://www.fujifilm.com/about/research/report/055/pdf/index/ff_rd055_all_en.pdf. (June 2012).

Inoue I, Tanaka N, Yamashita H, Yamaguchi T, Ishikawa H, Ihara H, 2003, 'Low-leakage-current and low-operating-voltage buried photodiode for a CMOS imager', *IEEE Trans. Electron Devices*, 50(1), 43–47.

Iwabuchi S, Maruyama Y, Ohgishi Y, Muramatsu M, Karasawa N, Hirayama T, 2006, 'A back-illuminated high-sensitivity small-pixel color CMOS image sensor with flexible layout of metal wiring', *ISSCC Dig. Tech. Papers*, February, 302–303.

Kawahito S, Park J-H, Isobe K, Suhaidi S, Iida T, Mizota T, 2008, 'A CMOS image sensor integrating column-parallel cyclic ADCs with on-chip digital error-correction circuits', *ISSCC Dig. Tech. Papers*, February, 56–57.

Ku T K, 2011, '3D TSV stacking IC technologies, challenges and opportunities', October. Available from: http://sites.amd.com/us/Documents/TFE2011_026ITR.pdf (June 2012).

Lee P P K, Gee R C, Guidash R M, Lee T-H, Fossum E R, 1995, 'An active pixel sensor fabricated using CMOS/CCD process technology', *IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, April.

Mabuchi M, *et al.*, 2004, 'CMOS image sensor using a floating diffusion driving buried photodiode,' *ISSCC Dig. Tech. Papers*, February, 112–113.

Matsuo S, Bales T, Shoda M, Osawa S, Almond B, Mo Y, Gleason J, Chow T, Takayanagi I, 2008, 'A very low column FPN and row temporal noise 8.9 M-pixel 60 fps CMOS image sensor with 14 bit column parallel SA-ADC', *Symp. VLSI Circuits Dig. Tech. Papers*, June, 138–139.

Mori M, *et al.*, 2004, 'A ¼″ 2M pixel CMOS image sensor with 1.75-transistor/pixel', *ISSCC Dig. Tech. Papers*, February, 110–111.

Murakami M, Masuyama M, Tanaka S, Uchida M, Fujiwara K, Kojima M, Matsunaga Y, Mayumi S, 2005, '2.8 mm-pixel image sensor nMaicovicon™', *IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, 13–14.

Nayar S, 2012, 'Focal sweep photography', presented at the IEEE International Conference on Computational Photography, April.

Ng R, 2006, 'Digital light field photography', PhD thesis, Stanford University, July.

Nitta Y, Muramatsu Y, Amano K, Toyama T, Yamamoto J, Mishina K, Suzuki A, Taura T, Kato A, Kikuchi M, Yasui Y, Nomura H, Fukushima N, 2006, 'High-speed

digital double sampling with analog CDS on column parallel ADC architecture for low-noise active pixel sensor', *ISSCC Dig. Tech. Papers*, February, 500–502.

Nobukazu T, 2012, 'Dark current and white blemish', presented at Image Sensors 2012, 20 March.

Panicacci R, Pain B, Zhou Z, Nakamura J, Fossum E, 1996, 'Progress in voltage and current mode on-chip analog-to-digital converters for CMOS image sensors', *Proc. SPIE, Solid State Sensor Arrays and CCD Cameras*, *2654*, February, 63–71.

Schuler C, Hirsch M, Harmeling S, Scholkoph B, 2012, 'Fast removal of non-uniform camera shake and blind removal of optical abberations', presented at the IEEE International Conference on Computational Photography, April.

Sony Corporation, 2010, 'Sony's Stacked CMOS Image Sensor Solves All Existing Problems in One Stroke'. Available from: http://www.sony.net/Products/SC-HP/ cx_news/vol68/pdf/sideview_vol68.pdf. (June 2012).

Sugiki T, Ohsawa S, Hoshino M, Miura H, Sasaki M, Nakamura N, Tomizawa Y, Arakawa T, Inoie I, 2000, 'A 60 mW 10 b CMOS image sensor with column-to-column FPN reduction', *ISSCC Dig. Tech. Papers*, February, 108–109.

Sun L, Hays J, 2012, 'Super-resolution from Internet-scale scene matching', in *Proceedings of the IEEE International Conference on Computational Photography*, April.

Takahashi H, *et al.*, 2004, 'A 3.9 mm pixel pitch VGA format 10 b digital image sensor with 1.5-transistor/pixel', *ISSCC Dig. Tech. Papers*, February, 108–109.

Takayanagi I, Shirakawa M, Mitani K, Sugawara M, Iversen S, Moholt J, Nakamura J, Fossum E R, 2005, 'A 1.25-inch 60-frames/s 8.3-M-pixel digital-output CMOS image sensor,' *IEEE J. Solid-State Circuits*, 40(11), 2305–2314.

Tanaka N, Hashimoto S, Sugawa S, Morishita M, Matsumoto S, Nakamura Y, Ohmi T, 1989, 'A 310 k pixel bipolar imager (BASIS)', *ISSCC Dig. Tech. Papers*, February, 96–97.

Teranishi N, Kohno A, Ishihara Y, Oda E, Arai K, 1982, 'No image lag photodiode structure in the interline CCD image sensor', *IEDM Tech. Dig.*, December, 113–116.

Wakabayashi H, *et al.*, 2012, '10.3 Mpix back-side illuminated imager', *IEEE ISSCC Dig. Tech. Papers*, February.

Wernersson M, 2012, 'Next generation camera phones – trends and technologies', presented at Image Sensors 2012, 20 March.

White M H, Lampe D L, Bllaha F C, Mack L A, 1974, 'Characterization of surface channel CCD image arrays at low light levels', *IEEE J. Solid-State Circuits*, SC-9(1), 1–13.

Yang W, Kwon O-B, Lee J-I, Hwang G-T, Lee S-J, 1999, 'An integrated 800 × 600 CMOS imaging system', *ISSCC Dig. Tech. Papers*, February, 304–305.

Yasutomi K, Itoh S, Kawahito S, Tamura T, 2009, 'Two-stage charge transfer pixel using pinned diode for low-noise global shutter imaging', in *Proc. 2009 International Image Sensor Workshop*.

Zhou Z, Pain B, Fossum E R, 1997, 'CMOS active pixel sensor with on-chip successive approximation analog-to-digital converter', *IEEE Trans. Electron Devices*, 44, October, 1759–1763.

8

Complementary metal-oxide-semiconductor (CMOS) image sensors for automotive applications

C. DE LOCHT and H. VAN DEN BROECK, Melexis Technologies NV, Belgium

Abstract: This chapter reviews the application of cameras in cars. It covers vision systems and sensing systems from a safety point-of-view before outlining the individual requirements of image sensors for optimum results. The future development of this technology is also discussed.

Key words: vision applications, sensing applications, warning, recognition, detection, safety.

8.1   Automotive applications

Cameras in cars serve a wide variety of applications. Two groups can be distinguished:

1. cameras that are used to show an image on a screen to the driver, so called vision applications, and
2. cameras that are used as input for computer algorithms that give a notification or warning to the driver, so called sensing applications.

In this chapter we will focus on cameras for passenger cars. Similar requirements exist for cameras on commercial vehicles.

The most common vision applications are:

- rear view
- surround view
- night vision and
- road crossing monitoring.

The most common sensing applications include:

- advanced driver assistance system (ADAS) applications
- blind spot detection (BSD) and lane change assist (LCA)
- parking assist and
- driver drowsiness monitoring.

235
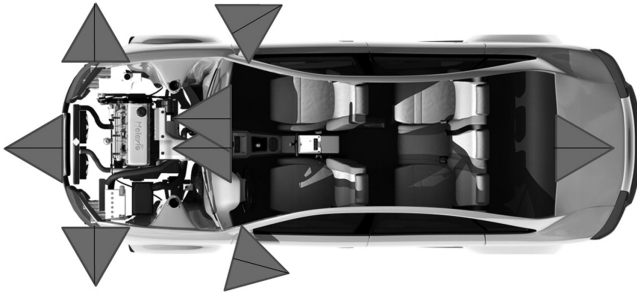
8.1 Common camera locations for the different applications.

Figure 8.1 shows some common camera locations for the different applications.

8.2   Vision systems

8.2.1   Rear view systems

Rear view systems are based on a rear view camera and a display for the driver where the rear view scene is depicted. The camera is often located near the rear central braking light or near the license plate. First generation systems are based on VGA (video graphics array, i.e. 640 × 480 pixels) resolution image sensors, featuring linear dynamic range (in contrast to high dynamic range image sensors (see Section 8.4.3)). First generation systems often feature an analog connection between the camera and the display, for example an NTSC (National Television Standards Committee) coaxial link.

Often a graphical overlay is presented on top of the camera image, which appears as colored lines for indicating the drive path and the distance. This overlay can either be static or dynamic. Dynamic overlays generate curved lines that adapt with the steering wheel angle. For dynamic overlay systems an electrical link is needed between the steering angle sensor and the rear view camera. As a common configuration, the steering angle sensor puts its angle information on the CAN (car area network) bus, which is then read by the rear view camera processor. In the US, possibly from September 2014 onward, the Cameron Gulbransen Kids and Cars Safety Act mandates that all new passenger cars must have a rear view camera to protect children and other vulnerable road users from potential back-up accidents.

8.2.2   Surround view systems

Surround view systems provide the driver with a top-view image, showing the car and the obstacles around the car. Multiple cameras with a wide field of view are used. A common configuration includes one front camera, one rear camera and two side cameras located in the side mirrors. The streams

from all cameras are stitched together by a centrally located processor (surround view electronic central unit – ECU), which generates the top-view image for the driver.

Surround view as well as rear view cameras need to cover a wide field-of-view of around 180°. Wide angle lenses like fish-eye lenses are used for this. The main drawbacks of these lenses are the strong image deformation and also the lower light intensity in the corners of the image, called vignetting. These issues can be solved by post-processing and require a higher image sensor resolution, especially in the corners of the image.

First generation systems are based on VGA image sensors with analog interfaces with analog NTSC as a common standard. Newer generation surround view and rear view systems offer feature enhanced image quality: higher dynamic range, higher sensitivity, higher resolution (1.2 Mpixel) and a digital interface (LVDS – low voltage differential signal or 100 Mbps Ethernet). In the near future, surround view and rear view systems will also include object and pedestrian detection. A warning or even an automatic braking decision can be generated based on the sensing algorithm's findings.

8.2.3 Night vision systems (Fig. 8.2)

To boost the performance of complementary-metal-oxide-semiconductor (CMOS) cameras at night, additional light should be used. In automotive, the light output of a car's front lights is limited to avoid blinding of oncoming traffic. Visual camera based automotive night vision systems apply extra near infrared (NIR) light to the scene. This light has a wavelength range of 800 μm to 1000 μm. The human eye is not sensitive for these wavelengths, so upcoming traffic is not blinded, while CMOS cameras are sensitive in the near infrared wavelength wave. With night vision systems with active light, the driver can see approximately three times further.[1] Low light sensitivity, high dynamic range and extended responsivity in the NIR band are key camera specifications for this application. Automotive night vision systems can also be based on far infrared (FIR) cameras, also called thermal

8.2 Example of an automotive night vision image.

cameras. These cameras detect thermal energy instead of light energy. FIR sensors are outside the scope of this book.

8.2.4 Road crossing monitoring

Based on two side looking cameras in the front bumper, a display shows these two camera images to the driver. Especially for cars with long hoods, the front placed cameras will see crossing traffic before the driver is able to do so.[2]

8.3   Sensing systems

8.3.1 ADAS

ADAS functions can be split in two groups: comfort functions and safety functions. The aim of the comfort functions is to warn the driver by triggering a warning, like a flashing light, sound, vibration or even a gentle steering suggestion. The aim of safety functions is to take action on the vehicle itself in cases where the driver is not responding to a potentially dangerous situation. Potential actions include brake pre-charging, safety belt preparation, hood lifting, automatic braking, evasive steering, etc.

Common comfort and safety ADAS functions include:

- lane departure warning (LDW)
- forward collision warning (FCW)
- automatic high beam assist (AHB)
- traffic sign recognition (TSR)
- object detection, pedestrian detection, potentially combined with automatic emergency braking (AEB).

These ADAS functions are based on one front camera or on a front stereovision camera. Sometimes the camera information is supplemented with information from other sensors like light detection and ranging (LIDAR) or radio detection and ranging (RADAR). ADAS cameras are located inside the car, against the front windshield, behind the central rear view mirror. The ADAS camera field of view is located in the wiper area to keep the glass in front of the camera as clean as possible. Sometimes, RADAR sensing, vision sensing and data fusion are combined in a single module.[3]

New Car Assessment Program (NCAP) safety rating for ADAS

The US NCAP (New Car Assessment Program) and EU NCAP attribute safety points to cars that have LDW and FCW functions on-board. From 2014 onwards, the EU NCAP[4] attributes safety points to cars that have
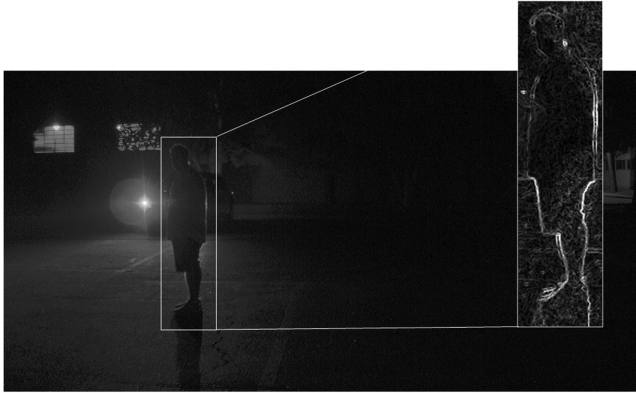
8.3 Example of a pedestrian detection application.

city-AEB and urban-AEB. City-AEB looks 6 to 8 m ahead, detects vehicles and large obstacles to avoid low speed impacts up to 20 km/h to avoid whiplash. Common technologies used are LIDAR or short range RADAR. Urban-AEB looks up to 200 m ahead and operates over the speed range of 50–80 km/h to avoid driver injuries. Common technology for this system is LIDAR or long range RADAR. From 2016 onward, the EU NCAP attributes safety points for cars that include interurban-AEB systems. These systems protect pedestrians and other vulnerable road users like cyclists by classification of road users and obstacles based on camera and RADAR (Fig. 8.3). AEB systems automatically brake the car in dangerous situations when the driver is not responding to a detected threat.

8.3.2 BSD and LCA

BSD systems give a warning to the driver if a car or motorbike is present in the blind spot area of the car mirrors. A LCA system not only covers the blind spot, but also looks at potentially overtaking cars at the left and right side of the vehicle. When the driver initializes a lane change, the LCA issues a warning when it is unsafe to do so[2]. Blind spot-like cameras implemented on the front bumper can also be used to spot oncoming traffic at intersections with poor visibility.

8.3.3 Parking assist

Based on information from the surround view system and from other sensors, the car provides parking assistance or self-parking capabilities to the driver. These systems are rapidly evolving towards complete autonomous parking systems, potentially without a driver behind the steering wheel.

8.3.4 Driver drowsiness detection

Driver drowsiness detection can be done in an indirect way, for example by analyzing the small forces a driver applies on the steering wheel, by means of a precise steering angle sensor. Direct monitoring can be done with an interior camera looking at the driver's head position and/or eyes. To also operate at night, the system requires a light source. As with night vision, NIR light sources are preferred as NIR light is invisible to the driver. Other technologies like low resolution FIR cameras or time-of-flight 3D cameras can support this application.

8.4   Requirements for automotive image sensors

8.4.1 Resolution

There are a number of significant differences between requirements for consumer and requirements for automotive applications. Consumer image sensors are following an evolution path towards ever higher resolutions. Current automotive sensors feature a (wide-)VGA or 1.3 Mpix resolution. With a 1.3 Mpix resolution, pedestrians can be detected at 100 m distance with a lens field-of-view in the range of 40–60°.

As a second example, the resolution of surround view systems is determined by the display resolution and by the amount of image distortion caused by the wide lens field-of-view. This application requires a typical lens field of view between 170–190°. These lenses are called fish-eye lenses and generate high image deformations in the corners. The deformation can be removed in post-processing provided there is sufficient resolution at the corners of the image. Newer surround view systems feature a camera resolution of 1.3 Mpix.

Another key aspect that defines resolution in automotive systems is system cost. With a higher resolution, sensing systems need to process more pixels and thus require larger and more expensive processors. Additionally, with higher resolution and fixed optical format the lens MTF must increase, leading to more expensive lenses, and thus further increases system cost.

8.4.2 Low light performance (sensitivity)

Pixel size and low light sensitivity go hand in hand: generally speaking bigger pixels, lead to higher sensitivity. As the applications requirements specify minimum resolution needs, pixel size will drive lens size (optical format) and thus module size. Car makers impose maximum limits to module sizes for aesthetics reasons and for mounting space reasons.

Especially in the side mirrors, where cameras are located for surround view or blind spot detection, the mounting space for a camera is extremely limited. At the time of writing, common optical formats for automotive cameras range from ¼″ to ⅓″. Combining this optical format with a resolution range between VGA and 1.2 Mpix leads to pixel sizes between 3.7 μm and 6 μm. A pixel size of 3.7 μm is considered state-of-the-art for automotive image sensors at the time of writing.

In consumer cameras, often a high gain factor is used to 'compensate' for the relative small pixel size (lower sensitivity) in dark scenes. Applying gain not only amplifies the signal but also the noise. In automotive sensing systems the detection algorithms are noise sensitive, so the key metric should be defined in terms of signal-to-noise ratio (SNR).

As an objective measure of low light performance, the European Machine Vision Association (EMVA1288)[5] recommends the use of SNR1 (signal-to-noise ratio) and SNR10. SNR1 provides the light irradiance level so that on the output of the image sensor the signal level is equal to the noise level. At the corresponding light irradiance level for SNR10, the ratio between signal and noise is a factor 10. The unit of SNR1 and SNR10 is $nW/(cm^2.s)$. These parameters can be measured without optics and are not influenced by gain or integration time. The wavelength and temperature needs to be specified, for example 535 nm (green) light at 25 °C for example.

8.4.3   Introduction to high dynamic range (HDR)

Next to low light sensitivity, a high dynamic intra-scene range is a second important key performance parameter for most automotive camera systems. The ratio of the highest to the lowest intensity of light in a scene is known as the intra-scene dynamic range. If the dynamic range of a camera is too narrow to accommodate the intra-scene dynamic range, the resulting image may miss important scene details. Once lost, these cannot be recovered through post-processing. Therefore HDR technologies are essential for providing the most reliable image information under a wide range of intensities and directions of illumination. Figure 8.4 shows a side-by-side comparison of a linear image and a high dynamic range image. The linear image shows a saturated zone with resulting loss of information because the scene dynamic range is larger than the dynamic range of the camera. The histogram measurements given in Fig. 8.5 show that during daytime intra-scene dynamic range can be as high as 90 dB. During night time the intra-scene dynamic range can even be higher, up to 120 dB.

For static scenes, a high dynamic range picture can be composed of multiple lower dynamic range images taken with different exposure times. Taking multiple pictures in a moving car would result in unacceptable motion artifacts in the image, hence the need for a high (120 dB) dynamic

8.4 Side-by-side comparison of a linear image and a high dynamic range image. The linear image shows a saturated zone with resulting loss of information because the scene dynamic range is larger than the dynamic range of the camera.
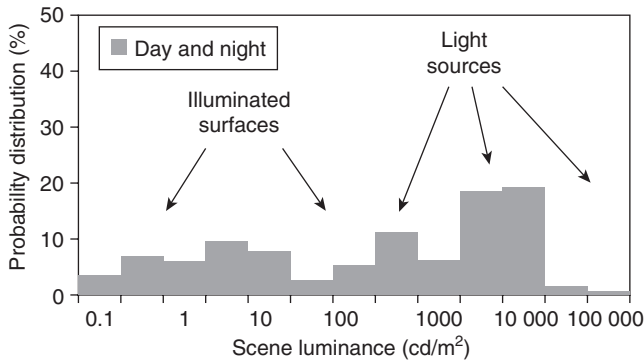
8.5 Measured dynamic range of a night scene.

range in every single frame. Examples of automotive scenes that require a high dynamic range camera include:

- driving in and out tunnels, parking garages, etc.
- driving towards the sun
- dusk
- driving at night with lights of upcoming traffic shining into the camera lens.

In the next paragraph we will show that the mentioned standard definition of dynamic range (ratio of the highest to the lowest detectable light intensity level) is not sufficient for automotive cameras. Automotive cameras need to guarantee minimum local contrast (minimum SNR) over the complete
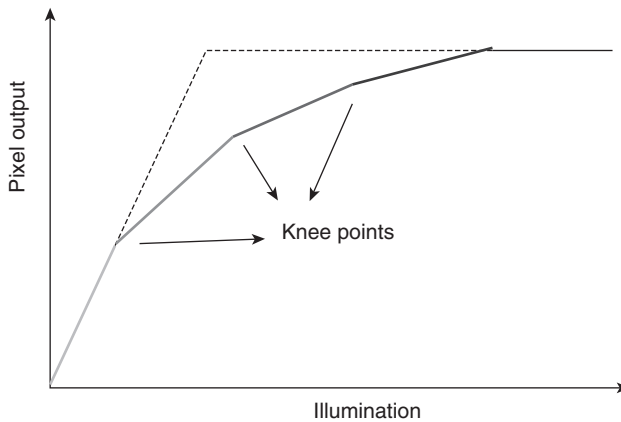
8.6 Example of a piece-wise linear response curve with three kneepoints.

light range of the scene. If this condition is not met, the detection algorithms will not be able to function properly for specific grey tone values.

HDR principle with non-linear response curve with kneepoints

For automotive image sensors, it is preferred to limit the number of bits on the output of the image sensor. A too high bit count would result in a more expensive image transport infrastructure. Up to 12 bit outputs are a good trade-off between cost and performance. How to map 120 dB HDR scenes to a 12 bit (72 dB) output? A logarithmic image sensor can do this by design, but has other limitations, like limited flexibility of the response curve.

A common way for automotive image sensors to map a HDR scene in a 12 bit output is to apply a piecewise linear response curve: pixels looking at a dark region of a scene should have the highest possible responsivity (slope of the response curve) to make low-light objects (pedestrians for example) visible in the output image. Pixels looking at a bright region of a scene can be given a lower responsivity, i.e. lower slope of the response curve. The different linear segments of the image sensor response curve are connected by kneepoints (Fig. 8.6).

Theoretical analyses and measurements show that the number of kneepoints is the most critical parameter in HDR image sensors. A too low number of kneepoints will result in information loss in some parts of the light range (Fig. 8.7). Depending on the kneepoint settings, information loss will occur in the low lights, mid tones or bright tones. Analyses show that for mapping a 120 dB scene into a 12 bit output, the minimum number of kneepoints in the response curve must be 5 (Fig. 8.8). For more information we refer to an HDR white paper.[6]
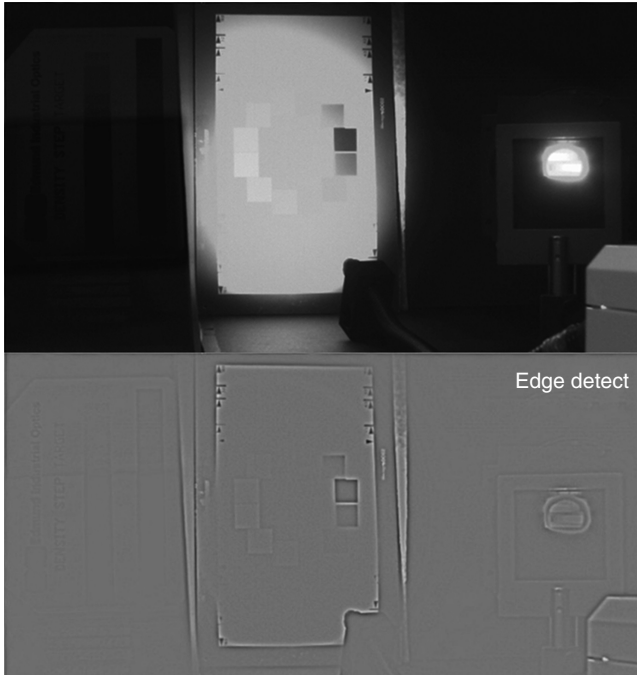
Edge detect

8.7 Picture taken with an HDR imager with two kneepoints: there is information loss in the mid tones when the dynamic range of the scene is larger than 80 dB: the edge detect algorithm does not detect some of the mid tones in the middle of the image.

8.4.4 Temperature range

Automotive image sensors must withstand a wider temperature range than their consumer counterparts, as the automotive parts are directly exposed to various atmospheric conditions, ranging from North Cape freezing temperatures to Sahara sun conditions. To make things worse on the high side of the temperature scale, automotive camera modules need to be fully closed to remain clean from moisture and dust. This implies active cooling strategies like fan blowers cannot be used. Another reason why an automotive image sensor can be exposed to higher temperatures is the camera module location. If we take for example a front camera for ADAS applications: these camera modules are generally attached to the front windshield, fully exposed to the sun. Module temperature is thus determined by the sun load and the heat generated by the electronic components, including image sensor, processor, power electronics and interface electronics. For these reasons, a typical automotive image sensor temperature range is between –40 °C and as high as 105 °C/115 °C.[7]
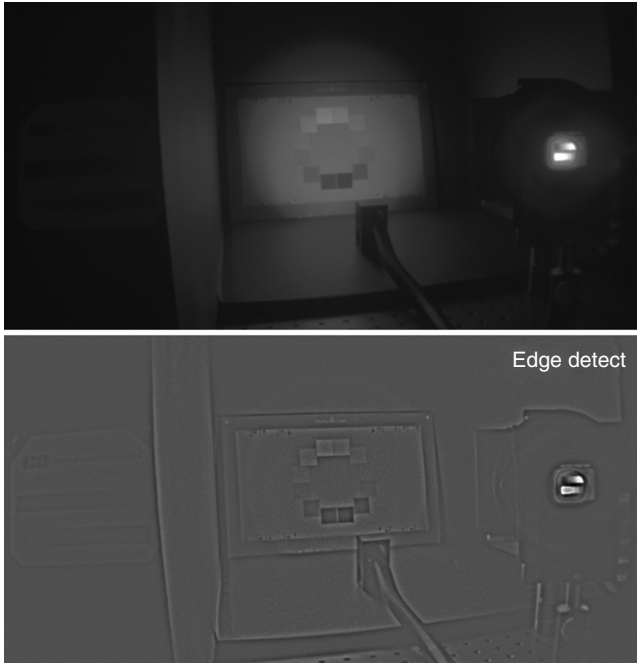
Edge detect

8.8 Picture taken with an HDR image sensor with six kneepoints: no information loss when the dynamic range of the scene is 120 dB: the edge detect algorithm detects low tone, mid tone and high tone information.

8.4.5 System integrity

As mentioned in Section 8.3.1 the group of driver assistance applications is evolving towards safety applications, where the driver assistance system is actively impacting the car's behavior, like for example triggering an AEB action when a pedestrian is detected in the drive path of the car. Triggering these automatic actions has an important impact on the driver, the car and the immediate environment of the car. Utmost care must be taken to avoid false positive actions. In automotive terms, system integrity is defined in the ISO 26262 standard[8] and defines four automotive system integrity levels ASILs: A, B, C and D. ASIL A defines the lowest system integrity.

For the automotive AEB function, the common request from car makers is that the AEB system should be ASIL B compliant. The image sensor can support a specific ASIL level by providing extra functions for communication integrity and device functional integrity. Communication integrity between the image sensor and the processor can be supported by applying cyclic redundancy codes (CRC) on all data exchanges. Device functional integrity can be supported by providing an on-chip temperature sensor, watchdog functions and additional pixel-to-output integrity check circuitry.

8.4.6   Image sensor interface

A common digital interface for automotive image sensors is the low-voltage TTL parallel output interface (for example providing 12 bits per pixel), offered together with a separate horizontal synchronization signal output (HSYNC), a vertical synchronization signal output (VSYNC) and a clock signal output. Some automotive HDR image sensors provide a linearized output (of 24 bits for example) by means of a high speed serial output or a double speed parallel output. Special care has to be taken against frequency noise increase because of the higher frequency output clock and thermal noise increase caused by additional power consumption.

8.4.7   Camera module interface

To transfer image data over a car's cable network, special care has to be taken for electromagnetic interference/electromagnetic compatibility (EMI/EMC). LVDS and LVDS variants are commonly used protocols provided by multiple automotive electronics products to transport the image data from a camera module to the display or camera processing unit (on ECU) in an uncompressed form. LVDS-based transport uses a shielded double twisted pair cable or a coaxial cable.

Optical technologies like 150 Mbps multimedia optical synchronous transport (MOST)[9] are available on the market. Higher data rates are currently being developed. A major advantage over electrical data transport is the immunity from electromagnetic radiation from other sources. Especially hybrid and electric cars power switching functions generate high amounts of electromagnetic radiation. However, not all car makers embrace optical transport technologies because of possible higher cable architecture costs.

A more recent alternative is the use of 100 Mbps automotive Ethernet over a single pair, unshielded cable. The Open Alliance Special Interest Group[10] established industry standards for this. 100 Mbps automotive Ethernet would enable cost reduction on cable architecture level, but requires compression of the image stream, with potential impact on delay and detection algorithm performance as a consequence. Reduced twisted pair gigabit Ethernet (RTPGE)[11] promises to support uncompressed in-car data transport by providing a frame rate up to 1 Gbps transport.

8.4.8   Color

Viewing applications will generally show a color image on a car display. Color processing can be done in a separate chip (integrated signal processor

(ISP) chip or companion chip) or can be done by adding digital circuitry on the image sensor (system-on-chip (SoC)). SoC-based system architectures allow smaller camera modules but suffer from lower sensitivity, as the power consumption for the digital processing increases heat on the image sensor die, and thus increases dark noise.

Some detection applications benefit from a certain amount of color awareness. Traffic sign recognition (TSR) and automatic high beam (AHB) functions benefit from access to the 'red' information. For these applications, a red-clear based color filter array (CFA) on the silicon is standard. A red-clear CFA pattern generally covers one out of four pixels with a red color filter while the remaining three pixels have no color filter (clear), also known as an 'RCCC' CFA. Full color based detection systems are used to detect different lane marking colors. The color system is built in the same way as for consumer cameras:

- NIR light is blocked on the lens or on the glass of the image sensor package
- a commonly used red, green, blue (RGB) Bayer CFA is applied on the image sensor silicon
- the colors are interpolated by post-processing, attribute a calculated three-color value to each pixel: R, G and B.

8.4.9 Combined colors with NIR light

As CMOS materials are also sensitive in the NIR wavelength 750–1000 nm, light energy in this wavelength can also be used for improving night performance of automotive viewing and sensing systems. The reason why most color systems block the near infrared light is because near infrared light may interfere with color quality. Recent advances in post-processing algorithms enable conserving color quality (for example during daytime) while boosting color system sensitivity at night time.

8.4.10 Optics

For robustness reasons, automotive camera systems avoid using moving parts. For lenses this implies the use of fixed focal point lenses. A typical automotive lens will use multiple glass or plastic elements. The F# typically varies between 1.5 and 2.5. Field of view (FOV) is application dependent and can range from 30° for night vision systems to more than 180° for surround and rear view cameras.

Special care needs to be taken for HDR functions. Under HDR conditions, a bright light source (for example, headlights) can impact the darker parts

of the scene, due to light energy diffusion inside the lens, causing halos, stray light and ghost objects. Lenses suitable for HDR applications typically apply anti-reflective coatings on each lens surface.

8.4.11 Automotive qualification test criteria

Automotive electronic components, including image sensors, are qualified according to the AEC-Q100 standard.[12] This standard defines qualification test criteria for tests including HTOL (high temperature operating lifetime), ESD (electronic discharge), thermal cycling, humidity testing, etc. There are no exceptions or specific standards foreseen in the AEC-Q100 standard for image sensors.

8.5   Future trends

There is a clear trend towards increased use of cameras in cars. A growing number of vision functions offer additional information to the driver by showing the car surroundings including objects and pedestrians on a display, potentially enriched by overlay graphics. Sensing systems based on computer algorithms are becoming commonplace, replacing display-based information by warnings, sounds, flashing lights or vibrations to attract the driver's attention to a potential dangerous situation. AEB functions based on pedestrian detection by camera and RADAR are beneficial for vulnerable road user's safety and will become commonplace from 2016 onwards. In case a braking action is insufficient to avoid an accident, automatic evasive steering actions could be triggered if the car is aware of the surrounding environment by means of surround cameras.

With camera based computer algorithms gradually supporting and sometimes taking over the driver function, autonomous cars are the logical next step. Will future drivers embrace this function? Will legal and liability issues get a sufficiently satisfactory solution? One thing is clear: thanks to advances in image sensor technology, our children will be raised in a – from an automotive viewpoint – safer world.

At the time of writing, work is ongoing at ISO (International Standardization Organization) to define a standard for the use of cameras and displays instead of rear view and blind spot mirrors: ISO16505 'Road vehicles – ergonomic and performance aspects of camera-monitor systems'.[13] As out-of-the-car placed cameras are generally much smaller than their mirror equivalents, this leads to better aerodynamics of the vehicle and thus saves energy.

In the near future, rear view and surround view systems will also detect obstacles and pedestrians. Based on the sensing algorithm's findings, a warning is then issued to the driver or an automatic braking decision is

triggered. Stereovision front cameras for ADAS are emerging to enhance distance detection, object and pedestrian sensing performance.

8.6   References

1. http://www.bosch-automotivetechnology.com/en/com/driving_comfort_com/driving_comfort_systems_for_passenger_cars_com/driver_assistance_systems_comfort_pc_com/driver_assistance_systems_com.html
2. http://www.clarion.com/xe/en/tech/technology/camera/index.html
3. http://delphi.com/manufacturers/auto/safety/active/racam/
4. http://www.euroncap.com/home.aspx
5. http://www.emva.org/cms/index.php
6. http://www.melexis.com/Assets/Autobrite-Whitepaper-5815.aspx
7. http://www.melexis.com/Optical-Sensors/Optical-Sensing/MLX75411-696.aspx
8. http://www.iso.org/iso/catalogue-detail?csnumber=43464
9. http://www.melexis.com/Optical-Sensors/Optical-Datacommunications/MLX75605-671.aspx
10. http://www.opensig.org/
11. http://www.ieee802.org/3/RTPGE/
12. http://www.aecouncil.com/AECDocuments.html
13. http://www.iso.org/iso/home.html

9
Complementary metal-oxide-semiconductor (CMOS) image sensors for use in space

J. B O G A E R T S, CMOSIS, Belgium

Abstract: This chapter discusses the use of complementary metal-oxide-semiconductor (CMOS) sensors for operation in space. It first introduces specific requirements for sensor use in space, then compares the benefits and drawbacks of CMOS sensors and CCDs in terms of typical sensor requirements. Finally, a number of developments in CMOS sensors for space applications are described in more detail.

Key words: CMOS sensor, space application.

9.1   Introduction

This chapter discusses the use of complementary metal-oxide-semiconductor (CMOS) sensors in space, introducing first the specific requirements for using sensors in space. In addition to the electro-optical performance of the detector, other requirements such as radiation hardness, power consumption and size and service life are of equal importance. The second section of the chapter compares the benefits and drawbacks of CMOS sensors and their charge coupled device (CCD) counterparts in terms of typical sensor requirements. Not only do technical specifications play a role when selecting the type of detector, but also economical and organisational aspects should be considered. Finally, a number of developments in CMOS sensors for space applications are described in more detail. Although the range of applications is too diverse to make a comprehensive list, examples of different CMOS sensor developments are categorised by their application, together with short descriptions of technological requirements.

Today, the mainstream technology for space imaging is still CCD technology. Given the inherent advantages of CMOS sensors with respect to CCD technology (e.g. radiation hardness, system size and power consumption, readout speed, integration) it is very likely that CMOS sensors will replace CCDs in a large number of applications. CMOS sensors will also enable new application areas that cannot be covered by CCDs. However, the adoption of CMOS and replacement of CCDs in space imaging will be much slower than that experienced in consumer markets in the last decade. This is partly due to specific technological requirements, but mainly because of industrial organisation in the space community, the long development

periods necessary in space programs, and stringent reliability and quali-fication requirements.

9.2   General requirements for use of complementary metal-oxide-semiconductor (CMOS) sensors in space

As discussed in previous chapters, the imaging landscape has changed drastically over the last few decades. Strengthened by their use in the mobile phone mass market, CMOS active pixel sensors continue to expand their already commanding lead of the overall image sensor market, pushing rival CCD sensors into an increasingly isolated space. One of such niche markets is space-borne imaging, or scientific imaging. Although the advantages of CMOS sensors were already recognised in the late 1990s, the introduction of CMOS sensors in space programs has developed at a much slower pace and CCDs still constitute the main part of silicon sensors for the visible wavelength range.

There are many reasons why the adoption of CMOS sensors is slower. First and most obvious is the economic background. In a consumer market, manufacturing cost reduction is a key driver. Since CMOS sensors are based on CMOS technology (being by far the dominant processing technology for electronic integrated circuits), they cash-in on the rapid progress of technology and scaling trends, and therefore lead to a lower system cost and size. For space-borne imaging, system cost plays a role, but is not of prime importance, especially since the manufacturing cost of the sensor or system is a relatively small part of the total space program budget. Other reasons can be found in the special requirements of sensors to be operated in space. These are similar for CMOS and CCDs, but to fully understand the impact on the final product, and to master all techniques to comply with these requirements, manufacturers need time to build-up experience and competences. Below is a list of specific space requirements:

- Radiation hardness. There are many different radiation-induced degradation effects. CMOS has some inherent advantages with respect to CCDs, but the diversity in effects may be even larger given the more complex circuitry, the many different technologies that are used, and so on. Testing and understanding all these effects in terms of type of damage, influencing parameters, and annealing effects, takes a lot of time.
- Packaging.   Space sensors are in many cases quite large. Many different aspects in packaging need to be taken into account, such as die flatness after assembly, planarity, window bow, placement accuracy, thermal expansion, qualification (hermeticity, shock and vibration), use of

restricted type of materials (adhesives, encapsulants and sealants) to prevent outgassing and resulting condensation on optical elements, and so on. Full space qualification of products can easily take years.

- Power dissipation, mass and system size.   This is an aspect where CMOS sensors out-perform CCD-based systems, which has led to rapid introduction of CMOS in some application domains such as attitude and orbit control system (AOCS) applications. For instrumental applications, this is of less importance with respect to electro-optical characteristics.
- On-board storage and download capability.   CMOS sensors can be read out at high frame rates, but obviously the sensor data bandwidth needs to be handled by the satellite infrastructure. CMOS sensors can open new possibilities for data gathering and different system architectures, but often in different ways that need to be compared to each other and agreed upon.
- Reliability and mission lifetime.   Space applications have to cope with extended temperature conditions (temperature range and cycling), handling and storage conditions, pressure, vibration and shock before and during launch or in orbit. In that respect, the space community most often chooses a conservative approach for risk mitigation, and because of cost and time of qualification. There is some similarity with reliability and packaging requirements for automotive applications (see Chapter 8).
- Long-term availability, design for obsolescence and access to foundry processes.   The technology used to manufacture CMOS sensors is determined by a number of technical and commercial aspects. Examples include availability of stitching, hybridisation, specific post-processing steps for back-side illumination (BSI), optical filters, and so on. Since the duration of an average space program is very long, there are concerns with respect to the long-term availability of and access to the processes. CMOS scaling evolves aggressively and the application to space is merely a small market within the whole CMOS technology.
- Detection of other than visual (VIS) wavelengths (UV, extreme ultraviolet (EUV)).   This typically also requires process technology development and can take a long time to mature.

9.3   Comparison of CMOS sensors and charge coupled devices (CCDs) for space applications

The mainstream technology used for space imaging today is still CCD technology. Optical requirements are stringent: high quantum efficiency and fill factor over a wide spectral range, low noise, low defects and high resolution. Despite continuous efforts in developing and optimising CCD

technology, it remains a specialised process and thus does not benefit from the much larger investments that are being made in CMOS processing technology, design and manufacturing. In the 1990s, many developments were started to investigate how CMOS sensors could replace CCDs in several low- to medium-end space imaging applications. These developments were driven by the idea that those applications could cash-in on CMOS sensors' inherent advantages over CCDs. On the other hand, CCD technology has built up a long heritage in space imaging and is a mature technology.

9.3.1 Radiation hardness

Radiation hardness is the most specific requirement related to space usage. It is difficult to summarise briefly the broad range of radiation degradation effects, but we aim to summarise the most prominent mechanisms in order to point out the main differences between CMOS and CCD sensors and introduce the typical requirements related to radiation tolerance. 'Radiation' is a rather comprehensive word. Anything in our immediate or far environment is continuously exposed to different types of radiation. CMOS and CCD sensors are actually built to interact with or measure a particular type of radiation. In this context, we are concerned with radiation that is possibly harmful to electronic circuits, particularly for CMOS or CCD sensors. The radiation in the space environment will affect the performance of electronic circuits and therefore poses a direct limit to the lifetime of its application in space.

   This radiation consists of many different particles that impinge on the devices and cause undesirable effects of a different kind. The most important particles are electrons, protons, alpha particles, photons, heavy ions, and so on. They can be separated into two main categories: photons and charged particles. The nature of the interaction of the incident particle and the target depends on several properties, as well of the particle that impinge (mass, energy, charge) and of the target material (mass, charge and density).

   The interaction of radiation with a sensor may be quite complex, considering the various types of particles, the wide range of particle energies, and the variety of interactions that occur. In a real environment, the effects that are produced in the irradiated components are usually the result of a combination of many different interactions, even if the type and energy of the impinging particle is clearly defined. Two basic mechanisms can be distinguished for radiation effects on electronic devices: generation of electron-hole pairs (ionisation, just like for visible photons) and the displacement of atoms from their lattice positions (displacement damage). Particles passing through a silicon device will lose their energy by ionisation of the silicon and the oxides and deposit the remainder of the energy into

Impinging particles
(e.g. electrons, protons, photons, heavy ions)

Ionisation

Displacement of
lattice atoms

1–100 MeV protons

Total ionising
dose

Single event
effects (SEE)

Co-60, X-rays, electrons   Heavy ions, Cf-252

Single event
upset (SEU)

Latch-up
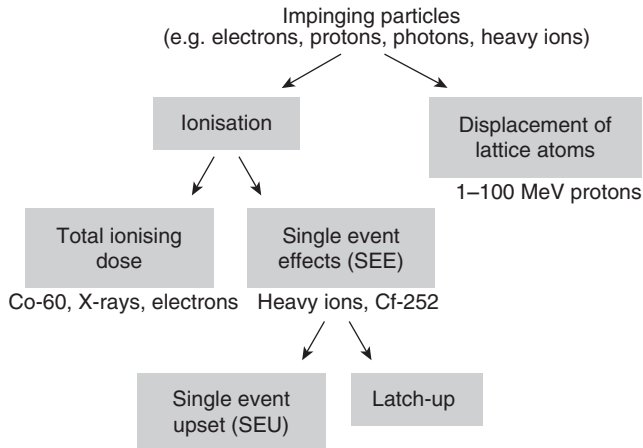
9.1 Overview of various radiation-induced degradation effects on
electronic components.

displacement. Figure 9.1 gives a simplified chart showing the possible
radiation-induced degradation effects on electronic components. In Fig. 9.1,
the particles typically used to study the different degradation effects are
mentioned.

Ionisation

The deposition of energy by means of ionisation can lead to transient or
permanent degradation due to an accumulation of a certain total ionising
dose. On the other hand, single heavy ions can generate enough electron-
hole pairs to cause single event effects (SEEs). Ionising radiation generates
mobile electrons and holes in both the silicon substrate and the insulator
in silicon devices and leads to a build-up of a certain dose. This dose from
ionising radiation is expressed in terms of gray (Gy), but many results are
still published in terms of the rad (radiation absorbed dose). One gray
equals 100 rad.

Total ionising dose (TID)

Ionising radiation gives rise to the production of positive charges in the gate
and field oxides of metal-oxide semiconductor (MOS) structures. As a
consequence of this positive charge build-up, the flatband and threshold
voltages of the particular MOS structure will decrease and this shift can
persist in time for hours to years and is not sensitive to the silicon surface
potential. Fortunately, in the presently used CMOS technologies, the
transistors have a very thin, high-quality gate oxide for which the build-up

Standard, linear   Circular or enclosed geometry

linear with [ ]-shape of field oxide   I-shape with [ ]-shape of field oxide

▨ Polysilicon gate on active (gate oxide)
▨ Polysilicon gate on field oxide
▨ Field oxide
■ n+ doped active region
■ p+ doped active region
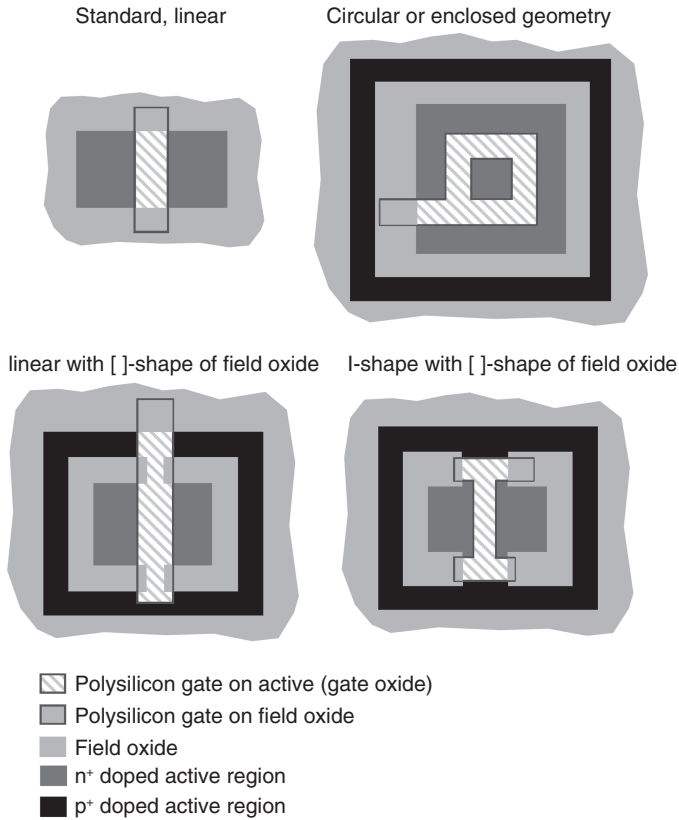
9.2 Examples of layout techniques to prevent source to drain leakage current in standard linear nMOS transistors (after Bogaerts (2002)).

is very small. In contrast, a local oxidation on silicon (LOCOS) or shallow trench isolation (STI) process is commonly used to isolate different devices. These oxides are typically much thicker than the gate oxides and their properties are not as well controlled. Therefore, relatively low total ionising doses produce large changes of threshold voltages. While the field oxide and p-well implantation provides isolation between adjacent devices before irradiation, the threshold voltage of the field transistors is not always sufficient to prevent leakage between source and drain in an ionising radiation environment.

For optimal radiation tolerance, CMOS sensors also need special design techniques such as gate-all-around transistors, for example. Figure 9.2 shows some possible solutions (Bogaerts, 2002). The traditional way to circumvent this problem is through the use of guarded circular transistors. This solution is, however, not so efficient in its use of silicon area, certainly when several transistors have to be combined with an additional photodiode in a small
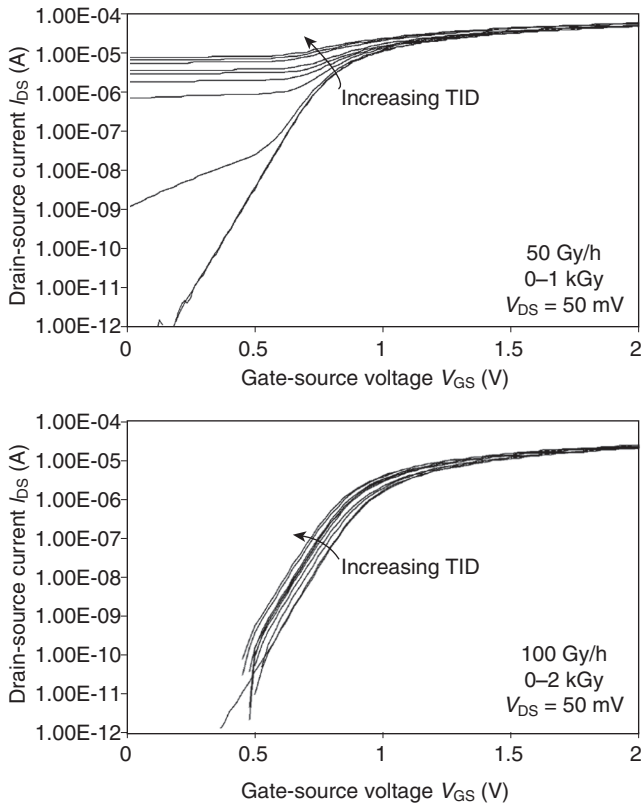
9.3 Input characteristics (I_{DS}-V_{GS}) of standard linear nMOS (top) and radiation-tolerant layouted nMOS (bottom) transistors under ionising radiation (both 0.5 µm CMOS technology) (after Bogaerts (2002)).

pixel area. Each solution is based on the consideration that every possible path between two n^+ regions that have to be electrically separated must be interrupted by a p^+ guard ring or by a silicon region with gate oxide on top.

Figure 9.3 shows the effect on the input characteristics (I_{DS}-V_{GS}) of the n-type metal-oxide-semiconductor (nMOS) transistors (in this case 0.5 µm CMOS technology). The top curves show a clear increase of sub-threshold leakage current in the case of the standard layout. As shown at the bottom, the curves of the radiation-tolerant designed transistor do not exhibit such an increase. The measurement shows only the threshold voltage shift, confirming that the gate oxide can easily withstand total ionising doses of 1 kGy, i.e. the threshold shift observed in gate oxides of modern CMOS technologies are very small.

There is an additional process of build-up of interface traps at the Si/SiO$_2$ interface. The latter build-up can result in a threshold voltage shift that is

dependent on the silicon surface potential. It also results in increased leakage current or pixel dark current, noise and/or image lag due to trapping of charges in the transfer region in the case of pinned photodiode pixels. Ionising radiation can also degrade the photo-responsivity of active pixels. This can be due to several causes: increase in the pixel capacitance and/or decrease in the gain of the readout transistors, recombination of photo-generated charges at the surface and an increase in the absorption and reflection in the dielectrics that cover the active pixel. More background information on radiation-induced degradation can be found in Bogaerts (2002), Goiffon *et al*. (2008, 2009a, 2009b, 2011) and Padmakumar (2009). The total ionising dose tolerance of standard CMOS sensors is usually limited to a few hundreds of grays, although newer technologies tend to show higher radiation tolerance. This is the same order of magnitude for the total ionising dose at which failure of most commercial CCDs is observed. It is therefore not necessarily correct to claim an inherently higher radiation-tolerance for CMOS sensors. There are, however, a number of potential hardening techniques that can be less easily implemented in CCDs. The gate oxide thickness, which is one of the most important parameters in the response to ionising radiation, is in present-day standard CMOS technologies considerably smaller than in CCD technologies. Flatband voltage shifts and associated performance degradation are thus less problematic. Radiation-induced surface state build-up can also pose a problem concerning the charge transfer efficiency in CCDs.

SEEs

SEEs is a general term that encompasses single event upset (SEU) and single event latch-up (SEL). Apart from the long-lived effects that are usually caused by a particular total ionising dose, an energetic heavy ion or proton may generate enough electron-hole pairs along its track that can affect the functionality of circuits. Such ionising particles can induce a current pulse in a p-n junction. The charge injected at, for example, a sensitive node in a logic circuit, may exceed the critical charge that is required to change the logic state of the circuit. Such a change of state is referred to as a SEU. The vulnerability to SEU increases with the technology evolution towards smaller device geometries as the critical charge decreases with higher densities. For a CMOS sensor, a SEU can occur in different forms. A flux of ionising particles impinging in the pixel array will cause a random arrangement of white speckles, because charges generated by the radiation are collected in the same way as charges from absorbed visible light. The sensor cannot discern different types of radiation. A SEU in the peripheral logic circuits can result in false read or reset pointers, etc. However, a SEU causes no permanent damage and the problem, in most

cases, can most easily be solved by appropriate system design, for example through increased frame rate, majority logic, etc.

Bulk CMOS structures contain parasitic vertical and lateral bipolar p-n-p and n-p-n transistors, and are vulnerable to radiation-induced latch-up. Positive feedback around the n-p-n-p loop can cause the structure to remain in a high-current low-impedance state, which can potentially result in burnout unless the power supply is turned off. Latch-up can therefore lead to permanent damage. To make CMOS sensors latch-up immune, special care has to be taken during the design, for example through the implementation of guard rings.

Displacement damage

Energetic photons and charged particles incident on silicon can also lose their kinetic energy by an energy transfer to the target nuclei and can in that way cause displacement of atoms. The atom that is displaced can in turn interact with other atoms and produce a displacement cascade. Various types of defects (vacancies, interstitials, defect clusters) can be created in irradiated silicon. These defects disturb the lattice periodicity and have therefore discrete energy levels associated with them that lie in the silicon bandgap. These radiation-induced energy levels can have a major impact on the electrical behaviour of the semiconductor device. They can lead to multiple processes: dark current generation through discrete levels in the bandgap, and recombination potentially degrading quantum efficiency. Furthermore, temporarily trapping of carriers followed by emission can induce transfer loss or random telegraph signal (RTS) phenomena. Increased leakage currents can result from defect-assisted tunnelling, etc.

Since CMOS sensors require only one charge transfer for signal readout, they are insensitive to charge transfer efficiency problems produced by displacement damage in CCDs. In general, the dark current increase after proton irradiation also appears to be significantly lower than in CCDs. The reason for this is found in the typically significantly smaller pixel-depletion volume that is sensitive to the number of elastic and inelastic recoils in each pixel in the displacement process.

More generally, radiation-induced degradation can also depend on the detector start material. For example, in the float zone growth technique the silicon molten zone is not in contact with any substances except for the ambient gas, only containing doping gas. Therefore, float zone wafers achieve higher purity and can be used for fully depleted detectors requiring high resistivity. Czochralski wafers, on the other hand, which are used in most commercially available foundry processes, have higher oxygen and carbon concentrations. Oxygen concentration can lead to the formation of unwanted electrically active defects, but on the other hand, can also act as

a gettering agent for metal impurities in the crystal (internal gettering) and radiation-induced defects. These differences become typically important for high proton or neutron fluencies (e.g. Mekki *et al.* 2010).

It is also generally known that p-channel CCDs are more tolerant to displacement damage than conventional n-channel (Bebek *et al.* 2002). It is therefore also expected that hole-based CMOS sensors would show improved radiation tolerance towards proton-induced displacement damage. Unfortunately, this is not yet widely studied since, with the exception of a few published works (Stevens *et al.* 2008), most CMOS image sensor (CIS) processes are electron-based.

9.3.2   System cost

CMOS sensors typically operate on a single low power supply or a few power supplies that can be derived from a single low power supply. Power supply requirements are typically achieved by use of only a few components and may also vary over a larger range without influencing the sensor performance. CMOS technology allows the on-chip integration of analogue and digital functionalities that require external components in case of CCDs. Examples are analogue-to-digital conversion, correlated double sampling, thresholding, signal processing, sequencing, etc. Furthermore, taking into account the limited number of available components, the cost of space qualification and long-term availability of such components, and the system power consumption, it is obvious that system dimensions and mass for CMOS-based systems can be significantly reduced. Payload size, power and mass have – in turn – a major impact on the cost of launch vehicles, and the economic viability of the mission. Thus, even if the cost of the CMOS detector or detector development may be higher in particular cases, the cost at system level can be significantly lower.

9.3.3   Technological progress

The rapid technological progress of CMOS is one of the most important reasons why CMOS sensors have taken over the command in the total imaging market during the last decade. Only a small fraction of electronic circuits are made using non-CMOS related technology. CMOS sensors can benefit from the afore-mentioned aggressive scaling trend, although most advanced CMOS image sensors nowadays require dedicated processing modules or modifications. For example, photodiodes are made through specific implant steps that are not required in non-imaging applications but that are needed to make pinned photodiodes with in-pixel charge transfer. Thin back-end processing steps, colour filters and micro lenses are used to optimise quantum efficiency and angular sensitivity.

However, the mobile phone mass market stimulates huge investments in CMOS imaging while CCD technology lacks such technology drivers. Obviously, CMOS image sensor development for space usage can benefit as well from this technological push, but here may also lie its weakness, at least in the short term. If we limit analysis to space applications, this is because of the different industrial organisation of CMOS sensor providers compared to CCD manufacturers. Historically, CCD providers are typically vertically integrated. They design, process, package and qualify their products themselves. For space applications, the manufacturers use relatively old front-end processes, but offer strongly optimised back-end processes like thinning, backside annealing and anti-reflection coatings. Because of their strong focus on space imaging, they also offer high flexibility for special sensor logistics and space qualification. On the other hand, the space market is too small for the vertically integrated CMOS image sensor manufacturers, that can afford the investments associated with state-of-the-art CMOS processing. Therefore, the development of CMOS image sensors for space applications is characterised by a much larger variety of companies, and much more specialisation is involved. The sensor provider does not provide the complete manufacturing chain.

The CMOS processing is typically done as a foundry service and no longer by the design company. Packaging, testing and/or qualification are not necessarily performed by the company that designed the sensor. This is somewhat in conflict with historical and current requirements from the space industry, where longevity of processes and long-term availability are crucial. This is further complicated by the fact that access to most advanced CMOS processes is only available for manufacturers that can guarantee or show the potential of sufficient wafer production volume at the foundry. Given the limited quantities of sensors for space use, those manufacturers can therefore not solely focus on the space market. It is for this reason that agencies are currently seeking, and setting up, shared and coordinated activities between all parties – agencies, industry, institutes – to ensure that access to the required CMOS image sensors capabilities for science and space community can be maintained and improved.

9.3.4 Technological features

Charge transfer

Since CMOS sensors do not require successive charge transfer through other pixels for signal readout, access to individual pixels or windows is greatly simplified and speeded up. This is required in applications where simultaneous tracking of several objects at a high frame rate is needed. This is, for example, the case in star trackers applications and robotics. Very high

frame rates can be achieved without penalty on noise performance as seen in CCDs. In most modern CMOS sensors, readout of signal occurs through single charge transfer in the pixel where charge is converted into a voltage. The charge transfer is introduced for low dark current and reset noise elimination due to full charge transfer from the pinned photodiode. CMOS sensors are therefore much less prone to charge transfer losses that can occur in CCDs after particle irradiation, although even the single transfer gate in the pixel can introduce dark current increase and/or RTS noise under irradiation (Goiffon *et al.* 2009b, Padmakumar 2009). Note that in this case – as in CCDs – the radiation-induced degradation may differ significantly depending on exact device implementation. This concerns layout of circuits, type of oxides, doping and doping levels that are used, and mode of operation.

Anti-blooming

The inherent anti-blooming available in CMOS sensors is definitely a major benefit for CMOS sensors in many applications. Bleeding or saturation effects in CCDs can render the measurements of objects of interest useless near the saturated pixels, e.g. in star trackers. Additional anti-blooming structures can be introduced in CCDs but often not without penalty with respect to quantum efficiency and full well capacity.

Stitching

Stitching is a method to make chips on the wafer that are larger than the reticle size. The maximum reticle size used in modern CMOS processing technologies is typically about 32 by 25 mm^2. When the final sensor needs to be larger than this reticle size, the sensor design is split up into several sub-blocks. One reticle block contains the repetitive structure that is repeated when the reticle is projected on the wafer. Figure 9.4 shows the general principle of stitching technology. The corresponding reticle layout is shown on the left. In this example the design is split into four parts on the reticle: A/C/I/G block, B/H block, D/F block and E block. The E block is the repetitive structure which contains the majority of the image sensor pixels while block H contains, for example, the column amplifiers, multiplexing logic and output drivers.

During CMOS processing in the foundry, block E, and to a lesser extent the other blocks, are projected several times at adjacent locations. The complete image sensor is shown at the right side of Fig. 9.4. Because blocks need to be repetitive, there are also various constraints to the layout. Sensors as large as a single wafer (6, 8 or 12 inch) have been produced with reasonable yield using this technology with main application field medical
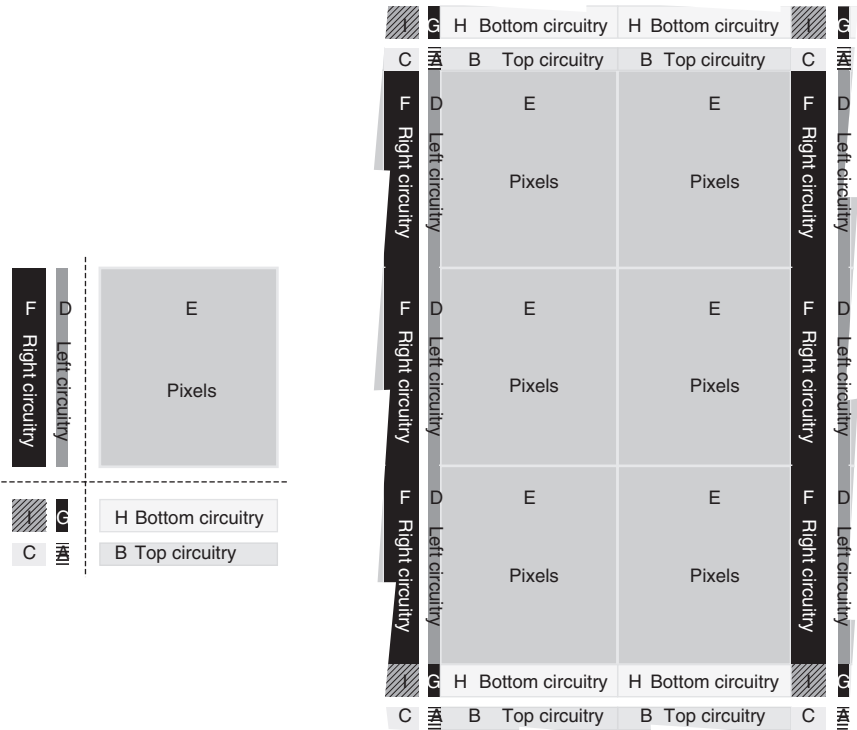
9.4 Example of stitching technology: block distribution on reticle (left), final sensors produced on wafer (right).

X-ray sensors (see also Chapter 12). Not all foundries offer this stitching technology, and certainly not for smaller wafer volumes typically encountered for space sensors. The processing of such wafers is slower because more lithographic exposures are required, including blading. This results, therefore, in a reduction of the foundry throughput.

High speed readout

The lack of multiple charge transfers in CMOS is one of the reasons why CMOS sensors can be read out at very high speed, but there is obviously more required to reach high frame rates. Several techniques exist to speed up the pixel signal readout, i.e. the use of multiple multiplex bus lines in parallel, the use of pipelined readout where pixel access is pipelined with respect to sampled data readout, eventually pipelined with analogue-to-digital conversion, etc.

   On-chip digitalisation, nowadays typically configured in a column wise manner, has the benefit of noise immunity further down the readout chain,

high speed digital interface, possibility for on-chip processing, and decreasing power consumption with scaling CMOS technologies, and so on. High performance imagers have been realised that achieve several Gigapixels per second or more than 20 Gbits per second readout rate. Basically, the limitation in speed will not lie in the sensor readout, but rather in the trade-off between power consumption, satellite on-board processing power, or storage and link download speed. One example of how this problem could be alleviated is the integration of intelligent on-board data reduction as presented in Bréart de Boisanger *et al.* (2011a) where a lightning imager architecture is presented that uses on- or off-chip processing to reduce the data throughput.

Special processing

Examples of special processing required for certain space applications are backside processing to achieve high quantum efficiency at visible wavelengths and UV, EUV or near infrared (NIR). This includes specific coatings, thinning steps, implant steps, thick epitaxial wafers, high resistivity material, etc., (see also Chapter 3).

9.3.5 Conclusion

It is probably fair to state that currently used CMOS sensors in space – typically designed five to ten years ago – may have lower electro-optical performance than scientific CCDs. But their potential lies either in enabling missions and in their specific functionalities like fast shutter, windowed readout, etc., that cannot be realised with CCDs, or by offering other improvements over CCDs. These include radiation hardening, higher frame rates, reduction of power consumption, easier electronics, etc. This means that different applications in the future may need far extended customisation of CMOS imagers or sensor architectures to fully exploit their capabilities. The final systems will definitely outperform CCD-based systems in certain applications. On the other hand, we believe that CCD technology, or CMOS combined with CCD features, will remain the preferred technology in other application areas such as noiseless additions in time delayed integration (TDI) imaging with a large number of stages.

9.4   CMOS sensors for space applications

For a long time, it was expected that high-quality scientific imaging would remain out of reach for CMOS sensors since they lacked the required performance in noise, non-uniformity and dark signal levels and hence could not compete with high-end scientific CCDs offering high quantum

efficiency, a large dynamic range, and special operation modes such as TDI and binning. However, CMOS imaging performs better than expected and will get into domains where even CCDs are not feasible. The following section illustrates several examples of CMOS detectors development in which the author was involved, together with discoveries and products made by other groups. It is not intended to give an exhaustive list of CMOS developments for space, neither in number nor in type of applications, since the amount of groups active in the field and variety in applications is much broader than for CCDs. It should, however, give a good impression of the potential of CMOS sensors for space applications and their advantages compared to CCD technology that is currently still used in the majority of imaging applications.

Because of the large variety in application fields, CMOS sensors and architectures can be categorised from different perspectives. They can either be ranked in terms of accuracy and level of electro-optical requirements, or by their need for specific technologies. In the first case, at the lower range of the application spectrum, there are low to medium accuracy sensors used for monitoring, inspection or visual telemetry. These, generally, do not require difficult-to-meet electro-optical specifications, but can take advantage of system integration or specific interfacing with the spacecraft. An example is the IRIS-2 sensor, developed in the late 1990s by Imec (Ogiers *et al*. 1998). Medium to high accuracy sensors are required for robotics (manoeuvring, docking), spacecraft optical guidance and navigation (AOCS), lander and rover imaging.

For example, such a sensor is used to support the automatic rendezvous and docking between the European automatic transfer vehicle (ATV) and the International Space Station (ISS). In the future, formation flying missions may become a reality in order to improve imaging performance through the use of distributed instruments. Such missions make use of several satellites that need to be kept in defined relative positions and would require very accurate attitude sensors. The relative position could be measured by use of laser diodes illuminating an optical retro-reflecting target that is mounted on the other spacecraft. Such sensor operates similar to a star tracker but with artificial targets instead of stars.

From a technological point of view, these sensors also require system integration for miniaturisation and low power consumption, and also need features like fast frame rates, multiple windowing, and so on. Finally, CMOS sensors will become more and more an alternative for CCDs as satellite main instruments in scientific missions such as earth and planetary observation, space science, spectroscopy, astronomy, etc. This is related to scientific imaging, discussed in Chapter 10. Different instrument architectures are being developed, such as high resolution pushbroom imaging or step and stare imaging. In terms of technological features, this may require TDI,

high resolution snapshot shutter sensors, stitching for very large size detectors, and so on.

Adaptive optics wavefront sensors demand large format arrays with fast readout (Saint-Pé *et al*. 2008). To achieve the high electro-optical requirements, quantum efficiency needs to be optimised through backside illumination, use of dedicated optical filters, or even CMOS readout detectors hybridised to optimised silicon diode arrays or other detector materials. CMOS sensors for astronomy will be large with small pixel pitch (down to 5 μm), for example backside thinned with UV enhanced sensitivity and only few readout noise electrons.

Low light imaging in space is usually realised by special CCDs featuring charge multiplication, but CMOS sensors are also being developed with very low noise, with or without similar electron multiplication or using single-photon avalanche diode (SPAD) technology. A specific example of CMOS benefits in an instrument detector is event detection at pixel or chip level for fast photon counting and ranging or lightning imager. An example can be found in Bréart de Boisanger *et al*. (2011a).

9.4.1 Sensors for AOCS

This is probably the most obvious example where CCD technology, in the last decade, has been replaced by CMOS sensors. Under contract by the European Space Agency (ESA), FillFactory (later Cypress and now On Semi) developed space-qualified STAR250 and STAR1000 sensors targeted towards AOCS applications (see, for example, Boldrini *et al*. (2004)). These were the first commercially available radiation hardened CMOS sensors (Bogaerts 2002). Later HAS (high accuracy star) tracker and LCMS (low cost and mass star) tracker sensors were developed to enable a highly accurate star tracker (Bogaerts *et al*. 2005b). By 'high accuracy', a noise level (from the star tracker) of less than one arcsecond is intended.

In order to reach these accuracies, a large number of pixels are combined with a relatively small sensor field of view (FoV). With the small FoV, there must still be sufficient stars present to be able to perform tracking and attitude determination, hence the sensitivity of the device has been increased and the noise floor – determining the star detection limit – had to be decreased. The pixel pitch is 18 μm. The sensor integrates a 1024 by 1024 3T pixel array, a dual addressable row shift register for rolling shutter operation, programmable offset and gain amplifier and an on-chip 12-bit pipelined analogue-to-digital converter (ADC). The HAS sensor allows for non-destructive readout, and therefore, easier simultaneous tracking of several windows combined with off-chip correlated double sampling (CDS). It supports the read out of 40 windows of 20 × 20 pixels each every 100 milliseconds, each having its own integration time, gain and offset. The HAS

is now the baseline sensor used by most star tracker manufacturers. The aim of the LCMS sensor was to enable the future production of very low cost, low mass and low power star trackers by means of a high degree of analogue and digital co-integration, even if this would possibly imply a reduction in terms of electro-optical performance. The LCMS development was therefore considered as a system-on-chip technology demonstrator, showing the future potential for AOCS applications. In order to achieve these goals the most obvious method was to reduce the number of parts required.

This is achieved by implementing the following digital components/functions on to the active pixel sensor (APS) chip. The detector timing control is fully integrated on-chip. The sensor basically supports two operation modes. In the acquisition phase, full frame images can be read at five frames per second. In tracking mode, up to 20 windows of 20×20 pixels with an update rate of 10 Hz can be taken with CDS using on-chip static random access memory (SRAM). Automatic background signal removal and thresholding (to reduce processor load) is implemented and illustrated in Fig. 9.5. On the right a horizontal cross-section of a captured scene of a bright vertical bar (about 30 pixels wide) in a darker background is shown.

In each part of Fig. 9.5, the curve shows the normal response of the scene with and without the post-processing function disabled. The logic core cells in the sensor were taken from a subset of the commercial libraries, pre-screened for low leakage current under radiation. All key registers have been implemented with regenerative triple redundancy and voting, providing protection against heavy ions. Heavy ions could also lead to soft errors in the SRAM. The data stored in the SRAM was made insensitive to single bit changes by the use of error detection and correction (EDAC) and periodic scrubbing.

Another interesting example of miniaturisation is the Digital Sun Sensor on a Chip (DSSoC development under ESA contract by Selex-Galileo, BAE Systems, and CMOSIS) (Fidanzati *et al*. 2011). Sun sensors are also used in attitude and orbit control and exist in many variants, determining the sun's position relative to the spacecraft frame. The most complex and accurate are digital imaging sensors with 2-D pixel array that calculate the photometric centre of gravity of the sun image to sub-pixel accuracy and at a high update rate. Traditional implementations comprise an image sensor with ADC, field-programmable gate array (FPGA)-based readout processing logic, supply conditioning, latch-up protection, optics, and housing, resulting in a system of roughly 500 cm^3, weighing 0.4 kg and consuming 1.5 W. The DSSoC integrates all radiation-hardened and single-event tolerant electronics on a single 11 mm $\times$ 11 mm die, including 512 $\times$ 512 pixel array, 10 bit ADC, processing logic, SpaceWire interface, supply regulators and clock oscillator.
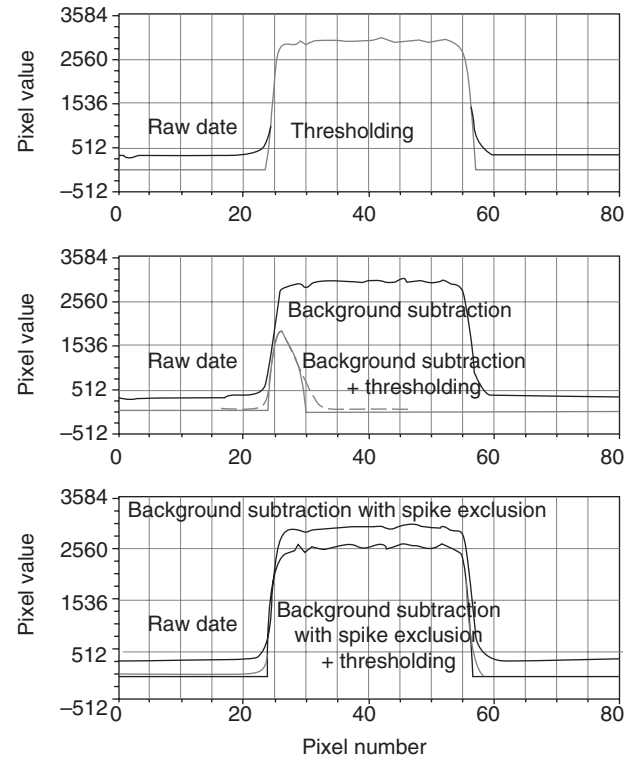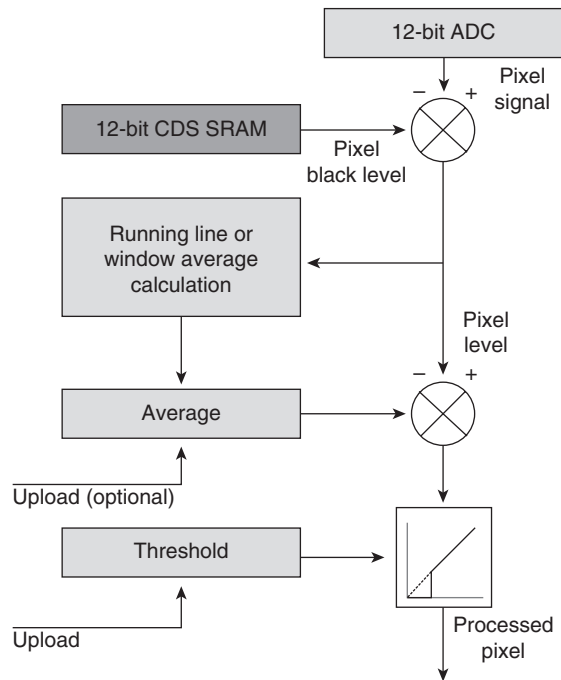
9.5 LCMS signal processing flow (left) and illustration of the different data processing options (right).
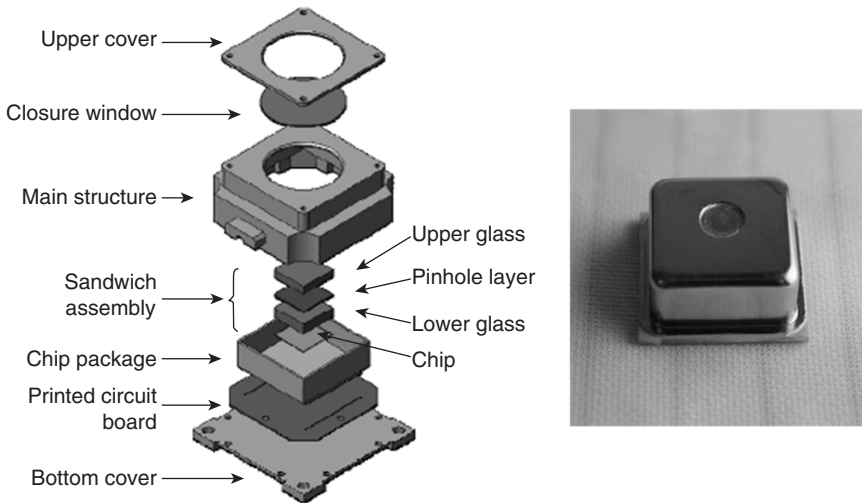
9.6 Digital sun sensor on a chip instrument assembly (left) and mock-up with integrated optics, and hermetic package lid (right). Images courtesy of Selex-Galileo, BAE Systems and CMOSIS.

Further, immersive optics are used: a fused silica block with embedded titanium pin hole is bonded directly on to the silicon die, at the same time offering a wide FoV of 128° (full cone) and radiation shielding. This assembly is then encapsulated in a hermetic enclosure; a handful of passive components and the external housing make up a complete instrument measuring 50 cm³, 65 grams, and 200 mW. However, there is no performance degradation over traditional implementations, with a resolution of 0.002° and a noise equivalent angle of just 10 arcseconds, and this at an update rate of over 60 frames per second. The manifold gains to be realized in terms of cost – of the instrument and of its footprint on the spacecraft – allows its deployment on a wide range of satellites, including nano-sats. Figure 9.6 shows the digital sun sensor on a chip instrument assembly (left) and mock-up with integrated optics, and hermetic package lid (right). A similar development for sun sensors has been performed by TNO and Delft University and described in Leijtens *et al.* (2007).

9.4.2   Imaging for solar physics and astrophysics

PROBA2 (PRoject for Onboard Autonomy) is an ESA micro-satellite, launched in late 2009. It is part of ESA's in-orbit technology demonstration program. Among the platform elements to be technologically demonstrated is a STAR1000 based star tracker. The science payload on-board PROBA2 contains a solar instrument SWAP (Sun Watcher using Active Pixel System

detector and Image Processing) that is based on the HAS sensor. The aim of the instrument is to identify and study all events on the sun that might have implications on the solar-terrestrial connection. In particular, the focus of the PROBA2 mission is the genesis and evolution of events that can affect space weather, such as coronal mass ejections, EUV waves, EUV dimming and solar flares. SWAP provides images of the solar corona at a temperature of roughly one million degrees, with a cadence of one image per one to two minutes, and FoV of 54 arcminute. The HAS detector for this application is coated with a scintillator layer to improve sensitivity in the EUV range, in this case at 17.5 nm, and is passively cooled with an external radiator viewing cold space. SWAP is the first, and at this moment only, solar space telescope using a CMOS sensor. The extreme ultraviolet normal incidence spectrograph (EUNIS) instrument (Daw *et al.* 2011) also uses a CMOS detector, but has flown on a sounding rocket. In early June 2012, the planet Venus partially eclipsed the Sun as seen from Earth. The SWAP instrument took the image shown in Fig. 9.7, where Venus shows up as a black circle that starts sliding in front of the Sun (top left).

PROBA2/SWAP 174 2012-06-05T22:16:10.264
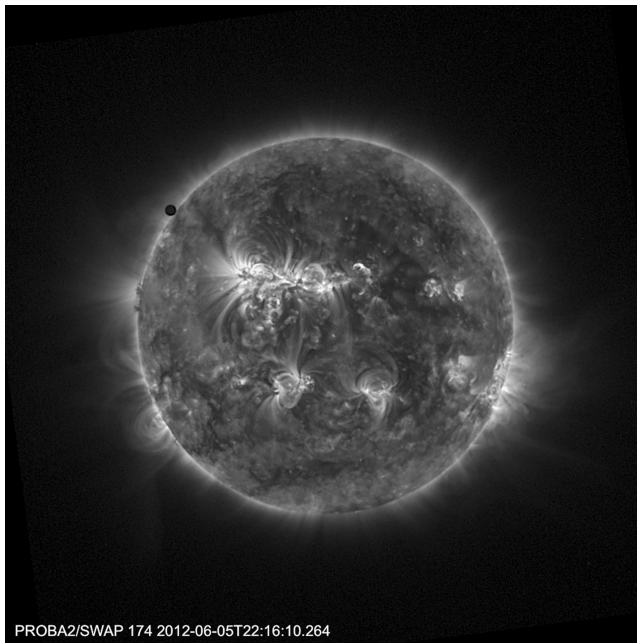
9.7 Image of Venus eclipse taken by SWAP instrument. Acknowledgement: SWAP is a project of the Centre Spatial de Liège and the Royal Observatory of Belgium funded by the Belgian Federal Science Policy Office (BELSPO). More information about SWAP instruments can be found at http://adsabs.harvard.edu/abs/ 2006AdSpR..38.1807B and http://proba2.oma.be/about/SWAP.

More recently, the APSOLUTE (APS optimised for low-noise and ultraviolet tests and experiments) prototype development was finalised (through a partnership between the Royal Observatory of Belgium, Centre Spatial de Liège and CMOSIS) (Benmoussa 2011). These prototypes were used to optimise noise performance (fewer than three to five noise electrons), pixel dynamic range (more than 85 dB) as well as backside-illumination with enhanced EUV sensitivity. Note that back-side illumination (BSI) – contrarily to BSI used for mobile phone sensors, as described in Chapter 4 – is not primarily used to avoid pixel crosstalk and improve angular sensitivity, but rather to increase the quantum efficiency outside the visible wavelength range. Finally, with solar orbiter specifications in mind, this has to led to 2k × 2k or 3k × 3k detectors with 10 μm pixel pitch using stitching technology. These sensors are all rolling shutter devices. They can be compared to the detector development for astrophysics described by Saint-Pé *et al.* (2008).

9.4.3   Hyperspectral imaging

Another development under ESA contract (in partnership between IMEC and FillFactory/Cypress) has been carried out for a high-end CMOS APS sensor, optimised for hyperspectral imaging in space. The goal of such space missions is to observe parameters encompassing agriculture, forestry, soil/geological environments and coastal zones/inland waters. These data can be used to improve our understanding of geospheric processes. The principle of hyperspectral imaging is illustrated in Fig. 9.8. The sensor consists of a 2-D pixel array. While the satellite moves with respect to the earth, the sensor's *x*-axis is used to retrieve the spatial information of a ground line (across-track) represented by the spectrometer slit. The second dimension (*y*-axis) is used for the reproduction of the spectral information. The resulting image data, together with the along-track information given by the motion of the satellite, will then constitute the so-called image cube.

The CMOS sensor combines two major technology developments. The first is backside thinning and hybridisation to achieve high quantum efficiency with optimisation of several post-processing steps, like anti-reflective coating for a broad wavelength range and techniques for cross talk reduction (Bogaerts *et al.* 2005a). The second is the pixel architecture with on-chip in-pixel correlated double sampling for low noise operation, combined with synchronous shutter operation.

The pixel pitch is 22.5 μm in both the *x* and *y* direction. Since the size of the device could be larger than the reticle size, the sensor also required the stitching technique. The pixel array consists of stitch blocks of 512 × 512 pixels and is stitchable up to 2048 × 2048 pixels. Hence, the resolution of the pixel array is not fixed. Columns are multiplexed in groups of 256 to a
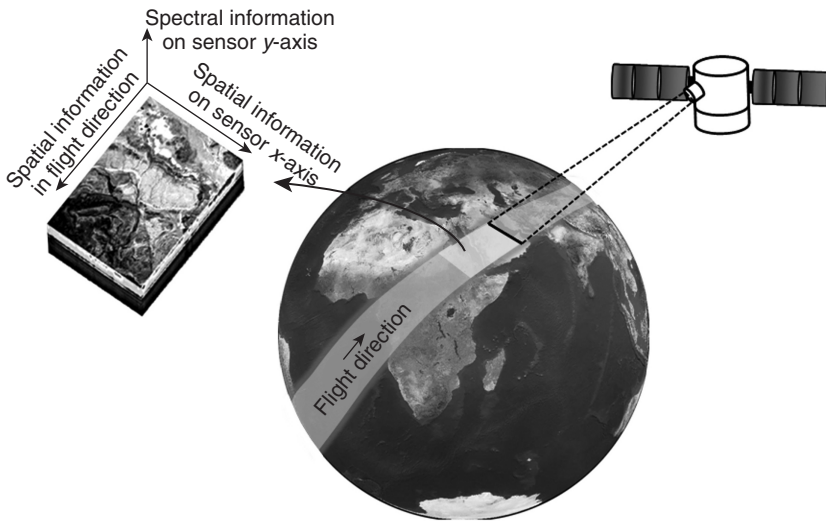
9.8 Principle of hyperspectral imaging.

pseudo differential output. Each output runs at a maximum readout frequency of 20 Mpixels per second such that the 2k × 2k detector could be read out at close to 30 frames per second. The sensor was designed using techniques to achieve high radiation tolerance. The sensor was implemented for use in a hybrid approach but also has the possibility to be used as a monolithic, backside thinned, backside illuminated image sensor on its own. In the hybrid approach, the readout circuit is combined with hybrid silicon diodes processed on wafers that are afterwards thinned, diced and flip-chipped on the readout sensor using 10 µm indium or CuSn bump technology.

The pixel consists of two stages: a light detecting stage and a sample and hold stage with three storage capacitors. This architecture, as shown in Fig. 9.9, allows for a true pipelined synchronous shutter with on-chip correlated double sampling (CDS), i.e. all pixels start and stop integrating at the same moment while the previous frame is still read out.

The pixel readout noise and dynamic range is determined by the size of the storage capacitors. Although the sensor normally operates as a synchronous shutter device, meaning that the integration starts and stops for all pixels at the same moment in time, the pixel architecture allows enough flexibility for the user to define a different integration time for each line individually. Furthermore, in hyperspectral imaging, a large full well is required to achieve the highest possible signal-to-noise ratio (SNR) in bright objects (e.g. clouds) in a particular wavelength range.
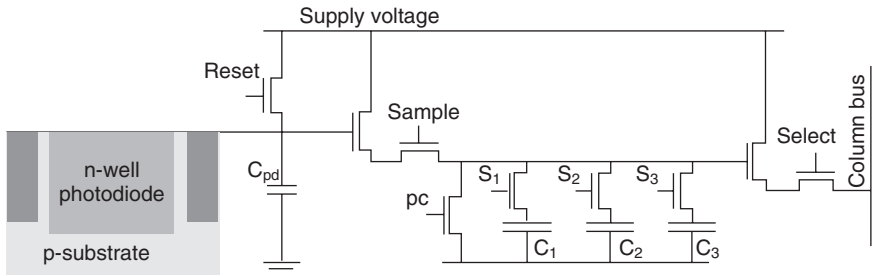
9.9 CMOS snapshot shutter pixel with three in-pixel storage capacitors.

On the other hand, a very low background noise is desirable for imaging of dim objects. Hence there is a need for a large full well and very low noise level. Typically, the two situations do not necessarily occur in the same wavelength range. For example, the lowest background noise is required at lower wavelengths with weak signals, while the highest full well is needed at longer wavelengths. The hybrid approach (in this case, in theory this could also be realised in monolithic sensors) allows the pixel capacitance to be optimised, depending on the wavelength, i.e. the pixel capacitance can vary along the *y* direction in the image array.

This optimisation can be obtained from the spectral response of the array and the expected spectral dynamic range of the irradiance. This is schematically depicted in Fig. 9.10. This is another example of pixel optimisation that is not feasible in CCDs, where all lines are read out with the same full well charge and readout noise. With this technique in CMOS, lower accuracy ADCs can be used while still obtaining a larger intra-scene dynamic range.

Other hyperspectral imagers are described in Dierickx *et al*. (2009), Vu *et al*. (2011) and Wang (2011). The latter sensor is used on a lunar mission (on board Indian satellite, Chandaryaan-1) launched in late 2008. It consists of a pixel array of 256×512 snapshot pixels at $50\,\mu m$ pitch and also features row wise gain compensation. It retrieves hyperspectral information at $80\,m$ ground resolution in the wavelength range of $421\,nm$ to $964\,nm$ in 64 spectral bands.

9.4.4 High resolution Earth and planetary observation

Current observation satellites for Earth and planetary observation and space science (for example, Hubble, XMM, HiRISE, Envisat, Pleiades, Spot, Helios-2, GCOM, etc.) currently all rely on CCD technology. For future missions, CCD technology is considered as the most mature technology, but many space agencies, for example, the French space agency CNES, (Materne 2011) support the development of CMOS sensors as a promising
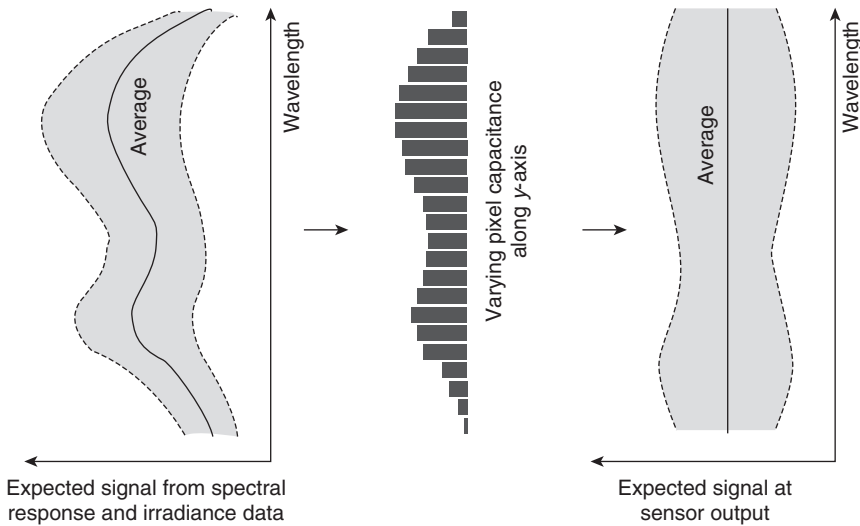
9.10 Principle of hyperspectral imager with varying and optimised pixel capacitance along the vertical direction.

alternative. The very high resolution requirements require very short integration times. Ground sampling distances close to the resolution limits imposed by the Earth's atmosphere are targeted. For example, depending on low Earth orbit (LEO) altitude, distances of 20 to 30 cm up to 1 m correspond to a maximum integration time of a few tens to a few hundreds of microseconds. A 400 km orbit has 7 km/s ground speed, so 1 m corresponds to 140 microseconds.

Such short integration times require architectures that enable very fast readout and accumulation of signals to improve signal-to-noise ratio. Moreover, the irradiance dynamic range can be extremely large. Very faint signals of only a few tens of photons should be measured, while reflecting objects in urban areas can quickly saturate the detector. Hence, besides good noise performance, there is the need for high dynamic range and good detector anti-blooming performance, which is inherently better in CMOS sensors. The general roadmap for detector development for these applications consists of two different types of sensor architectures: TDI and high resolution 2-D snapshot shutter focal planes.

High resolution pushbroom imaging instrument

TDI is a particular imaging mode that has been used for decades, and uses 2-D pixel arrays instead of 1-D arrays, even though the pushbroom principle would bring the spatial information in the flight direction. In a TDI sensor the pixel signals delivered by the pixels of the same column (along-track

direction) are adequately delayed and added synchronously with optical scanning. Thus, the light from a given point in the scene impinges successively on each pixel of the given column and the signals from each of the rows are therefore accumulated to increase the final SNR.

The TDI principle has typically been addressed with CCD sensors where the TDI functionality is more or less intrinsically available by shifting the charge packets along the CCD synchronously with the moving image. This addition process is noise-free in a CCD, but is not so easily implemented in CMOS sensors. Signal summation should take place outside the pixel in the voltage domain. This requires high speed pixel array readout linearly proportional to the number of TDI stages. In order to keep the noise floor low, careful design of the column circuitry is mandatory. Different architectures can be used depending on application requirements (Lepage et al. 2009). Recent developments have demonstrated that CMOS is a viable alternative to CCDs, unless the number of TDI stages becomes too large, e.g., 100 stages depending on reference line time. However, they have the clear advantage of low power consumption and simpler system design. In the future, CMOS technology, combined with CCD-like operation options – allowing a limited number of charge transfers within the pixels – may turn out to be the optimal choice.

In Lepage et al. (2011), the CMOS sensor design has been optimised to take into account final dynamic modulation transfer function (MTF), quantum efficiency and read noise. The pixel pitch is 13 μm in across track direction. The sensor uses a 4T rolling shutter pixel readout with temporal (41 μs instead of 50 μs reference line time) and spatial (10.8 μm) oversampling in the along track direction. This is done to restore the dynamic MTF with only minor degradation of quantum efficiency and noise performance. The addition process in the CMOS sensor can be performed using analogue accumulators, or eventually be made off-chip. Another possibility is to perform analogue-to-digital conversion early in the signal chain and perform the addition in the digital domain (Lepage et al. 2009). In case the required speed and accuracy can be reached in the ADC, this addition process turns out to be much simpler than in the analogue domain.

It also offers some additional advantages. In CCD technology, the dynamic range is determined by the full well capacity of the pixel or the output register, and the readout noise. Bright objects in the scene that saturate the pixel after accumulation in the pixel array but before readout will not generate useful information, i.e. n_{TDI} lines are used in total, but pixel signal saturates after less than n_{TDI} additions. Since the TDI accumulation process does not necessarily happen within the pixel in case of CMOS sensor, the final charge handling capacity can be much larger than the full well charge of a single pixel. This also means that the charge handling capacity of a CMOS pixel in a TDI sensor can be made relatively small. The pixel can
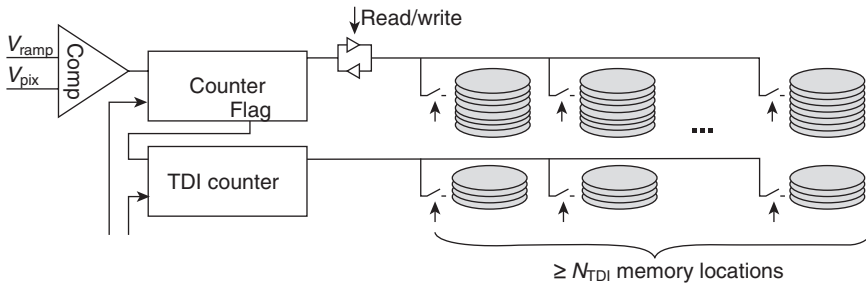
9.11 TDI architecture based on ramp ADC with local counter and dynamically controlled number of additions.

be operated with high conversion gain, which is in general beneficial for its noise performance. The charge handling capacity will in this case be limited by the charge handling capacity of the accumulators. In case of digital addition, the addition process can be controlled dynamically without much complexity. The idea is to check after each addition whether the next addition should take place or not.

This can, for example, be implemented in the column. The principle is illustrated in Fig. 9.11. If the counter value exceeds a certain pre-defined or programmable value, the addition of the next pixel signals is suppressed. At that moment the TDI depth is memorised with no further memory access for subsequent additions. This means that each pixel can have its own optimal TDI level and at the same time use the full dynamic range offered by the accumulator. Due to the need for large charge handling capacity at the output, the output gain of a CCD needs to be low. Note therefore that – although the charge addition in CCD is noise-free – the readout noise of the CCD can be larger than in a CMOS sensor where the read noise increases with the square root of the number of additions, but the read noise of a single reading can be very small due to the high pixel conversion gain. CMOS pixels thus need to be optimised for high gain and low image lag in combination with their relatively large size.

Other CMOS sensor developments for pushbroom imaging in space can be found in Lepage *et al*. (2006), Bréart de Boisanger *et al*. (2011a) with non-TDI line detectors, and Lo *et al*. (2011) with TDI 16/8 stages. Figure 9.12 shows an image of part of the Moon taken by the terrain mapping camera (TMC) on-board Chandrayaan-1 (Wang 2011). This is based on a linear 4000 pixel sensor with 7 μm pitch.

Step and stare imaging instruments

An alternative solution for high resolution observation applications is the use of 2-D pixel arrays (Materne, 2011). Two different strategies can be
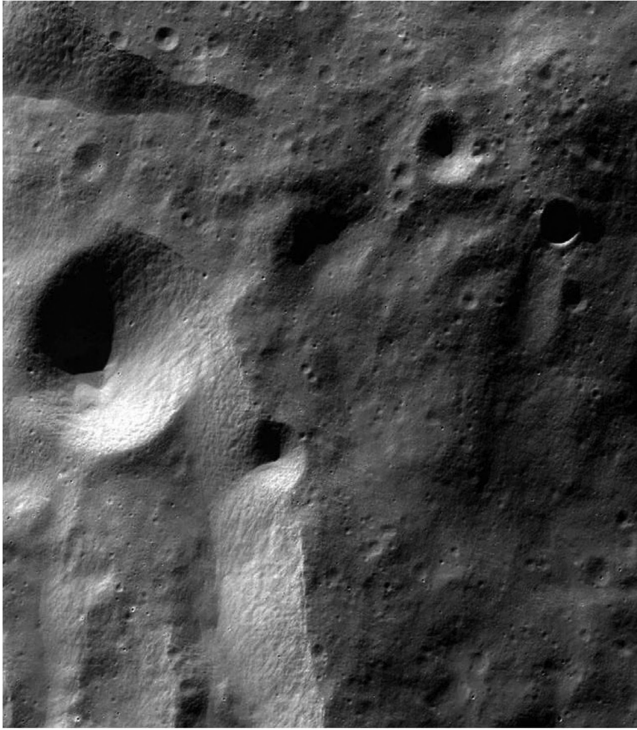
9.12 Moon image from terrain mapping camera (TMC) on board Chandrayaan-1 (100 km orbit). Image courtesy of CMOS Sensor Inc. (Wang 2011).

followed. In case a motion compensating mirror is used, the effective integration time can be prolonged. The advantage is that CMOS sensor development is simplified; however, it results in a more complex system. Due to pointing stability limits, integration times can be increased to about one millisecond without loss of MTF. Alternatively, signals are accumulated off-chip from multiple images with short integration times. Advantageously, geometric corrections due to the satellite pointing axis drift or instability during scanning, are no longer required. However, to achieve the same performance, this requires snapshot shutter imagers with low read noise and high frame readout rates. Shutter efficiency is in such case a key parameter as well. It measures the ability of the sensor to stop integration while still being under illumination and to keep the information stored in the pixel during readout.

The required technology has already been demonstrated for machine vision application (Bogaerts *et al*. 2011). Figure 9.13 shows the pixel architecture and operation. During the frame overhead time (FOT, global
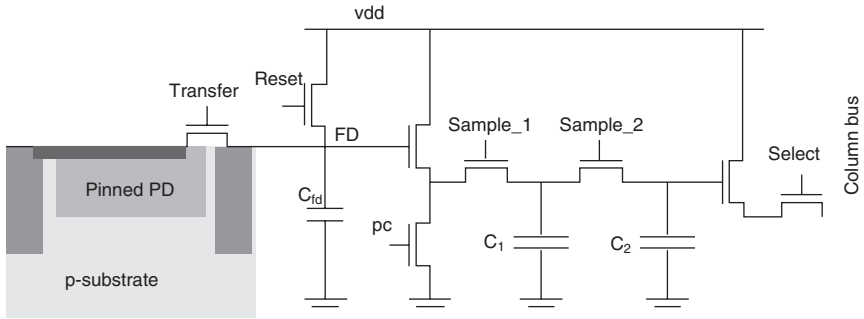
9.13 8T CMOS pipelined snapshot shutter pixel with CDS.

shutter operation), the reset voltage of the floating diffusion (FD) is first sampled for all pixels simultaneously on pixel capacitor C_2. Charges from the pixels are then transferred from the pinned photodiode to FD and the value is sampled on capacitor C_1. This ends the global sampling operation and the next exposure period can start. Meanwhile, the memorised frame is read out row by row by first reading the reset value from C_2 capacitors. The sample_2 switch is then used to obtain charge sharing between the two in-pixel capacitors. Subtraction of both signals in the readout path cancels the pixel kTC reset noise and offset non-uniformity. This pixel has significant advantages. The read noise of this pixel is drastically reduced since the reset noise of the FD is cancelled. The noise of the pixel is mainly determined by the size of the two in-pixel capacitors that are used to store both the reset and the signal level. Thanks to the CDS operation, conversion gain and dynamic range are no longer linked and the pixel can be optimised for high sensitivity – as required for this high speed application – without sacrificing the dynamic range. Shutter efficiency is another important advantage of this pixel architecture. Shutter efficiency has been measured to be better than 99.999%. This is obtained by the low light sensitivity of the gate capacitor, the differential readout operation of the pixel, and by the large capacitance ratio between the FD and storage capacitors. Even for BSI, this pixel architecture exhibits very good shutter efficiency (Meynants *et al.* 2011).

   Other CMOS snapshot shutter sensors have been demonstrated or are used for space applications (Wang, 2011, Lepage *et al.* 2007). Rolling shutter sensors have been proposed for earth observation from geostationary earth orbit (GEO) (Bréart de Boisanger *et al.* 2011b). Note that readout speed in many cases is a key parameter to achieve the required performance. In that respect, CMOS imagers will further benefit from the processing scaling trend, that will in time bring higher speed at lower power consumption.

9.5   References

Bebek CJ, Groom DE, Holland SE, Karcher A, Kolbe WF, Lee J, Levi ME, Palaio NP, Turko BT, Uslenghi MC, Wagner V, Wang G (2002), 'Proton radiation damage in high-resistivity n-type silicon CCDs', *SPIE* 4669, 2002.

Benmoussa A (2011), Developments of back side illuminated CMOS APS for EUV Solar Observations, *2nd Workshop CMOS Image Sensors for High Performance Applications*, CNES, 6–7 December, Toulouse, France.

Bogaerts J (2002), 'Radiation-induced degradation effects in CMOS active pixel sensors and design of radiation-tolerant image sensor', PhD. Thesis.

Bogaerts J, Dierickx B, De Moor P, Sabuncuoglu Tezcan D, De Munck K, Van Hoof C (2005a), 'High-end CMOS active pixel sensor for hyperspectral imaging', *Proc. 2005 IEEE Workshop on CCDs/AISs*, Japan.

Bogaerts J, Ogiers W, Cools T, Diels W (2005b), 'Next generation CMOS active pixel sensors for attitude and orbit control systems', *6th International ESA Conference on Guidance, Navigation and Control Systems*, 17–20 October, Loutraki, Greece.

Bogaerts J, Ruythooren K, Gvozdenovic A, Van Esbroeck K, Ceulemans B, Ogiers W, Arsinte G, Wang X, Meynants G (2011), 'High speed 36 Gbps 12Mpixel global pipelined shutter CMOS image sensor with CDS', *2011 International Image Sensor Workshop*, 8–11 June, Hokkaido, Japan.

Boldrini F, Procopio D, Airy S P, Giulicchi L (2004), 'Miniaturised star tracker (AA-STR) ready to fly', *Proceedings of The 4S Symposium: Small Satellites, Systems and Services (ESA SP-571)*, 20–24 September, La Rochelle, France.

Bréart de Boisanger M, Saint-Pé O, Larnaudie F, Guiry S, Magnan P, Gonthier P M, Corbière F, Rolando S, Goiffon V, Guyatt N, Pike A, Lord B (2011a), 'CMOS devices for ESA LEO and GEO Earth observation operational missions', *2nd Workshop CMOS Image Sensors for High Performance Applications*, CNES, 6–7 December, Toulouse, France.

Bréart de Boisanger M, Saint-Pé O, Larnaudie F, Berdin E, Magnan P, Gonthier P M, Corbière F, Rolando S, Guyatt N, Pratlong J, Jerram P, Biffi J M, Bardoux A (2011b), 'CMOS devices for GEO ocean colour imager', *2nd Workshop CMOS Image Sensors for High Performance Applications*, CNES, 6–7 December, Toulouse, France.

Daw AN, Brosius J, Criscuolo E, Davila J, Haas JP, Hilton G, Linard D, Plummer T, Rabin D, Thomas R, Varney D, Wang T (2011), 'New capabilities of the EUNIS sounding rocket instrument', *Bulletin of the American Astronomical Society*, 43. Available from: http://adsabs.harvard.edu/abs/2011SPD….42.1502D (Accessed 21 June 2012).

Dierickx B, Dupont B, Jerram P, Fryer M, Pratlong J, Walker A, Defernez A (2009), 'CMOS Synchronous shutter backside illuminated image sensor for hyperspectral imaging', *2009 International Image Sensor Workshop*, 26–28 June, Bergen, Norway.

Fidanzati P, Boldrini F, Monnini E, Ogiers W, Pritchard A, Airey SP, Kowaltschek S (2011), 'Miniature digital sun sensor on a chip performance evaluation', *8th International ESA Conference on Guidance, Navigation and Control Systems*, 5–10 June, Karlovy Vary, Czech Republic.

Goiffon V, Magnan P, Saint-Pé O, Bernard F, Rolland G (2008), 'Total dose evaluation of deep submicron CMOS imaging technology through elementary device and

pixel array behavior analysis', *IEEE Transactions on Nuclear Science*, 5 (6), 3494–3501.

Goiffon V, Magnan P, Saint-Pé O, Bernard F, Rolland G (2009a), 'Ionization versus displacement damage effects in proton irradiated CMOS sensors manufactured in deep submicron process', *Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment*, 610 (1), 225–229.

Goiffon V, Hopkinson G R, Magnan P, Bernard F, Rolland G, Saint-Pé O (2009b), 'Multilevel RTS in proton irradiated CMOS image sensors manufactured in a deep submicron technology', *IEEE Transactions on Nuclear Science*, 5 (4), 2132–2141.

Goiffon V, Cervantes P, Virmontois C, Corbière F, Magnan P, Estribeau M (2011), 'Generic radiation hardened photodiode layouts for deep submicron CMOS image sensor processes', *IEEE Transaction on Nuclear Sciences* 58(6), 3076–3084.

Leijtens J, Theuwissen A, Padmakumar R, Wang X, Xie N (2007), 'Active pixel sensors: the sensor of choice for future space applications', *Proceedings of SPIE*, 6744.

Lepage G, Dantes D, Diels W (2006), 'CMOS long linear array for space application', *Proc. SPIE*, 6068.

Lepage G, Materne A, Renard C (2007), 'A CMOS image sensor for Earth observation with high efficiency snapshot shutter', *2007 International Image Sensor Workshop*, 7–10 June, Ogunquit, Maine, USA.

Lepage G, Bogaerts J, Meynants G (2009), 'Time-delay-integration architectures in CMOS image sensors', *IEEE Transactions on Electron Devices*, 56 (11), 2524–2533.

Lepage G, Materne A, Gilbert T (2011), 'Characterization of a digital TDI CMOS demonstrator for high resolution Earth observation', *2nd Workshop CMOS Image Sensors for High Performance Applications*, CNES, 6–7 December, Toulouse, France.

Lo R, Ling J, Wang B (2011), 'Monolithic, 120mm CMOS sensor for multi-mode RSI application', *2nd Workshop CMOS Image Sensors for High Performance Applications*, CNES, 6–7 December, Toulouse, France.

Materne A (2011), 'CNES roadmap for CMOS image sensors in space applications', *2nd Workshop CMOS Image Sensors for High Performance Applications*, CNES, 6–7 December, Toulouse, France.

Mekki J, Moll M, Glaser M, Dusseau L, Saigné F, Roche NJ-H (2010), 'Radiation response of forward biased float zone and magnetic Czochralski silicon detectors of different geometry for 1-MeV neutron equivalent fluence monitoring', *IEEE Transactions on Nuclear Science*, 57 (6), 3483.

Meynants G, Bogaerts J, Wang X, Vanhorebeek G (2011), 'Backside illuminated global shutter CMOS image sensors', *2011 International Image Sensor Workshop*, 8–10 June, Hokkaido, Japan.

Ogiers W, Dierickx B, Scheffer D, Habinc S (1998), 'A single-chip CCSDS packet telemetry and telecommand based microcamera', in *Sensors Update Volume 4* (eds Baltes H, Goepel W, Hesse J), Wiley-VCH. Available from: http://microelectronics.esa.int/vhdl/doc/Iris.pdf (Accessed 30 June 2012).

Padmakumar Rao (2009), 'Charge transfer CMOS image sensors: device and radiation aspects', PhD. Thesis.

Saint-Pé O, Bréart de Boisanger M, Tulet M (2008), 'CIS developments within astrium/ISAE and expected outcomes for astrophysical instrumentation', *Astrophysics Detector Workshop*, 17–20 November, Nice, France.

Stevens E, Komori H, Hung D, Fujita H, Kyan F, Parks C, Shi G, Tivarus C, Wu J (2008), 'Low-crosstalk and low-dark-current CMOS image-sensor technology using a hole-based detector,' *Solid-State Circuits Conference, Digest of Technical Papers*, ISSCC, pp. 60–1.

Vu P, Fowler B, Balicki J, Liu C, Mims S, Do H, Li W, Appelbaum J, Venus H, Schwarzer H, Amigues X, Becker P, Degen K-H, Neidhardt M (2011), 'High-speed BSI CMOS image sensor for space applications with 1.1Me- full well capacity and 28e- rms read noise', *2nd Workshop CMOS Image Sensors for High Performance Applications*, CNES, 6–7 December, Toulouse, France.

Wang B (2011), 'CMOS sensor for lunar mission', *2nd Workshop CMOS Image Sensors for High Performance Applications*, CNES, 6–7 December, Toulouse, France.

10

Complementary metal-oxide-semiconductor (CMOS) sensors for high-performance scientific imaging

R. T U R C H E T T A, Science & Technology Facilities Council, UK

Abstract: This chapter deals with high-performance complementary metal-oxide-semiconductor (CMOS) image sensors. It starts with a reminder of the basic principles and sensor parameters for the detection of charged particles and photons, in particular high-energy photons. It is then followed by two sections, the first one on charged particles and the second on the detection of X-rays. In each section, first the main applications are introduced, then the history and state-of-art of CMOS image sensors in the field are presented. As in many of the applications, large area sensors are needed, some notes on the stitching process are also included. The chapter is concluded with a brief review of future trends as well as some suggestions for further reading.

Key words: CMOS, imaging, charged particles, X-rays, large area sensors, stitching, low-noise.

10.1   Introduction

Since the modern invention of complementary metal-oxide-semiconductor (CMOS) sensors (Mendis *et al.*, 1994) in the early 1990s, their detecting performance has been continuously improving. Parameters like quantum efficiency, dark current and noise have been significantly improved, sometimes by design but often by technology improvement introduced directly by the foundry. The original claim that good CMOS sensors could be designed and manufactured in a standard technology can still be considered valid, although some of the manufacturing steps today existing in 'standard' CMOS image sensors (CIS) processes were not originally present in the flow. It was also obvious since the beginning that CMOS sensors could deliver other advantages, for example in terms of higher integration, then reduced cost and/or added performance, low power consumption and radiation hardness. It was then natural to try to address high-performance, scientific applications with this type of technology.

   As other chapters will focus on low light imaging or space applications, we will here review the progress in the detection of higher energy photons and particles. This chapter will first briefly summarise the principle of

detection of this type of radiation in silicon, and it will then review the development of CMOS sensors for both areas. We will also briefly review the history and recent developments in stitching, a technology step that allows making large-area sensors, up to the full size of a single CMOS wafer. Although stitching is also important for other applications, e.g., to make full-frame sensors for DLSR, large area sensors are often needed for the detection of high-energy radiation because of the lack of any efficient lens. A section on future trends will conclude each of the two main sections dedicated, respectively, to the detection of charged particles and high-energy photons. The chapter will finish with a short review of further reading.

10.2   Detection in silicon

In this section we will briefly summarise the principles of detection in silicon for high-energy radiation. Although the ways these two types of radiation interact with silicon are different, the process can be summarised by the loss of ΔE energy by radiation, this loss going into the creation of N_{eh} electron-hole pair, where the relation between these two quantities is set by the quantity W:

$$N_{eh} = \frac{\Delta E}{W} \qquad [10.1]$$

W is roughly proportional to the bandgap of the material (Klein, 1968), and for silicon is equal to 3.62 eV/pair. The quantity W does not depend on the environmental conditions, like temperature or pressure, and neither depends on the energy of the incoming radiation. This energy requirement is larger than the silicon bandgap as the conservation of momentum requires that some energy transfers to the excitation of vibrations and photons (i.e. heat).

10.2.1 Detection of charged particles

When a charged particle traverses a material, it interacts with it and loses energy through the electromagnetic interaction with the material. For heavy particles or low-energy electrons, the average energy loss $-\left\langle \dfrac{dE}{dx} \right\rangle$ is well described by the Bethe equation (Bethe, 1930). This is given by

$$-\left\langle \frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \qquad [10.2]$$

where e is the charge of the electron, z is the charge of the incident particle, Z and A are respectively, the atomic number and mass of the material, β
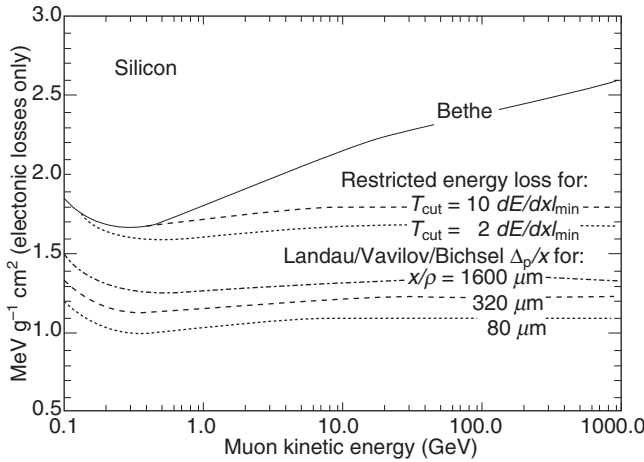
10.1 Bethe energy loss, as well as two examples of restricted energy loss and the Landau most probable energy loss per unit thickness in silicon. The incident particles are muons (Reprinted figure from J. Beringer *et al.*, Phys. Rev. D86, 010001, 2012. Copyright 2012 by the American Physical Society).

and γ are the usual relativistic kinematic quantities, c is the speed of light in vacuum, $m_e c^2$ is the mass of the electron, T_{max} is the maximum energy transfer in a single collision, I is the mean excitation energy, $\delta(\beta\gamma)$ is the density effect correction to ionisation energy loss and $K = 4\pi N_A r_e^2 m_e c^2$ with N_A the Avogadro's number and r_e the classic radius of electrons. Figure 10.1 plots eqn [10.2] for muons in silicon. It also showed corrections to the original Bethe formula, which produce a better agreement with the experimental data.

This formula shows that low-energy particles lose energy vary rapidly. When their energy becomes comparable with their rest mass, the energy loss goes through a minimum. It then increases again due to the so-called density effect. For solid material, this effect is due to the screening action of material atoms over the electric field generated by the incoming particle, thus resulting in an increased energy loss for relativistic particles.

Statistical fluctuations in energy losses are in first approximation described by the so-called Landau distribution (Landau, 1944; Vavilov, 1957). As shown in Fig. 10.2, this distribution is highly skewed and so a good parameter to describe is the most probable energy loss (Bichsel, 1988)

$$\Delta_p = \varepsilon \left[\ln \frac{2mc^2\beta^2\gamma^2}{I} + \ln \frac{\varepsilon}{I} + j - \beta^2 - \delta(\beta\gamma) \right] \qquad [10.3]$$

where $\varepsilon = (K/2)<Z/A>(x/\beta^2)$ MeV for a detector with a thickness x in g cm^{-2} and $j = 0.200$.
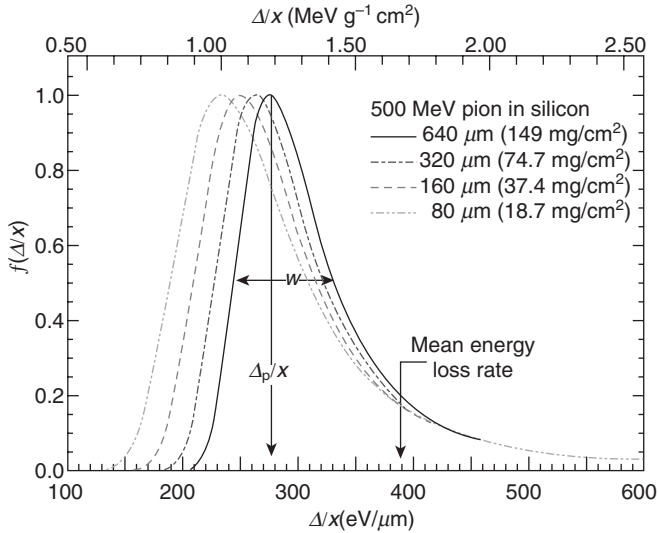
10.2 Energy loss distribution (straggling function) in silicon for 500 MeV pions (Reprinted figure from J. Beringer *et al.*, Phys. Rev. D86, 010001, 2012. Copyright 2012 by the American Physical Society).

The overall energy loss is mainly due to interactions of the impinging particles with the material resulting in small energy losses. However, every now and then a particle could undergo head-on or nearly head-on collisions with electrons, thus transferring a significant amount of energy to a single electron that can then travel into the material thus losing energy. These electrons are also called δ-rays. As they can travel in silicon for a few microns before stopping, they can significantly affect the measurement of the impact point of the particle. High-energy transfer is responsible for the tail in the Landau distribution (Damerell, 1995) and it was shown that this is also linked to lower spatial resolution measurements in the sensor (Colledani *et al.*, 1996).

For thin materials, the Landau solution is not any longer valid as a detailed treatment, which includes the different contributions from atomic orbitals, that have to be considered. The width of the energy loss distribution is larger than predicted by the Landau curve. For example, for a thickness of 32 μm, the width of the distribution is about twice that predicted by the simple Landau theory (Bak *et al.*, 1987). Also the most probable energy loss is smaller than predicted by the original theory. This result is summarised in Fig. 10.3, which shows the most probable energy loss as scaled to the average loss of a minimum ionising particle. This latter quantity is equal to 388 eV/μm and corresponds to the creation of 164 electron-hole pairs per micron. It should be noticed that while eqn [10.1] can be used for the energy
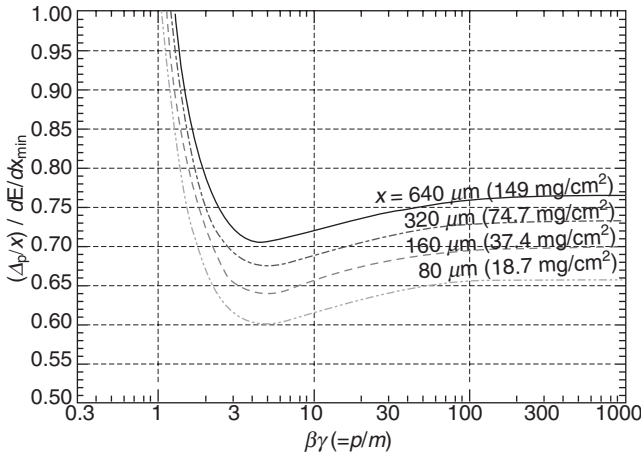
10.3 Most probable energy loss in silicon, scaled to the mean loss of a minimum ionising particle, for different thicknesses of the absorbing material (Reprinted figure from J. Beringer *et al.*, Phys. Rev. D86, 010001, 2012. Copyright 2012 by the American Physical Society).

loss of charged particles in general, at low energy some discrepancies have been measured, indicating a slightly higher quantum yield (Shouleh *et al.*, 1998).

While losing energy, the particle also changes direction as an effect of many small angle scatters (Bethe, 1953). It is common to define a root mean square scattering angle θ_0

$$\theta_0 = \frac{13.6 \, MeV}{\beta c p} z \sqrt{x/X_0} \left[1 + 0.038 \ln (x/X_0)\right] \qquad [10.4]$$

where X_0 is the so-called radiation length of the material. This quantity corresponds to the length over which the electron energy is reduced to $1/e$ of its original energy by *bremsstrahlung* (Beringer *et al.*, 2012), a phenomenon in which the particle is decelerated and the lost energy goes into the generation of photons flying in the same direction as the particle (Heitler, 1949). The radiation length in silicon is equal to 94.461 mm.

Equation [10.4] shows that the scattering is inversely proportional to the momentum of the particle. If the thickness of the sensor is sufficiently important and the energy of the particle is sufficiently low, then the particle could be even reflected back in the material. This is well shown in simulation in McMullan *et al.* (2009c) as well as indirectly in experimental events shown in McMullan *et al.* (2009a).

For high-energy electrons the dominant mechanism for the energy loss is *bremsstrahlung*. The combination of this effect with the electron-hole pair

creation by high-energy photons (see Section 10.2.2) leads to the generation of electromagnetic showers for particles or photons of sufficiently high energy (Perkins, 2000). Starting, for example, from a particle, this can generate high-energy photons through *bremsstrahlung*. These photons can in turn convert back into electron-positron pairs that would then travel through the material losing energy and again creating photons by *bremsstrahlung*. This cascading effect generates showers of particles. It is important for high-energy particles and is used in experiments to measure these high-energy particles through a suitable detector called a calorimeter.

10.2.2 Detection of photons

Only photons with energy higher than 1.1 eV, corresponding to 1,100 nm are absorbed in silicon. As the energy increases, silicon rapidly becomes more effective in stopping photons as shown in Fig. 10.4 by the rapid decrease of the absorption length. This trend is still valid in the visible range, but beyond about 200 nm, the opposite happens. For this wavelength, the absorption is very shallow being of only a few nanometres, but it then increases rapidly. For higher energy, the absorption coefficient would go like E^3, except for energies corresponding to the so-called edges, where the absorption coefficient suddenly drops. This behaviour is shown in Fig. 10.4, which covers the wavelength range between about 0.2 and 1,000 nm, corresponding to photon energies between 1.2 eV and 6.2 keV.

For high-energy photons three different types of interactions need to be considered: photoelectric effect, Compton scattering and pair creation. In
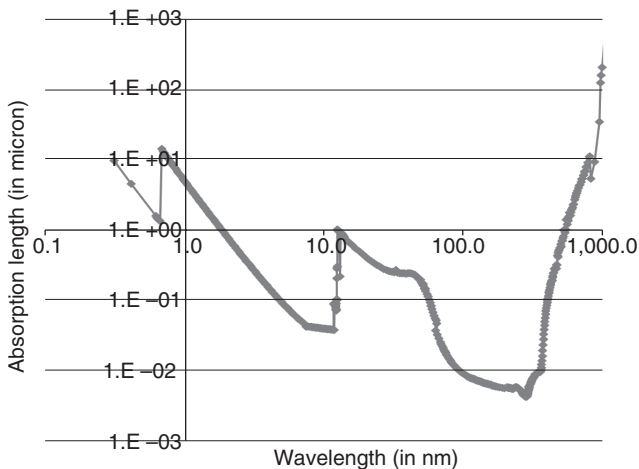
10.4 Absorption coefficient as a function of energy in silicon (adapted from Palik, 1985).

the photoelectric effect the photon is absorbed by one of the electrons in silicon. The electron gains the full energy from the photon and travels into the material releasing its energy. In Compton scattering, the photon loses only part of its energy by inelastic scattering with one of the electrons of the material. Part of the photon energy is given to this electron, resulting in an overall reduction in the energy of the incident photon, as well as a change in its travelling direction. Pair creation can only happen for photons with energy higher than two times the rest energy of an electron or 1.02 MeV. In this case, the photon annihilates generating a pair electron-positron. The linear absorption coefficient relative to the different processes is shown in Fig. 10.5. The cross-section for photoelectric effect varies like $1/E^3$ and the one for Compton effect varies like $1/E$, so that for $E > 10$ MeV the process of pair production, which has a cross-section grossly independent of energy,
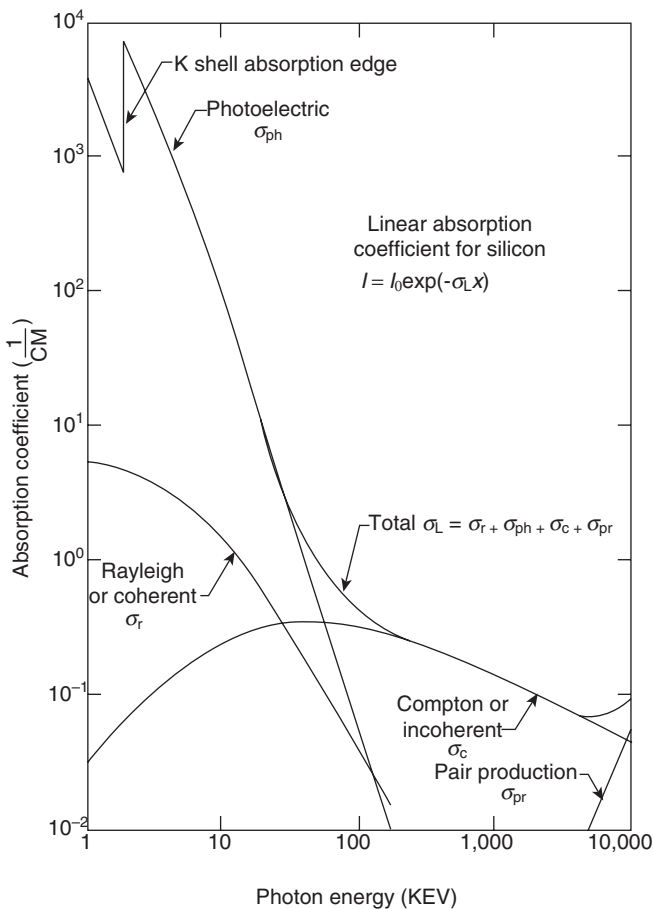
10.5 Linear absorption coefficient in silicon (from Spieler 1998).

becomes dominant. This effect is important for the measurement of high-energy photons and particles because of its role in the creation of showers, as explained in the previous section. Compton effect is used in special gamma cameras as it can provide an electronic way of detecting the travelling direction of the absorbed photons thus getting rid of aperture masks that would decrease the efficiency of detection. Because of its dominance at low energy, photoelectric effect remains the most important effect for the detection of photons.

For wavelengths above about 400 nm, corresponding to an energy of 3.1 eV, only one electron-hole pair is generated by each absorbed photon. Below about 400 nm, the quantum yield η, i.e. the number of electron-hole pairs generated by each incoming photon, increases. The exact dependence of the quantum yield on the wavelength depends on the electric field in the sensor, then it is not a universal constant (Janesick, 2007), but for energies above 10 eV, the quantum yield is given by

$$\eta = \frac{E}{W} \tag{10.5}$$

where W is the same constant that appears in eqn [10.1].

When considering noise sources, it is important to take into account the noise related to the photon beam. In the detection of multiple photons N, photon shot noise is associated to it and its magnitude is given by $\sqrt{N}$. For the detection of a single high-energy photon, fluctuations in the amount of generated charge need to be considered as well. With the number N_{eh} of electron-hole pairs generated by one photon given by eqn [10.1], the variance σ_{eh}^2 of this number is given by

$$\sigma_{eh}^2 = F * N_{eh}$$

where F is the so-called Fano factor (Fano, 1947). In all materials, F tends to be smaller than 1 and is equal to 0.115 in silicon. It is less than one because there are some correlations in the generation of charge. It is worth pointing out that, as the energy of the photons increases the photon shot noise as measured in electron-hole pairs increases, thus making the relative importance of the intrinsic electronic noise less prominent.

10.2.3 Indirect detection of particles and photons

Because of silicon absorption properties (see Fig. 10.4) and the availability of suitable substrates in CMOS (see Section 10.2.4), efficient direct detection of photons in silicon is possible only in a relative limited wavelength range, although this range does cover visible light. In the UV range, photons are absorbed so quickly that basically any material, including passivation material, positioned inbetween the detecting area and the photon beam

would significantly affect the quantum efficiency of the sensor. One option is to use a scintillating material like lumogen, whose absorption coefficient is in excess of $10^5 \, cm^{-1}$, corresponding to an absorption length of less than 100 nm, for most wavelengths between 200 nm and its absorption edge at 460 nm. Upon absorption of UV radiation, lumogen emits visible light in the range between 500 to 650 nm. Lumogen makes possible using front-illuminated sensors for the detection of UV radiation and also provides some protection against damage from UV radiation.

As silicon becomes quickly fairly transparent for increasing energy of photons, scintillators are commonly used, for example, for the detection of X-rays, where they provide a suitable stopping power. The most commonly used material is CsI (Zhao *et al.*, 2004), often doped with a heavy material like tallium and structured to provide some channelling to the light, thus improving the spatial resolution (Arvanitis *et al.*, 2009). In some cases a fibreoptic plate (FOP) can also be added to the scintillator to protect the silicon against the absorption of high-energy photons and then radiation damage.

Indirect detection is also used in the detection of particles. This is shown in Tietz *et al.*, (2008) for the indirect detection of electrons. The detection of neutrons can also be achieved by using scintillating material (Bollinger *et al.*, 1959; Bollinger *et al.*, 1962; Knoll *et al.*, 1988), or suitable materials converting electrons into charged particles (Aoyama *et al.*, 1992; Maneuski *et al.*, 2008).

10.2.4 Complementary metal-oxide-semiconductor (CMOS) substrates

Most of the discussions are valid for silicon in general and do not take into account the specific substrates used for CMOS image sensors. The most common substrate to be used is epitaxial, typically with a resistivity of a few tens of ohm cm and a thickness of a few microns. The chosen thickness matches well the absorption of light in silicon. For longer wavelengths or for X-rays, thicker substrates need to be used. Because of the low voltages typical in CMOS and the resistivity used, higher resistivities have to be considered. Availability of uncontaminated reactors and the economy of the process make it difficult to achieve controlled resistivity in excess of 1 kOhm cm as well as thickness beyond 20 μm. In order to go beyond this limit, silicon-on-insulator (SOI) substrates can also be considered. SOI material was used experimentally for the manufacturing of sensors with 100% fill factor for visible light applications (Wrigley *et al.*, 2000; Pain 2009; Pain and Zheng, 2002) and can also be used for backside illumination (Pain 2009, Edelstein *et al.*, 2011) as the buried oxide can provide a convenient stopping layer for the etching process. SOI materials are also being explored

for the detection of charged particles and X-ray detectors as the handle wafer can be thick thus increasing the signal or the overall detection efficiency as will be shown in Section 10.3.4. For shallow penetrating radiation, e.g., UV or low-energy electrons, substrates need to be backthinned for backside illumination, as shown in Chapter 4.

10.3 Complementary metal-oxide-semiconductor (CMOS) sensors for the detection of charged particles

10.3.1 Particle physics

Charged particles are used in a number of applications, both scientific and industrial. Firmly embedded in the realm of 'big science' is particle physics, thanks to the need for complex and costly accelerators and instrumentation, including detectors. Most particle physics experiments are based around accelerators although cosmic rays are also still used, mainly in underground experiments. For experiments using accelerators a further division is between fixed target and collider experiments. In a fixed target experiment (Fig. 10.6), high-energy beams are directed against a target made of a suitable material. An elementary collision occurs between the particle in the primary beam and the protons or neutrons of the material and the particles generated in this collision are then detected. In a collider experiment (Fig. 10.7), two beams of particles, one positively and the other negatively charged, often particle and antiparticle, are made to circulate in a circular accelerator and then steered against each other in selected collision areas. Collisions between the elementary particles forming the two beams take place and the resulting particles need to be detected. In both fixed target and collider experiments, the experimental apparatus is composed of a multitude of detectors, with different functionalities. Closer to the interaction point, are generally detectors that measure the trajectory of the particles. They tend to be further divided into vertex and tracker detectors, with the vertex being closer to the interaction point. Both vertex and tracker detectors need to have a good spatial accuracy in measuring the impact point of the particle as well as being thin in order to disturb the particle's trajectory as little as possible. At present the most common solution is with hybrid silicon detectors, either pixels or microstrips (Turala, 2005; Hartmann, 2012), but this is an area where CMOS image sensors are emerging as will be explained later. Continuing in our description of a standard particle physics experiment, further away from the interaction point, other types of detectors can be found, used for example for helping in identifying the particle or for measuring its energy. These latter detectors are called calorimeters and work on the principle of a shower as described above.
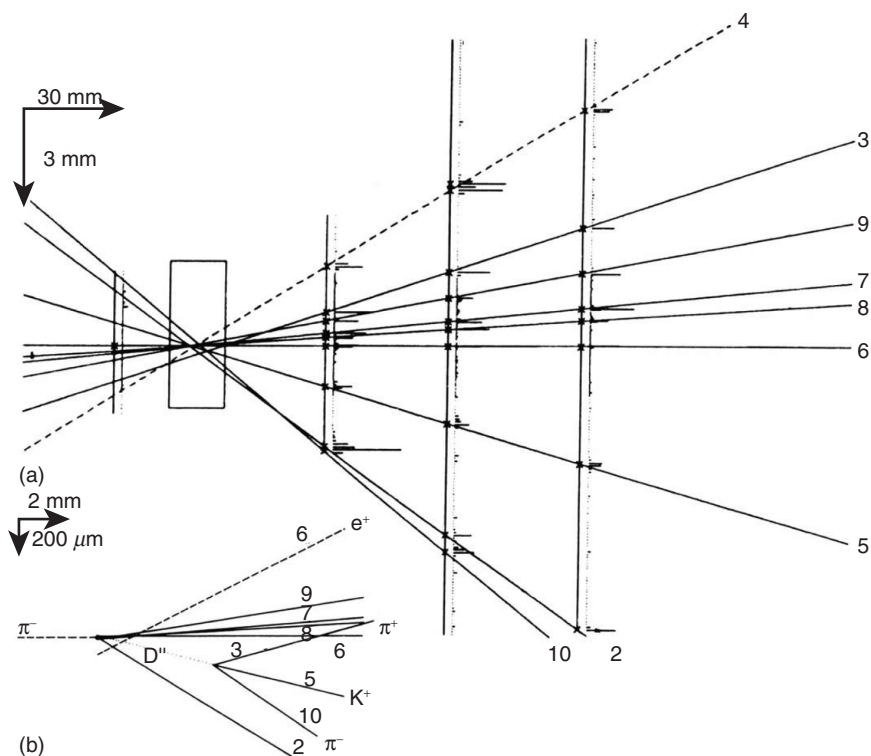
10.6 Event display of an interaction in Be target of NA11 fixed target experiment at CERN, showing tracking and vertexing capabilities (two planes of detectors in front and six beyond the target are marked; horizontal lines attached to these planes represent amplitudes of recorded signals) (Belau *et al.*, 1983; Bailey *et al.*, 1985).

10.3.2 Electron microscopy

Two large categories of electron microscopes exist: transmission (TEM) and scanning electron microscopes (SEM). The typical geometry of a transmission electron microscope is shown in Fig. 10.8. A beam of electrons is accelerated and made to traverse a sample. By using a combination of electric fields to focus the beam, an image of the sample is formed over a focal plane. Up until recently, detectors used in this field were either films or scintillators coupled to a charged coupled device (CCD) (Faruqi and Henderson, 2007). Electrons are normally accelerated to an energy of a few 100 keVs, up to 1 MeV for some microscopes. The optical properties of the system can be conveniently studied by considering that electrons can also be described as waves according to the particle-wave duality principle. The equivalent wavelength λ of an electron of energy E is $\lambda = hc/E$, with h equal to the Plank's constant and c to the speed of light. Although the beam is essentially

CMS detector
Total weight: 14,000 tonnes
Overall diameter: 15.0 m
Overall length: 28.7 m
Magnetic field: 3.8 T

Steel return yoke
12,500 tonnes

Silicon trackers
Pixel (100 x 150μm) ~16m^2 ~66M channels
Microstrips (80 x 180μm) ~200m ~9.6M channels

Superconducting solenoid
Niobium titanium coil carrying ~18,000A

Muon chambers
Barrel: 250 drift tube, 480 resistive plate chambers
Endcaps: 468 cathode strip, 432 resistive plate chambers

Preshower
Silicon strips ~16m^2~137,000 channels

Forward calorimeter
Steel and quartz fibres ~2,000 channels

Crystal
electromagnetic
calorimeter (ECAL)
~76,000 scintillating PbWO$_4$
crystals

Hadron calorimeter (HCAL)
Brass and plastic scintillator ~7,000
channels
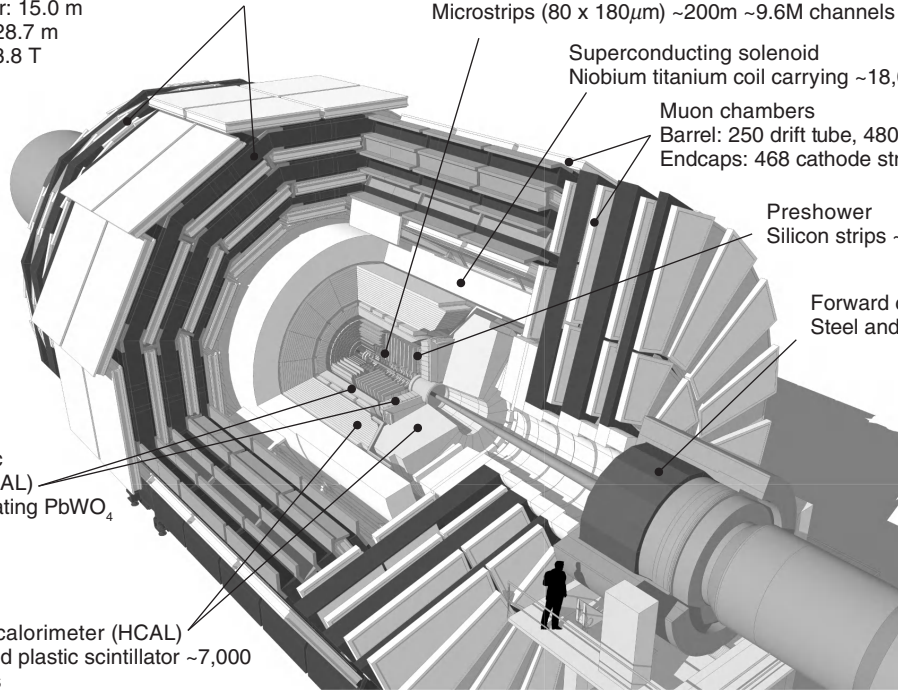
10.7 The CMS experiment at the Large Hadron Collider (LHC) in CERN (this image was produced by CMS collaborator Tai Sakuma and is © CERN, for the benefit of the CMS Collaboration.).
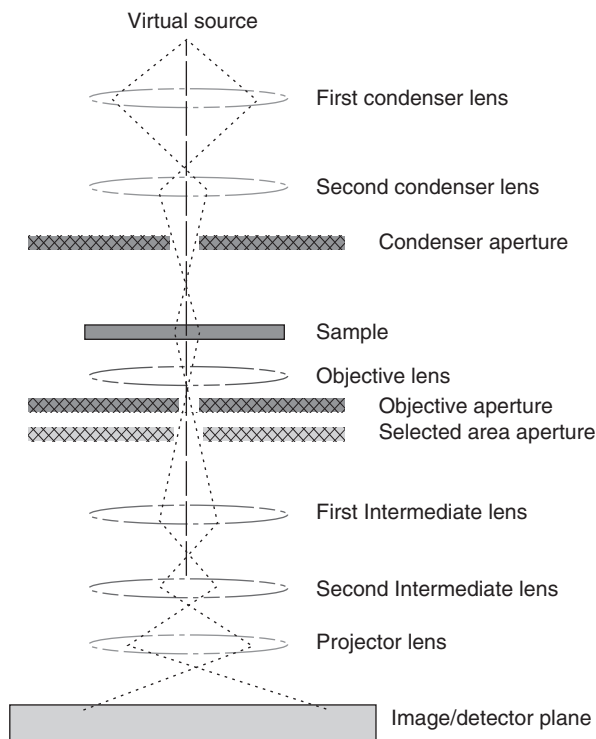
10.8 Schematic diagram of a transmission electron microscope (TEM).

monochromatic, a small spread in the energy, hence in the wavelength, of the electrons exists so that lenses for correcting chromatic aberration need to be used. The equivalent wavelength of electrons is much smaller than that of visible light, thus allowing imaging of very fine structures, like viruses. TEM is also used for diffraction studies, much in the same way as X-rays. Examples of diffraction studies are found in (Henderson and Unwin, 1975).

The first TEM was built by Max Knoll and Ernst Ruska in 1931 (Ruska and Knoll, 1932; Ruska 1987), with this group developing the first TEM with resolving power greater than that of light in 1933 and the first commercial TEM in 1939. TEM is mainly used for biology and material science. In biology, beams are low intensity as they would otherwise damage the sample to be imaged. Higher intensities can be used for material science studies.

In SEM, lower energy electrons are used. Normally the beam is scanned over the surface of a material and transmitted or backscattered electrons are detected, normally by a single point detector. Several SEM techniques exist and SEM apparatus can normally only cope with a number of them. The first SEM image was obtained by Max Knoll, who in 1935 obtained an

image of silicon steel showing electron channelling contrast (Knoll, 1935). The SEM was further developed by Professor Sir Charles Oatley and his postgraduate student Gary Stewart and was first marketed in 1965 by the Cambridge Scientific Instrument Company as the 'Stereoscan'.

10.3.3 Other applications

Mass spectroscopy (MS) is an analytical technique to determine the charge-to-ratio e/m ratio of an ion. In the basic configuration, a constant electric field is used to accelerate the ions and a magnetic field perpendicular to their trajectory is used to bend their trajectory with a radius which is proportional to their e/m ratio. The position of arrival of the ions on a sensor would be used to measure their e/m ratio. Linear sensors can be used in this case. Time-of-flight mass spectroscopy can also be performed. In this apparatus, no magnetic field is needed and the time of arrival of the ion on a detector is measured. As this time is inversely proportional to the square root of the e/m ratio, this quantity can then be extracted. There are also imaging techniques for mass spectroscopy like SIMS (secondary ion mass spectrometry) or MALDI (matrix-assisted laser desorption/ionisation) which are used respectively, for the analysis of solid state material or biological samples. Imaging time-of-flight mass spectroscopy requires a pixellated detector which is also able to record the arrival time of the ions with high resolution. Early experiments in this field were performed using framing cameras (Brouard *et al.*, 2008), but recently CMOS image sensors have been used as explained below. Hybrid pixel detectors have also been explored for this application (Granja *et al.*, 2009).

Thin tissue autoradiography is an imaging modality where *ex-vivo* tissue sections are placed in direct contact with autoradiographic film. These tissue sections contain a radiolabelled ligand bound to a specific biomolecule under study. This radioligand emits beta– or beta+ particles. High spatial resolution autoradiograms can be obtained using low-energy radioisotopes, such as ^{3}H. Film is still the dominant technique but digital alternatives, including CMOS sensors are being considered. Silicon-based imaging technologies have demonstrated higher sensitivity compared to conventional film. CMOS images sensors are also being explored and promising results obtained, showing high efficiency and good spatial resolution (Cabello and Wells, 2010).

Particles, together with photons, are also used for cancer treatment. The radiation absorbed by the cells is effective in stopping the development of cancers. While photons absorption is exponential, particles absorption has a characteristic Bragg peak. It is then possible to reduce damage to nearby healthy cells by the appropriate use of beams of particles. Nowadays proton therapy is the dominant particle therapy. Imaging detectors are required as

there is a need to understand the exact structure of the beam in order to optimise the dose delivered to the patient.

10.3.4 History

The history of CMOS sensors for the detection of charged particles can be dated back to the 1980s (Heijne *et al.*, 1988). At that time, silicon detectors in particle physics were mainly based around high-resistivity silicon microstrips wire-bonded to a readout application specific integrated circuit (ASIC). Thanks to the continuous scaling of microelectronics, it was then possible to think about integrating conventional signal conditioning electronics into a pixel of a size useful for a particle physics experiment, of the order of a couple of hundreds of microns. Since the beginning, two types of approaches were envisaged: hybrid and monolithic. With the estimated noise performance of the electronics, a thick substrate of high-resistivity silicon was needed in both cases and, for the monolithic approach, the use of SOI was then proposed (Pengg 1996; Dierickx *et al.*, 1993).

A group at IMEC in Belgium (Vanstraelen *et al.*, 1988; Vanstraelen *et al.*, 1989; Vanstraelen *et al.*, 1991) also developed fully functional metal-oxide-semiconductor (MOS) transistors on high-resistivity material, obtaining low leakage currents in the pin diodes on the same substrate by having a clean process and relatively low processing temperatures. However, in the monolithic detector that was built afterwards, the readout circuitry was prone to catch some of the charge generated by the incident radiation, causing regions within the active area of the detector to become insensitive. A group in LBL was able to build good pin diodes and MOS devices on the same chip (Holland, 1989a; Holland, 1989b). They also integrated an amplifier on to the detector wafer, but only a single pixel element, rather than an array (Holland and Spieler, 1990).

After the initial two-prong, hybrid and monolithic, pixel approach, the CERN group focused on the development of hybrid sensors. The first hybrid pixel detector was installed in the Delphi experiment at the Large Electron Positron (LEP) Collider in CERN (Delpierre *et al.*, 1994; Becksa *et al.*, 1997; Becksa *et al.*, 1998a, b). Today, most experiments at the Large Hadron Collider (LHC) in CERN have hybrid pixel detectors (Aad *et al.*, 2008; Kotlinski, 2009; Kluge *et al.*, 2007). This technology is now considered well-established for this application; for a recent review article, which also covers other pixel technologies for particle physics, we recommend Wermes (2009).

An American group, based in Stanford and Hawaii University, started developing the monolithic approach, using high-resistivity material as a substrate (Parker, 1989; Snoeys and Parker, 1995). The substrate was of P-type, with an N-implant on the backside to create the necessary pn junction (Fig. 10.9). On the front-side, N- and P- wells were implemented

Not to scale
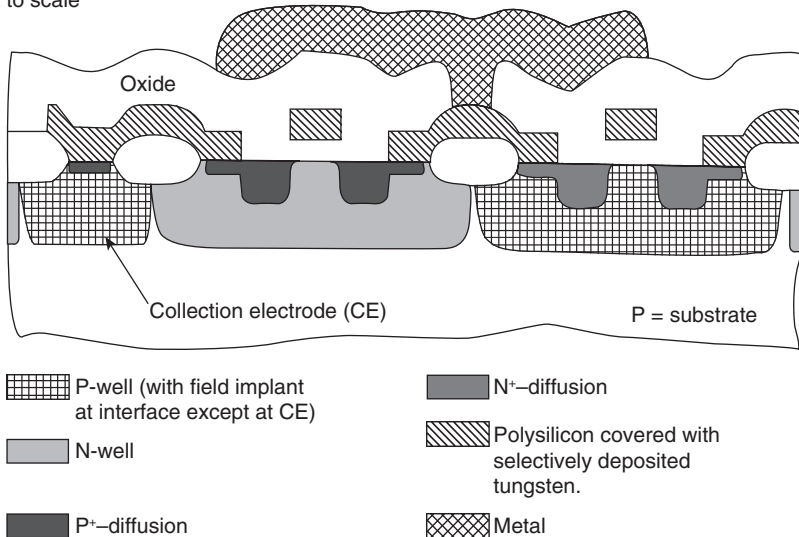
10.9 Schematic cross-section of the process used on the front side of the wafer (Snoeys, 1992). As shown in the figure, the process had both P- and N-well and a single poly and single metal level for routing. The substrate was of P type, doped at approximately $10^{12}\,cm^{-3}$. P+ and N+-diffusion areas are used to create the source/drain implants of the MOS transistors.

in order to obtain well-behaved transistors. The collecting electrode was also of a P-type. The circuitry included an analogue front-end and logic for managing the particle hits. A sensor of a few mm size with rectangular $34\,\mu m \times 125\,\mu m$ pixels was manufactured and successfully demonstrated. Good noise performance was obtained and the sensor was also tested in a beam test where it obtained a spatial resolution of $2.0\,\mu m$ (Kenney *et al.*, 1993; Snoeys *et al.*, 1993). Despite these good results, the development of the devices was not continued, possibly due to the difficulty in obtaining larger devices with good yield.

It is only after Turchetta *et al.* (2001) that the monolithic approach came back with renovated vigour. The proposal was to use standard CMOS technology and collect the charge from the epitaxial substrate with an N-well to p-epi diode. As already pointed out by Dierickx *et al.* (1997), the difference in doping concentration between on one side the heavily doped P-substrate or the P-wells and, on the other side, the lowly doped P-epitaxial layer creates a potential barrier which, although small, is sufficiently high to maintain the radiation-generated electrons within the P-epitaxial layer. Eventually the electrons are collected by the anode. Given the fact that high-energy particles easily traverse metal layers, this means that 100%

efficiency can be achieved in the detection of charged particles (Deptuch *et al.*, 2001). A small device of 64×64 pixels at $25\,\mu m$ pitch, called MIMOSA, meaning minimum ionising particles MOS array, was used to demonstrate the concept. The first results showed it could simultaneously achieve high detection efficiency and high spatial resolution (Deptuch *et al.*, 2000).

For many, it was immediately evident that this monolithic approach could provide a valid solution for pixel detectors for particle physics. The Strasbourg group continued developing the MIMOSA family with a series of different devices, mostly test structures with different architectures for the pixel and the readout circuitry (Hu-Guo *et al.*, 2009). Among the results achieved, a continuous reset structure was designed (Degerli *et al.*, 2005). It makes use of a diode to provide a continuous reset path for the collecting diode. The time constant associated with this reset path is long enough so that the reset does not affect the measurement of the charge when a particle hits a pixel but it is also short enough not to affect the efficiency of the sensor. Double correlated sampling is used and noise performance as low as a few e⁻ rms is achieved. This architecture would not be suitable for standard imaging applications but it is effective in applications where the pixel is hit relatively infrequently, so that most of the time it only has the dark signal. The time constant associated with the continuous reset is then chosen to be long enough not to disturb the signal generated by a particle but short enough to fully reset the diode before another particle arrives. A full size device (MIMOSA-26) was designed for a so-called beam telescope application. A beam telescope comprises a set of detectors arranged along the axis of a beam of particles. By measuring the impact point of the particles, their trajectory can be reconstructed. The sensor format is 1152 by 576 with pixels disposed on a $18.4\,\mu m$ pitch (Baudot *et al.*, 2009). Correlated double sampling is implemented inside each pixel and the sensor works in rolling shutter mode. At the periphery of the pixel array, an offset compensated discriminator compares the signal with an adjustable threshold to select pixels with a hit. The corresponding logic circuitry takes care of selecting only hit pixels. In this way a compression factor between 10 and 10,000 can be achieved, depending on the occupancy levels. The sensor achieves a frame period of $112\,\mu s$.

This sensor further evolved in the ULTIMATE, also known as MIMOSA28, sensor for the STAR pixel upgrade (Besson *et al.*, 2011; Greinera *et al.*, 2011). This upgrade, planned to start taking data in 2014 at RHIC in Brookhaven, would be the first example of a monolithic CMOS sensor being used in a particle physics experiment. Like MIMOSA26, the ULTIMATE sensor is designed and manufactured in the AMS $0.35\,\mu m$ 'Opto' process. It features 928 (rows) $\times$ 960 (columns) pixels, $20.7\,\mu m$ pitch, for a total area of about $20 \times 23\,mm^2$. The epitaxial layer has a thickness of $15\,\mu m$ and a relatively high resistivity of 400 ohm cm. A fast binary readout

and zero suppression are also features of this sensor, whose readout time can be as short as 200 μs. Because of the targeted radiation tolerance of ~ 150 kRad and a few 10^{12} n_{eq}/cm^2/year, radiation tolerant design techniques were used for this sensor.

While STAR at RHIC is the first experiment to use CMOS monolithic active pixel sensors, other experiments are developing them. For example, for the upgrade of the Alice Inner Tracking System at LHC in CERN, CMOS monolithic active pixel sensors have been selected (Musa *et al.*, 2012).

While CMOS sensors are reaching their maturity and starting to be used in real experiments, they still suffer from limitations compared to the more mature technology of hybrid pixel sensors. In these sensors, the detector substrate is different from the electronic readout circuit and connected to it through bump-bonding. This modular approach allows separate optimisation of the detector and the electronics. More complex electronics can be used to condition the signal at high-speed and with good noise performance and digitise the analogue information locally in the pixel. In most CMOS image sensors, the detecting junction is formed by an N-doped area, e.g. N well, in the P-doped epitaxial substrate. N-wells in which p-type metal-oxide-semiconductor (PMOS) transistors sit can also collect the radiation-induced charge and this effect results in a severe loss of charge collection efficiency (Ballin *et al.*, 2008). In order to isolate these unrelated N-wells from the substrate, additional structures need to be introduced (Turchetta *et al.*, 2011). The two main approaches so far adopted are:

1. the use of a deep P-implant to form a deep P-well and
2. the use of SOI.

The first approach was introduced in Crooks *et al.* (2007) and Stanitzki *et al.* (2007). This prototype sensor was designed for an electromagnetic calorimeter for the International Linear Collider (Augustin 2004; ILD 2010). Each pixel is 50 μm square and consists of four diodes connected in parallel, a charge amplifier followed by a shaper with a peaking time of about 120 ns. The analogue signal is discriminated by a two-stage comparator and, on the detection of a hit, a 14-bit time code corresponding to the moment when the hit is detected is stored in a memory, sitting at the periphery of an area with 48 × 92 pixels. The pixel also includes logic circuitry to reset the pixel once a hit is detected, a six-bit DAC to locally adjust the threshold applied to the comparator, thus overcoming the unavoidable threshold spread, a seven-bit memory to store the six-bit value for the DAC as well as a pixel enable flag. In total, there are over 160 active devices, mainly transistors, in each pixel. The prototype sensor worked according to specifications, achieving a noise performance of 22 e$^-$ rms.

Because of the overall status of the ILC project, no further development of this sensor happened, but the concept of using a deep P-well to isolate PMOS transistors found further applications in the PImMS (pixel imaging for mass spectroscopy) sensor for mass spectroscopy (Nomerotski *et al.*, 2010). The pixel architecture is similar to the previous example, but in this case each pixel also includes four 14-bit memory cells that store the time codes of the hits. Because of this addition, the pixel size was increased to 70 μm and the number of active devices increased to over 600. A first prototype with 72×72 cells has already been demonstrated (John *et al.*, 2012) and a reticle-size sensor, consisting of 384×384 pixels has already been manufactured, is working well and is currently being characterised (Sedgwick *et al.*, 2012). All the sensors mentioned here and featuring a deep P-well for protection were manufactured in the180 nm CMOS image sensor process from TowerJazz Semiconductors.

As mentioned earlier, the use of SOI substrates to build monolithic CMOS sensors was one of the two approaches originally proposed by the CERN group in the 1980s. However, this approach was abandoned to follow the hybrid route. It was only about twenty years later, after standard CMOS was proposed for monolithic sensors and in order to overcome the substrate limitations mentioned above, that SOI technology was again considered for particle physics (Arai *et al.*, 2006; Ikeda *et al.*, 2007). In this approach 150 nm SOI process from Lapis (formerly OKI) was modified to provide connections between the electronic layer and the handle wafer. For the latter, high resistivity was used in order to achieve thick depleted layers, thus increasing the signal generated from traversing particles. This development was slowed down by the discovery of the so-called backgate effect. Under irradiation, positive fixed charge is left in the buried oxide and can perturb the behaviour of the transistors. By using a deep P implant (Kochiyama *et al.*, 2011), this effect could be cured and several sensors are now in development, both for particle and X-ray detection (Hara *et al.*, 2010; Ryu *et al.*, 2011; Ono *et al.*, 2013).

In particle physics, each single particle needs to be detected and some information about it, like hit position or time, retrieved. However in transmission electron microscopy useful information can already be gathered by conventional, integrating imaging. CCDs have long been used coupled to a phosphor or scintillator to provide digital images. The spread of the light coming from the scintillator limits the spatial resolution so that for high spatial resolution imaging, conventional films are used (Faruqi and McMullan, 2011). Direct detection of electrons in a CMOS sensor was first demonstrated by (Faruqi *et al.*, 2005; Milazzo *et al.*, 2005). It was immediately evident that CMOS sensors could provide single electron sensitivity as well as high spatial resolution. By thinning the substrates in order to limit the electron backscattering, the spatial resolution could even be better than film

(McMullan *et al*., 2009a). The first commercial camera with a CMOS sensor working in direct detection was launched in 2009 (FEI, 2009). The sensor features 4096 × 4096 pixels at 14 µm pitch and it is capable of 40 frames per second. It was designed with enclosed geometry transistors and its radiation resistance is about 20 MRad (Guerrini *et al*., 2011a, b). An American team also developed a CMOS active pixel sensor for the same application (Battaglia *et al*., 2010). In their design, smaller 5 µm pixels are integrated in a 4k × 4k sensor (Contarato *et al*., 2011). This sensor works at 400 fps (Gatan, 2012). Although most images are static, there are advantages in developing a fast sensor. The first advantage is in the reduction of the leakage-induced signal, which effectively means enhanced radiation hardness. The second advantage is that at low dose the resulting frame would consist only of hits from individual electrons. This information can then be processed to achieve even better spatial resolution (McMullan *et al*., 2009b).

10.3.5 Future trends

CMOS image sensors are just starting to be used for the detection of charged particles. After a decade of R&D, the first particle experiment equipped with this type of sensor is about to start taking data. Cameras equipped with 16 Mpixel CMOS image sensors working in direct detection can now be purchased for transmission electron microscopes. While becoming mature for use in field, there are still several R&D programmes in progress in the world, aiming at further improving the performance of this technology.

As the signal from a particle is proportional to the thickness of the substrate, it seems natural to try to increase its thickness and, in order to avoid pixel crosstalks, to introduce a drift field by depleting the sensor. This approach led to the proposal of using a high-voltage process, where up to 100 V could be applied to the diode. With this voltage and with the resistivity of the substrates, about 10–20 µm of silicon can be depleted. The process, available in 180 and 350 nm, also features a deep N-well which allows having both N- and PMOS transistors in the pixel. The electronics sit in the N-well which is also used for detection. Despite its large area, the capacitance is still low because of the extent of the depletion region. The concept was demonstrated in Peric (2007), showing the improvement in signal-over-noise ratio for the detection of single particles (Peric, 2012).

Another approach is also presented in Coath *et al*. (2010). Standard 180 nm CMOS image sensor technology is used but high-resistivity substrates, in the order of 1 kohm cm, are used to enhance the drift field in the detection volume. Noise is reduced by the use of a 4T pixel that achieves noise performance better than 5 e⁻ rms. A similar approach is also used in a sensor for a radiation monitor (Guerrini *et al*., 2012).

10.4   CMOS sensors for X-ray detection

10.4.1 Advanced applications

While Chapter 12 deals with the detection of X-ray for medical imaging, in this chapter we will look at some of the other applications where X-rays are used. A well-known application is luggage scanning; X-rays are also used for industrial analysis with laboratory machines. The structure of molecules or materials can also be studied with X-rays: diffraction studies with laboratory machines are possible, while the best experimental conditions and the most precise observations can be done at synchrotrons. In this type of particle accelerator, high brilliance, high coherence X-ray beams are generated. There are a few tens of synchrotron machines in the world, each with several beamlines where applications as diverse as material science, medical imaging or life science are covered. The energy range at a synchrotron is very wide, extending in the extreme ultraviolet (EUV) range at low energy (100 s of eV), mainly for surface studies, and up to 100 s of keV for studying hard materials. Low-energy X-rays are also used to study the composition of materials in artistic and archaeological objects. X-rays are used for understanding the properties of plasmas generated in fusion reactors (Fujita *et al.*, 1989). In astronomy, the observation of X-rays generated by celestial objects reveals important information about their structure.

As explained above, CMOS substrates are relatively thin, so that direct detection of X-rays for energies above a few keV is inefficient. For low-energy photons, while the thickness of the substrates is adequate, the absorption from the material covering the surface can drastically reduce the performance of the sensor. In this case, backside illumination becomes necessary to restore the overall detective quantum efficiency. Most applications need fairly high-energy X-rays for which a converter, like a scintillator or a phosphor is necessary. Once this step is introduced, sensors developed for visible light detection can be, and are, used most of the time. These sensors are extensively covered in other chapters of this book. However, in this field it is often necessary to achieve large area coverage as it is very difficult to focus X-rays. Detectors specifically developed for X-ray detection then tend to be larger than the standard reticle and stitching is used. This is topic of the next section, which will then be followed by a brief review of the state-of-art of silicon, non-CMOS detectors used for this application.

10.4.2 Stitching

In many applications with X-rays a large field of view is needed. As X-ray lenses are still very difficult to manufacture, the sensor has to cover a large

area. Having large sensors also means that large pixels can be designed thus helping in obtaining a large full well capacity. The need for large pixels and large sensors is also present in applications requiring the detection of charged particles, as it is not possible to focus high-energy particles efficiently. For lower energies, below 1 MeV for electrons, focusing is somehow possible but, because of the details of the charge generation and collection process, large pixels are still useful to obtain high spatial resolution, thus leading again to large sensors.

In modern CMOS processes, a 5:1 reduction in the mask exposure is very common (Bosiers *et al.*, 2008) and the area covered by a single reticle exposure is of the order of about 20 mm × 20 mm, reaching 25 mm × 35 mm for some equipment (Cohen *et al.*, 1999). This area is still smaller than what is required in many applications and much smaller than a widely used 200 mm wafer. In order to manufacture a large sensor, a process known as 'stitching' is used (Theuwissen *et al.*, 1991; Kreider *et al.*, 1995). In stitching, the design is organised in a modular way (see Fig. 10.10 showing the typical arrangement of a reticle and the overall sensor floorplan). The different modules are then put together on a single or multiple mask sets. In a non-stitched design, a single mask would be fully exposed and then moved to the next location on the wafer where it is again fully exposed. In a stitched design, only the area corresponding to one or a few modules would be exposed in one shot, with the rest of the reticle being bladed out. In this way it is possible to create a large focal plane by multiple exposure of the
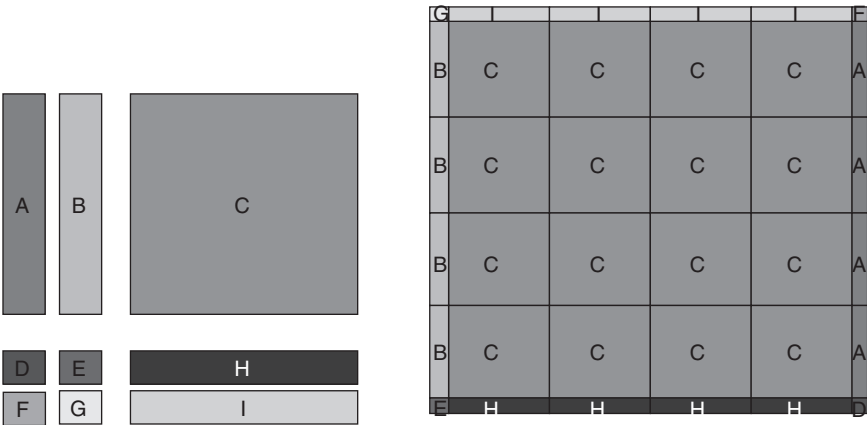
10.10 Reticle (a) and sensor floorplan (b) for a stitched sensor (adapted from Guerrini *et al.*, 2011a, © IOP Publishing Ltd and Sissa Medialab srl; doi: 10.1088/1748-0221/6/03/C03003). Each block on the reticle is labelled with a different letter. The same convention is used for the floorplan of the sensor, showing the replication of the different blocks to achieve the desired size.

pixel array module and then create all the periphery electronics in a similar manner. The end result is a sensor that can be as large as a full CMOS wafer, as shown for example in Reshef *et al.* (2009), Korthout *et al.* (2009), Takahashi *et al.* (2011) and Sedgwick *et al.* (2013). In order for the stitching process to work, some care is required in the photolithography with special design rules needed for the edge, or cut, areas on each module. It also requires a different, modular design approach and special attention to design for high yield.

10.4.3 Other silicon, non-CMOS detectors

As explained above, the substrate of CMOS image sensors is limited to about 20 μm. In order to expand the energy range where efficient detection of X-rays is possible other types of silicon detectors are also used. We have already mentioned hybrid pixel detectors for the detection of charged particles. They are also used in some X-ray applications. The Medipix (Ballabriga *et al.*, 2007) chip is one example. Although originally developed for medical imaging, it has found applications in material science (Firsching *et al.*, 2008). Another similar device has been developed by PSI (Eikenberry *et al.*, 2003; Kraft *et al.*, 2009). In both of these sensors, the in-pixel electronics features low-noise, fast analogue processing, followed by a digital counter. The counter content is read at the end of the exposure to provide a virtually noise-free image. This type of approach is particularly useful for images where pixels can have a very small, down to single photon, signal, as in this case the detector noise is next to zero and the sensor works in a quantum limited regime.

Above we also mentioned that a CMOS monolithic sensor made in an SOI technology can be used for X-ray detection, by using the thick handle wafer as the detecting medium. Other monolithic detectors are built directly in high-resistivity, detector-grade silicon wafers. The typical thickness of these wafers is 300 μm, making them effective for detection of energies up to about 20 keV. The silicon drift chamber was proposed in 1984 (Gatti and Rehak, 1984). The charge generated by the radiation is drifted towards a linear arrangement of anodes by an electric field generated by field-shaping electrodes created on both surfaces of the detector. Thanks to the low anode capacitance, low noise performance can be achieved (Lechner *et al.*, 1996). CMOS electronics can be developed and very low noise can be achieved. In some cases and in order to improve the signal-over-noise ratio, the input transistor of the amplifier is integrated on the high-resistivity substrate of the sensor (Lechner *et al.*, 2004). In another detector built on a high-resistivity silicon wafer, the radiation-induced charge affects the transconductance of a specially designed transistor that generates a fairly high current swing as a result (Kemmer *et al.*, 1990). This device, called a

depleted p-channel field effect transistor (DEPFET), also provides non-destructive readout as well as charge storage. Very low noise can be achieved so that Fano-limited spectra of low noise X-rays can be recorded (Treis *et al.*, 2005). Pixel arrays based around DEPFETs have also been proposed for particle physics experiments (Richter *et al.*, 2003).

10.5   Future trends

While the history of pixel sensors in advanced scientific applications predates the invention of the modern CMOS image sensors, the development of CMOS sensors for high-performance scientific imaging is only recently starting to yield results. CMOS image sensors are starting to be used in particle physics experiments and, as mentioned above, the Alice experiment at LHC is planning for a CMOS sensor for its upgrade. Also the Super-B experiment (Rizzo *et al.*, 2007) is developing monolithic CMOS sensors. If the International Linear Collider or a similar machine comes back on the roadmap of particle physics, CMOS sensors will be again a very good candidate as already shown by the R&D that took place in recent years (Claus *et al.*, 2001). The main technology challenges for the development of such sensors are low noise, radiation hardness and self-triggering, i.e. the ability of each pixel to independently detect the arrival of a particle which leads to further integration of electronics within the pixel.

In the field of X-ray detectors, while CMOS sensors developed for visible light applications can also be used if coupled to a scintillator, specific developments have already taken place. High-dynamic-range and fast readout is required at synchrotron machines. CMOS sensors used in indirect detection can allow accessing new areas of the phase space, see for example new sensors developed for the Diamond Light Source in the UK (Tartoni, 2012). The wavelength domain between UV and low-energy X-ray is very difficult to access as it requires backthinning for backside illumination as well as thick substrates, given that the absorption length rapidly increases from sub-micron to several microns. New developments using CMOS sensors have started to appear (Wunderer *et al.*, 2012; Hoenck *et al.*, 2013). For synchrotrons, high-speed and dynamic range are again of interest, but similar developments could also become interesting for astronomy, especially in adaptive optics (Downing *et al.*, 2012).

10.6   Sources of further information and advice

Being such a young field, there is yet no book to comprehensively cover the development of CMOS sensors for scientific applications. Information is scattered around in conference proceedings, mainly in conferences where detectors for these types of applications are presented. The largest

conference for particle and X-ray detectors is the IEEE Nuclear Science Symposium, but other, smaller conferences can provide a better reference to the underlining technology. Pixel takes place every 2–3 years and is an international workshop on semiconductor pixel detectors for particles and imaging. Some developments especially on transmission electron microscopy has also been recently presented at Iworid, another international workshop on radiation imaging detectors.

10.7　References

Aad G. *et al*. (2008) 'ATLAS pixel detector electronics and sensors', *Journ. of Instr.* 3 P07007.

Aoyama T., Oka Y., Honda K. and Mori C. (1992) 'A neutron detector using silicon PIN photodiodes for personal neutron dosimetry', *Nucl. Instr. and Meth.* 314(3), 590–594.

Arai Y. *et al*. (2006) 'Monolithic pixel detector in a 0.15 μm SOI technology', *IEEE Nucl. Sci. Symp.*, Conference record, N34-4.

Arvanitis C.D., Bohndiek S.E., Blakesley J., Olivo A. and Speller R.D. (2009), 'Signal and noise transfer properties of CMOS based active pixel flat panel imager coupled to structured CsI:Tl', *Med Phys.* 36(1), 116–26.

Augustin J.-E. (2004) 'Detector concepts for a linear collider', *Nucl. Instr. and Meth. A* 535, 36–47.

Bailey R. *et al*. (1985) 'A silicon strip detector telescope for the measurement of production and decay of charmed particles', *Nucl. Instr. and Meth.* 226, 56–58

Bak J.F., Burenkov A., Petersen J.B.B., Uggerhøj E., Møller S.P. and Siffert P. (1987), 'Large departures from Landau distributions for high-energy particles traversing thin Si and Ge targets', *Nucl Phys B* 288, 681–716.

Ballabriga R., Campbell M., Heijne E.H.M., Llopart X. and Tlustos L. (2007) 'The Medipix3 prototype, a pixel readout chip working in single photon counting mode with improved spectrometric performance', *IEEE Trans. on Nucl. Sc.* 54(5), 1824–1829.

Ballin J.A. *et al*. (2008) 'Monolithic active pixel sensors (MAPS) in a quadruple well technology for nearly 100% fill factor and full CMOS pixels', *Sensors* 8(9), 5336–5351.

Battaglia M. *et al*. (2010) 'Characterisation of a CMOS active pixel sensor for use in the TEAM microscope', *Nucl. Instr. and Meth. A* 622, 669–677.

Baudot J. *et al*. (2009) 'First test results of MIMOSA-26, a fast CMOS sensor with integrated zero suppression and digitized output', *IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, 1169–1173.

Becksa K.H. *et al*. (1997) 'The DELPHI pixels', *Nucl. Instr. and Meth. A* 386, 11–17.

Becksa K.H. *et al*. (1998a) 'Commissioning of the DELPHI pixel detector', *Nucl. Instr. and Meth. A* 418, 15–21.

Becksa K.H. *et al*. (1998b) 'First results from the DELPHI pixel detector', *Nucl. Instr. and Meth. A* 409, 229–231.

Belau E. *et al*. (1983) 'Silicon detectors with 5 μm spatial resolution for high energy particles', *Nucl. Instr. and Meth.* 217, 224–228.

Beringer J. *et al.* (Particle Data Group) (2012) 'The review of particle physics', *Phys. Rev.* D86, 010001.

Besson A. *et al.* (2011) 'CMOS sensors with high resistivity epitaxial layer', XXIst International Europhysics Conference on High Energy Physics.

Bethe H.A. (1930), 'Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie', *Ann Phys* 5, 325.

Bethe H.A. (1953) 'Moliere's theory of multiple scattering', *Phys Rev* 89 1256–1266.

Bichsel H. (1988), 'Straggling in thin silicon detectors', *Rev Mod Phys* 60(3), 663–669.

Bollinger L.M., Thomas G.E. and Ginther R.G. (1959), 'Glass scintillators for neutron detection', *Rev. Sci. Instrum.* 30, 1135–1136.

Bollinger L.M., Thomas G.E. and Ginther R.G. (1962), 'Neutron detection with glass scintillators', *Nucl. Instr. and Meth.* 17(1), 97–116.

Bosiers J. *et al.* (2008) 'Challenges and innovations in very-large CCD and CMOS imagers for professional imaging', invited talk at Silicon Photonics and Photonic Integrated Circuits, *Proc. of SPIE* 6996, 69960Z.

Brouard M., Campbell E.K., Johnsen A.J., Vallance C., Yuen W.H. and Nomerotski A. (2008) 'Velocity map imaging in time of flight mass spectrometry', *Rev. Sci. Instrum.* 79, 123115; http://dx.doi.org/10.1063/1.3036978.

Cabello J. and Wells K. (2010) 'The spatial resolution of silicon-based electron detectors in beta-autoradiography', *Phys Med Biol.* 55(6), 1677–1699.

Claus G., Colledani C., Deptuch G., Dulinski W., Gornushkin Y., Hu Y. and Winter M. (2001) 'Monolithic active pixel sensors for a linear collider', *Nucl. Instr. and Meth. A* 473(1), 83–85.

Coath R.E. *et al.* (2010) 'A low noise pixel architecture for scientific CMOS monolithic active pixel sensors', *IEEE Trans. on Nucl. Sc.* 57, 2490–2496.

Cohen D., Koltin E., Ayelet M.I., Shacham A. (1999) 'Stitching design rules for forming interconnect layers', *US Patent No.* 6,225,013.

Colledani C., Dulinski W., Turchetta R., Djama F., Rudge A. and Weilhammer P. (1996), 'A submicron precision silicon telescope for beam test purposes', *Nucl Instr and Meth in Phys Res A* 372(3), 379–384.

Contarato D., Denes P., Doering D., Joseph J. and Krieger B. (2011) 'Direct detection in transmission electron microscopy with a 5 µm pitch CMOS pixel sensor', *Nucl. Instr. and Meth. A* 635, 69–73.

Crooks J.P. *et al.* (2007) 'A novel CMOS monolithic active pixel sensor with analog signal processing and 100% fill factor', *Proceedings of the IEEE Symposium on Nuclear Science* 2, 931–935.

Damerell C. (1995), 'Vertex detectors: the state of the art and future prospects', *RAL-P-95-008*, December, available at http://hepwww.rl.ac.uk/damerell/2003/Vertex-detector-95-008.pdf

Degerli Y., Deptuch G., Fourches N., Himmi A., Yan Li, Lutz P., Orsini F. and Szelezniak M. (2005) 'A fast monolithic active pixel sensor with pixel-level reset noise suppression and binary outputs for charged particle detection', *IEEE Trans. on Nucl. Sci.* 52, 3186–3193.

Delpierre P. *et al.* (1994) 'Large scale pixel detectors for DELPHI at LEP200 and ATLAS at LHC', *Nucl. Instr. and Meth. A* 342, 233–239.

Deptuch G. *et al.* (2000) 'Design and testing of monolithic active pixel sensors for charged particle tracking', *IEEE Nuclear Science Symposium Conference Record* 1, 103–110.

Deptuch G., Winter M., Dulinski W., Husson D., Turchetta R. and Riester J.L. (2001) 'Simulation and measurements of charge collection in monolithic active pixel sensors', *Nucl. Instr. and Meth. A* 465, 92–100.

Dierickx B. *et al.* (1993) 'Integration of CMOS-electronics and particle detector diodes in high-resistivity silicon-on-insulator wafers', *IEEE Trans. on Nucl. Sci.* 40, 753–758.

Dierickx B., Meynants G. and Scheffer D. (1997) 'Near-100% fill factor standard CMOS active pixels', in *Proc. 1997 IEEE Workshop on CCDs and Advanced Image Sensors*, Bruges, Belgium.

Downing M. *et al.* (2012), 'Backside-illuminated, high-QE, 3e-RoN, fast 700fps, 1760 × 1680 pixels CMOS imager for AO with highly parallel readout', in *SPIE Astronomical Telescopes+ Instrumentation*, International Society for Optics and Photonics. pp. 84530C-84530C.

Edelstein R.S. *et al.* (2011) 'Process integration aspects of back illuminated CMOS imagers using smart stacking™ technology with best in class direct bonding', *Proceedings of the International Image Sensor Workshop (IISW)*.

Eikenberry E.F., Brönnimann C., Hülsen G., Toyokawa H., Horisberger R., Schmitt B., Schulze-Briese C. and Tomizaki T. (2003) 'PILATUS: a two-dimensional X-ray detector for macromolecular crystallography', *Nucl. Instr. and Meth. A* 501(1), 260–266.

Fano U. (1947), 'Ionization yield of radiations. II. The fluctuations of the number of ions', *Phys. Rev.* 72, 26–29.

Faruqi A.R. and Henderson R. (2007) 'Electronic detectors for electron microscopy', *Curr Opin Struct Biol.* 17(5) 549–555.

Faruqi A.R. and McMullan G. (2011) 'Electronic detectors for electron microscopy', *Quarterly Reviews of Biophysics* 44(3), 357–390.

Faruqi A.R., Henderson R., Prydderch M., Turchetta R., Allport P. and Evans A. (2005) 'Direct single electron detection with a CMOS detector for electron microscopy', *Nucl. Instr. and Meth. A* 546, 170–175.

FEI (2009) http://investor.fei.com/releasedetail.cfm?ReleaseID=399045

Firsching M., Takoukam Talla P., Michel T. and Anton G. (2008), 'Material resolving X-ray imaging using spectrum reconstruction with Medipix2', *Nucl. Instr. and Meth. A* 591(1), 19–23.

Fujita J., Moritaa S. and Sakuraia M. (1989) 'X-ray diagnostics for fusion plasmas', *Laser and Particle Beams* 7, 483–486.

Gatan (2012) http://www.gatan.com/K2/

Gatti E. and Rehak P. (1984) 'Semiconductor drift chamber – an application of a novel charge transport scheme', *Nucl. Instr. and Meth.* 225(3), 608–614.

Granja C., Jakubek J., Platkevic M., Pospisil S. and Vykydal Z. (2009) 'Detection and real time spectroscopy of charged particles with the timepix pixel detector', *AIP Conference Proceedings* 1204, 75–79.

Greinera L. *et al.* (2011) 'A MAPS based vertex detector for the STAR experiment at RHIC', *Nucl. Instr. and Meth. A* 650, 68–72.

Guerrini N., Turchetta R., Van Hoften G., Henderson R., McMullan G. and Faruqi A.R. (2011a) 'A high frame rate, 16 million pixels, radiation hard CMOS sensor', *Journal of Instrumentation* 6(03), C03003.

Guerrini N. *et al.* (2011b) 'A 61 mm × 63 mm, 16 Million pixels, 40 frames per second, radiation-hard CMOS image sensor for transmission electron microscopy', *Proceedings of the International Image Sensor Workshop (IISW)*.

Guerrini N. *et al.* (2012) 'Design and characterisation of a highly miniaturised radiation monitor', *Nucl. Instr. and Meth. A* 731, 154–159.

Hara K., Kochiyama M., Koikea K., Sega T., Shinshoa K., Araib Y., Fujitab Y., Okiharah Miurah M. and Motoyoshii M. (2010) 'Development of INTPIX and CNTPIX silicon-on-insulator monolithic pixel devices' *Proceedings of Science 19th International Workshop on Vertex Detectors*.

Hartmann F. (2012) 'Silicon tracking detectors in high-energy physics', *Nucl. Instr. and Meth. A* 666, 25–46.

Heijne E.H.M., Jarron P., Olsen A. and Redaelli N. (1988) 'The silicon micropattern detector: a dream?', *Nucl. Instr. and Meth. in Phys. A* 273(2–3), 615–619.

Heitler W. (1949) 'Theory of meson production', *Rev. Mod. Phys.* 21, 113.

Henderson R. and Unwin P.N.T. (1975) 'Three-dimensional model of purple membrane obtained by electron microscopy', *Nature* 257, 28–32.

Hoenck M.E. *et al.* (2013a), 'The DUV stability of superlattice-doped CMOS detector arrays', *Proceedings of the International Image Sensor Workshop (IISW)*.

Holland S.E. (1989a) 'Fabrication of detectors and transistors on high-resistivity silicon', *Nucl. Instr. and Meth. A* 275, 537–541.

Holland S.E. (1989b) 'An IC-compatible detector process', *IEEE Trans. on Nucl. Sci.* 36, 283–289.

Holland S.E. and Spieler H. (1990) 'A monolithically integrated detector-preamplifier on high-resistivity silicon', *IEEE Trans. on Nucl. Sci.* 37, 463–468.

Hu-Guo C. *et al.* (2009) 'CMOS pixel sensor development: a fast read-out architecture with integrated zero suppression', *JINST* 4 P04012.

Ikeda H. *et al.* (2007) 'Deep sub-micron FD-SOI for front-end application', *Nucl. Instr. and Meth. A* 579, 701–705.

Janesick J.R. (2007) *Photon Transfer DN* $\rightarrow \lambda$, Bellingham, SPIE.

John J.J. *et al.* (2012) 'PImMS, a fast event-triggered monolithic pixel detector with storage of multiple timestamps', *Journ. Of Instr.* 7, C08001.

Kemmer J., Lutz G., Prechtel U., Schuster K., Sterzik M., Strüder L. and Ziemann T. (1990) 'Experimental confirmation of a new semiconductor detector principle', *Nucl. Instr. and Meth.* 288(1), 92–98.

Kenney C., Parker S., Snoeys W., Plummer J., Rosseel G. and Aw C.H. (1993) 'Performance of a monolithic pixel detector', *Nucl. Instr. and Meth. A* 232, 460–467.

Klein C.A. (1968) 'Bandgap dependence and related features of radiation ionization energies in semiconductors', *J Appl Phys* 39, 2029–2038.

Kluge A. *et al.* (2007) 'The ALICE silicon pixel detector', *Nucl. Instr. and Meth. A* 582, 728–732.

Knoll M. (1935) 'Aufladepotentiel und Sekundäremission elektronenbestrahlter Körper', *Zeitschrift für technische Physik* 16, 467–475.

Knoll G.F., Knoll T.F. and Henderson T.M. (1988), 'Light collection in scintillation detector composites for neutron detection', *IEEE Trans. on Nucl. Sc.* 35(1), 872–875.

Kochiyama M. *et al.* (2011) 'Radiation effects in silicon-on-insulator transistors with back-gate control method fabricated with OKI Semiconductor 0.20μm FD-SOI technology', *Nucl. Instr. And Methods A* 636; S62–S67.

Korthout L., Verbugt D., Timpert J., Mierop A., de Haan W., Maes W., de Meulmeester J., Muhammad W., Dillen B., Stoldt H., Peters I. and Fox E. (2009) 'A wafer-scale

CMOS APS imager for medical X-ray applications', *Proceedings of the International Image Sensor Workshop (IISW)*.

Kotlinski D. (2009) 'Status of the CMS pixel detector', *Journ. Of Instr.* 4, P03019.

Kraft P., Bergamaschi A., Broennimann C., Dinapoli R., Eikenberry E.F., Henrich B. and Schmitt B. (2009) 'Performance of single-photon-counting PILATUS detector modules', *J. of Synchrotron Radiation* 16(3), 368–375.

Kreider G. *et al.* (1995) 'An mK × nK modular image sensor design', *IEDM Dig. Techn. Papers*, 155–158.

Landau L.D. (1944), 'On the energy loss of fast particles by ionization', *J. Exp. Phys. (USSR)* 8, 201–205.

Lechner P., Eckbauer S., Hartmann R., Krisch S., Hauff D., Richter R. and Sampietro M. (1996) 'Silicon drift detectors for high resolution room temperature X-ray spectroscopy', *Nucl. Instr. and Meth.* 377(2), 346–351.

Lechner P., Pahlke A. and Soltau H. (2004) 'Novel high-resolution silicon drift detectors', *X-Ray Spectrometry* 33(4), 256–261.

McMullan G., Faruqi A.R., Henderson R., Guerrini N., Turchetta R., Jacobs A. and Van Hoften G. (2009a) 'Experimental observation of the improvement in MTF from backthinning a CMOS direct electron detector', *Ultramicroscopy* 109, 1144–1147.

McMullan G., Clark A.T., Turchetta R. and Faruqi A.R. (2009b) 'Enhanced imaging in low dose electron microscopy using electron counting', *Ultramicroscopy* 109(12), 1411.

McMullan G., Chen S., Henderson R. and Faruqi A.R. (2009c), 'The detective quantum efficiency of electron area detectors in electron microscopy', *Ultramicroscopy* 109, 1126–1143

Maneuski D., Eklund L., Kralik M., Laing A., O'Shea V., Pospisil S. and Vykydal Z. (2008) 'Evaluation of silicon monolithic APS as a neutron detector', *IEEE Nucl. Sc. Symposium Conference Record*, 2383–2386.

Mendis S., Kemeny S.E. and Fossum E.R. (1994) 'CMOS active pixel image sensor', *IEEE Trans on Electron Devices* 41(3), 452–453.

Milazzo A. *et al.* (2005) 'Active pixel sensor array as a detector for electron microscopy', *Ultramicroscopy* 104, 152–159.

Musa L. *et al.* (2012) 'Conceptual design report for the upgrade of the ALICE ITS', *CERN-LHCC*-2012-005; LHCC-G-159.

Nomerotski A. *et al.* (2010) 'Pixel imaging mass spectrometry with fast and intelligent pixel detectors', *Journ. Of Instr.* 5, C07007.

Ono Y., Ishikawa A., Yamamoto H., Arai Y., Tsuboyama T., Onuki Y., Iwata A., Imamura T. and Ohmoto T. (2013) 'Development of the pixel OR SOI detector for high energy physics experiments', *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 731, 266–269.

Pain B. (2009) 'Backside illumination technology for SOI-CMOS image sensors', *Proceedings of the International Image Sensor Workshop (IISW)*.

Pain B. and Zheng X. (2002). *US Patent* No. 6,380,572.

Palik E.D. (1985) *Handbook of Optical Constants of Solids*, New York, Academic Press.

Parker S. (1989) 'A proposed VLSI pixel device for particle detection', *Nucl. Instr. and Meth. A* 275, 494–516.

Pengg F. (1996) 'Monolithic silicon pixel detectors in SOI technology', Ph.D. Thesis, Johannes Kepler University, Linz, Austria.

Peric I. (2007) 'A novel monolithic pixelated particle detector implemented in high-voltage CMOS technology', *Nucl. Instr. and Meth. A* 582, 876–885.

Peric I. (2012) 'Active pixel sensors in high-voltage CMOS technologies for ATLAS', *Journal of Instrumentation* 7(08), C08002.

Perkins D.H. (2000) *Introduction to High Energy Physics*, Cambridge, Cambridge University Press.

Reshef R., Leitner T., Alfassi S., Sarig E., Golan N., Berman O., Fenigstein A., Wolf H., Hevel G., Vilan S. and Lahav A. (2009) 'Large-format medical X-ray cmos image sensor for high resolution high frame rate applications', *Proceedings of the International Image Sensor Workshop (IISW)*.

Richter R.H., Andricek L., Fischer P., Heinzinger K., Lechner P., Lutz G. and Wermes N. (2003) 'Design and technology of DEPFET pixel sensors for linear collider applications', *Nucl. Instr. and Meth.* 511(1) 250–256.

Rizzo G., Batignani G., Bettarini S., Bosi F., Calderini G., Cenci R. and Zoccoli A. (2007) 'Recent development on triple well 130 nm CMOS MAPS with in-pixel signal processing and data sparsification capability', *Nuclear Science Symposium Conference Record* 2, 927–930.

Ruska E. (1987) 'The development of the electron microscope and of electron microscopy', *Bioscience Reports* 7(8), 607–629.

Ruska E and Knoll M (1932) 'Das Elektronenmikroskop', *Z Phys* 78, 318–339.

Ryu S.G. *et al.* (2011) 'First performance evaluation of an X-ray SOI pixel sensor for imaging spectroscopy and intra-pixel trigger', *IEEE Trans. on Nucl. Sci.* 58, 2528–2536.

Sedgwick I. *et al.* (2012) 'PImMS: a self-triggered, 25 ns resolution monolithic CMOS sensor for time-of-flight and imaging mass spectrometry', *Proceedings of the IEEE NEWCAS 2012 Conference*.

Sedgwick I., Das D., Guerrini N., Marsh B. and Turchetta R. (2013), 'LASSENA: a 6.7 megapixel, 3-sides buttable wafer-scale CMOS sensor using a novel grid-addressing architecture', *Proceedings of the International Image Sensor Workshop (IISW)*.

Shouleh N., Yu Q., Smith A.L., Jones T.J., Tombrello T.A. and Elliott S.T. (1998), 'Direct detection and imaging of low-energy electrons with delta-doped charge-coupled devices', *Appl. Phys. Lett.* 73, 3417–3419.

Snoeys W. (1992) 'A new integrated pixel detector for high energy physics', *Technical Report 92-017*, Stanford University.

Snoeys W. and Parker S. (1995) 'Integrated VLSI radiation/particle detector with biased pin diodes', *US Patent* No. 5,465,002.

Snoeys W., Plummer J., Rosseel G., Aw C.H., Kenney C. and Parker S. (1993) 'First beam test results from a monolithic silicon pixel detector', *Nucl. Instr. and Meth. A* 326, 144–149.

Spieler H. (1998) 'Introduction to radiation detectors and electronics', *Lecture Notes – Physics 198*, Spring Semester 1998 – UC Berkeley.

Stanitzki M. *et al.* (2007) 'A tera-pixel calorimeter for the ILC', *Proceedings of the IEEE Symposium on Nuclear Science* 1, 254–258.

Takahashi H., Yamashita Y., Kikuchi S., Fujita M., Hirayama S., Kanou T., Hashimoto S., Shigeta K., Aoki T., Momma G. and Inoue S. (2011) 'A 300 mm wafer-size

CMOS image sensor for low-light-level imaging', *Proceedings of the International Image Sensor Workshop (IISW)*.

Tartoni N. (2012) private communication.

The ILD Concept Group (2010) 'The international large detector: letter of intent', DESY 2009-87, Fermilab-Pub-09-682-E, KEK Report 2009-6.

Theuwissen A. *et al.* (1991), 'A 2.2 Mpixel FT CCD imager according to the Eureka HDTV standard', *IEDM Techn. Dig.*, 167–170.

Tietz D., Tietz H., Nickell S., Baumeister W. and Plitzko J.M. (2008), 'Characterization of a fiber-optically coupled 8k CCD/CMOS device', *Proceedings of the EMC 2008 14th European Microscopy Congress*, Aachen, Germany, I1, I1.3, 89–90.

Treis J., Fischer P., Halker O., Harter M., Herrmann S., Kohrs R. and Wermes N. (2005) 'DEPMOSFET active pixel sensor prototypes for the XEUS wide field imager', *IEEE Trans. on Nucl. Sc.* 52(4), 1083–1091.

Turala M. (2005) 'Silicon tracking detectors—historical overview', *Nucl. Instr. and Meth. A* 541, 1–14.

Turchetta R., Berst J.D., Casadei B., Claus G., Colledani C., Dulinski W., Hu Y., Husson D., Le Normand J.P., Riester J.L., Deptuch G., Goerlach U., Higueret S. and Winter M. (2001) 'A monolithic active pixel sensor for charged particle tracking and imaging using standard VLSI CMOS technology', *Nucl. Instr. and Meth. A* 458, 677–689.

Turchetta R., Villani G. and Prydderch M. (2011) 'Accelerated particle and high energy radiation sensor', *US Patent* No. 7,944,012.

Vanstraelen G., Debusschere I., Claeys C. and Declerck G. (1988) 'New concepts for integrated solid state detector electronics', *Nucl. Instr. and Meth. A* 273, 625–629.

Vanstraelen G., Debusschere I., Claeys C. and Declerck G. (1989) 'Fully integrated CMOS pixel detector for high energy particles', *Nucl. Instr. and Meth. A* 275, 574–579.

Vanstraelen G., Debusschere I., Claeys C. and Declerck G. (1991) 'Design implications of a p-well CMOS technology for the realization of monolithic integrated pixel arrays', *Nucl. Instr. and Meth. A* 305, 541–548.

Vavilov P.V. (1957), 'Ionization losses of high-energy heavy particles', *Sov Phys JETP* 5, 749.

Wermes N. (2009) 'Pixel detectors for charged particles', *Nucl. Instr. and Meth. A* 604, 370–379.

Wrigley C., Yang G. and Pain B. (2000) 'High responsivity CMOS imager pixel implemented in SOI technology', *Proceedings of the 2000 IEEE SOI Conference*, 138–139.

Wunderer C. *et al.* (2012), 'The PERCIVAL soft X-ray imager', presented at the *11th International Conference on Synchrotron Radiation Instruments (SRI)*.

Zhao W., Ristic G. and Rowlands J.A. (2004), 'X-ray imaging performance of structured cesium iodide scintillators', *Med Phys.* 31(9) 2594–605.

11

Complementary metal-oxide-semiconductor (CMOS) sensors for fluorescence lifetime imaging (FLIM)

R. K. HENDERSON, University of Edinburgh, UK,
B. R. RAE, ST Microelectronics, UK and
D.-U. LI, University of Strathclyde, UK

Abstract: This chapter begins by reviewing the theory, techniques and applications of fluorescence lifetime sensing. It then looks at existing instrumentation and the ways in which complementary metal-oxide-semiconductor (CMOS) technology has sought to complement these technologies by providing low cost, robust and miniaturised sensors with increased throughput and dynamic range. A number of pixels and sensor architectures are compared with a view to future fully-integrated, lifetime imaging systems based on CMOS sensors.

Key words: CMOS image sensor, fluorescence lifetime imaging, time-correlated single photon counting, single photon avalanche diode, charge coupled device.

11.1 Introduction

Fluorescence lifetime sensing constitutes one of the most demanding applications of solid-state imaging. In addition to the high sensitivity and image resolution expected by microscopists to observe the finest features of live biological cells, the technique demands extremely high temporal resolution. While the latter performance requirements are amply served by a wide choice of solid-state technologies such as electron multiplying charge coupled device (EMCCD), intensified charge coupled device (ICCD) or most recently scientific complementary metal-oxide-semiconductor (sCMOS) sensors (Theuwissen and Seitz, 2011), fluorescence lifetime imaging (FLIM) is still performed mostly by vacuum tube or image intensifier based approaches (Becker, 2005). However, this situation is changing; in the last decade a number of adaptations of complementary metal-oxide-semiconductor (CMOS) or chage coupled device (CCD) image sensors have emerged, capable of delivering the simultaneous extremes of single photon sensitivity, megapixel spatial and picosecond temporal resolution to observe large cohorts of single molecules in dynamic cellular processes (Esposito, 2011). Such imaging technology holds enormous

312

potential to assist researchers in improving our understanding of the processes underlying normal cellular function, and their alteration in disease states.

The intense scientific interest in FLIM resides in the fact that the fluorescence lifetime of a fluorophore depends on its molecular environment but not on its concentration (Lakowicz, 2006). As a result, effects at molecular scale can be studied independently of the highly variable concentration of the fluorophore. The theory and applications of fluorescence lifetime sensing will be reviewed before looking at existing instrumentation. FLIM has currently acquired two main techniques broadly classified into time-domain or frequency-domain. Time-domain FLIM is performed by time-correlated single photon counting (TCSPC) or by gated image intensifiers. Frequency-domain FLIM is gain-modulated photomultiplier tubes or image intensifiers. The operating principles and performance of these systems have inspired most of the solid-state approaches and therefore will be reviewed prior to examining their CMOS equivalents. Finally, the aspect of lifetime estimation will be considered. The computation of the exponential decay time (or times) and coefficients can be a complex process almost exclusively rendered by silicon computation of some description (Verveer *et al.*, 2000). This may range from software post-processing, real-time field programmable gate array (FPGA) based algorithms to on-chip lifetime determination. This final stage in the image processing pipeline is less often discussed within treatises on time-resolved hardware. However, we deem this a critical aspect in this context, as the capability of integrated circuits to perform high speed digital computation is capable of providing video rate or real time fluorescence lifetime estimation. Thereby, new applications of FLIM are likely to emerge in high-throughput screening. near-infrared spectroscopy, cell-sorting or live-cell imaging, enabled entirely by solid-state implementations.

11.2   Fluorescence lifetime imaging (FLIM)

11.2.1 Theory

Fluorescence-based methodologies are at the core of many modern instrumentation technologies, especially in the life sciences (Michalet *et al.*, 2003). Originally, the interest was in imaging of specific labelled biological samples, more recently with the advent of deoxyribonucleic acid (DNA) sequencing and micro-array applications. The equipment needed for such spectroscopic instrumentation includes a narrow wavelength source to excite the fluorophore of interest. The resulting fluorescence must pass through an optical filter to separate the excitation light from the fluorescence emission, before being detected by a light sensor.

Time-resolved fluorescence analysis is the measurement of the temporal properties of a fluorophore sample. Fluorescence lifetime detection provides a method of differentiating between spectrally overlapping samples which exhibit different lifetime properties (Cubeddu *et al.*, 2002). The sensitivity of a sample's lifetime properties to the micro-environment provides an extremely powerful analysis tool. The equipment required to perform FLIM includes a pico-second pulsed or modulated light source (often a laser), a sensitive detector such as a micro-channel plate photo-multiplier tube (MCP-PMT) or single photon avalanche diode (SPAD) detector along with associated signal processing electronics and software, most commonly within a microscope containing a variety of lenses and filters.

11.2.2 Fluorescence lifetime

Fluorophores have an exponential fluorescent decay transient after the removal of the excitation source, which defines their characteristic lifetime. Due to the random nature of fluorescence emission, a fluorescent sample's associated lifetime is the average time the molecules in a sample spend in the excited state before photon emission occurs (Fig. 11.1).
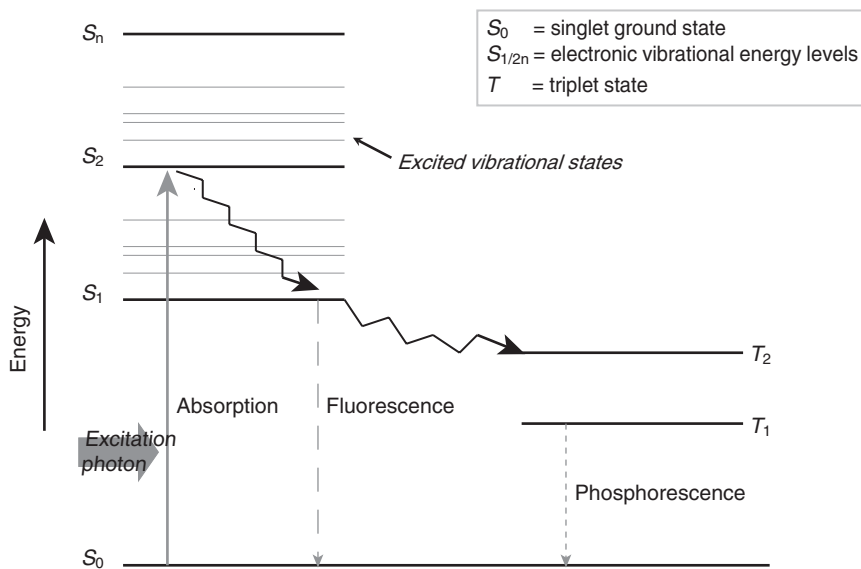
11.1 Jablonski diagram. Promotion of an electron to a higher energy state by excitation light and its subsequent relaxation to the ground state, coupled with the emission of a lower energy photon.

A sample's fluorescence lifetime, τ, is determined by the rate at which the sample leaves the excited state (Equation 11.1). The transition can occur via two mechanisms, either by fluorescence emission (at rate Γ) or by competing non-radiative processes (represented collectively as K_{nt}).

$$\tau = \frac{1}{\Gamma + \Sigma K_{nt}} \tag{11.1}$$

A fluorophore's quantum yield (Θ) is the ratio of emitted photons to the number of absorbed photons. This can be represented by Equation 11.2.

$$\Theta = \frac{\Gamma}{\Gamma + \Sigma K_{nt}} \tag{11.2}$$

For a given excitation light intensity, a fluorophore's brightness (molecular brightness, q) can be calculated if the molecular absorption coefficient (ε) is known (Equation 11.3).

$$q = \varepsilon \times \Theta \tag{11.3}$$

The absorption coefficient of a fluorophore is usually constant; therefore, changes in a fluorophore's brightness can usually be attributed to changes in the sample's quantum efficiency. Therefore, from Equations 11.2 and 11.3, if the fluorescence intensity changes this will usually result in a change in sample lifetime. Due to the fact that fluorescence intensity is a composite property of a sample, dependent on sample quantity and concentration as well as instrument set-up, it is very sensitive to sample variation and is subject to interference from scattered light. This makes the observation of small intensity changes very difficult. Conversely, fluorescence lifetime is an intrinsic fluorophore property, independent of sample volume and concentration. Lifetime analysis is also less sensitive to instrument setup. Fluorescence lifetime is therefore a more robust analysis method compared to intensity measurement, capable of observing subtle changes in sample conditions (Turconi et al., 2001). The rate of non-radiative recombination is dictated by the fluorophore's electron structure and its interaction with the environment. Non-radiative decay mechanisms include (Lakowicz, 2006):

- inter-system crossing
- collisional or static quenching
- solvent effects
- resonance energy transfer.

Fluorescence intensity is related to lifetime according to Equation 11.4 (for a mono-exponentially decaying sample). The equation assumes that the sample has been excited by an infinitely short (δ-function) light pulse. The time-dependent intensity at time t, I(t), is given by:

$$I(t) = I_0 \exp\left(\frac{-t}{\tau}\right)$$ [11.4]

Fluorescence lifetime is independent of fluorophore concentration but dependent on the sample's local environment. Thus, lifetime detection allows precise quantitative data about both the fluorophore distribution and local environment to be obtained, while avoiding the problems related to fluorescence intensity imaging such as photo-bleaching (Christenson and Sternberg, 2004). Fluorescence lifetime detection can also be used to differentiate between fluorophores with overlapping spectra, but exhibiting different decay characteristics. Typical fluorescence decay times of organic compounds fall between a few hundreds of picoseconds and several nanoseconds. There are a number of different imaging experiments for which time-resolved detection can be used; these include, multiple fluorophore labelling (Pepperkok et al., 1999), quantitative detection of ion concentrations and oxygen and energy transfer characteristics using fluorescence resonance energy transfer (FRET) (Prasad, 2003).

11.2.3 Lifetime measurement techniques

There are two main techniques for measuring the fluorescence lifetime of a sample: the frequency-domain method and the time-domain method. In the frequency domain method a sample is excited by an intensity modulated light source. The fluorescence emission is modulated at the same frequency, but with a phase shift due to the intensity decay law (Equation 11.4) of the sample (Lakowicz, 2006; Chodavarapu et al., 2005) and a reduction in the modulation depth. In the time domain method the intensity decay of a fluorescent sample is directly measured as a function of time, following absorption of a short excitation pulse.

Fluorescence lifetime imaging is achieved by two methods: wide-field imaging and point scanning (detector or light source). Wide-field imaging makes use of a multi-pixel detector, collecting data for each pixel location simultaneously, while point-scanning relies on an x-y stage to scan the laser or detector location over the region of interest in order to build up the image one pixel at a time.

TCSPC is a time-domain, point scanning lifetime measurement technique that relies on single photon sensitive detectors to obtain photon arrival time information for each gathered photon. Each detected photon is logged along with a time stamp denoting the photon's arrival time relative to a repetitive synchronisation pulse from the pulsed light source (Fig. 11.2). This process is repeated many times in order to produce a decay histogram. To produce a fluorescence lifetime image this process is performed either simultaneously at different detector pixel locations (Veerappan et al., 2011)
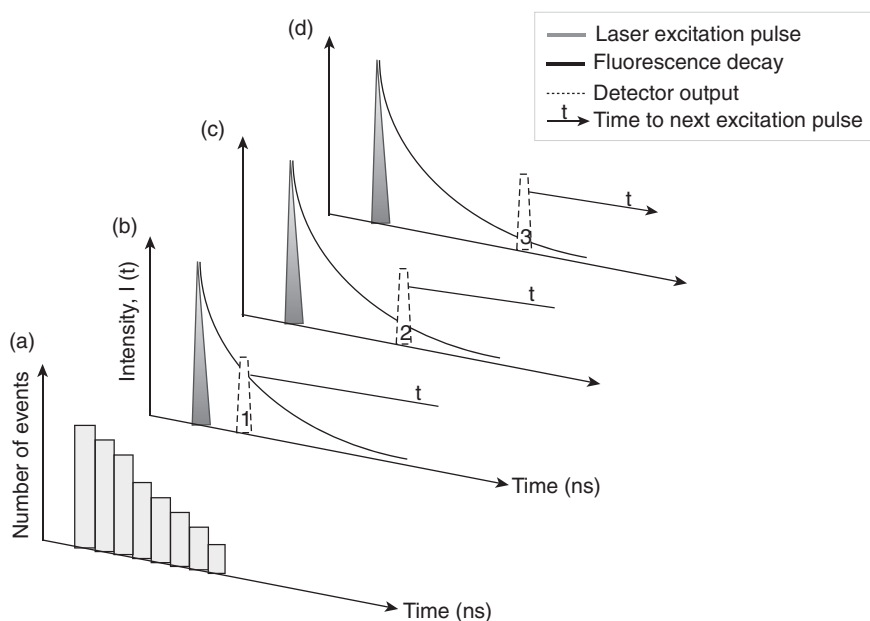
11.2 Reverse start-stop TCPSC principle. The sample of interest is excited by a short laser pulse and the time between the first fluorescence photon detected and the subsequent excitation pulse is recorded. A histogram of photon arrival time (a) is generated by repeating this measurement many times (b–d).

or sequentially as the detector or light source is scanned across the region of interest.

In a scan based system the maximum frame rate that can be achieved using TCSPC is limited by the rate at which the detector head or light source can be scanned and both techniques are limited by the relatively low count that is required in order to avoid such issues as pulse pile up. However, TCSPC is very photon efficient, with data from all gathered photons being processed. This minimises sample exposure to excitation light and provides time information on each photon, which is important for items such as single-molecule detection, fluorescence correlation spectroscopy and phosphorescence lifetime imaging.

Photon counting applications require detectors of single-photon sensitivity; these include micro-channel plate PMTs, high-speed amplified PMTs, discrete photodiodes and avalanche photodiodes. These devices tend to be discrete components, requiring separate power supplies and a communication interface. Furthermore, they tend to be physically large and delicate. These devices achieve single photon sensitivity through
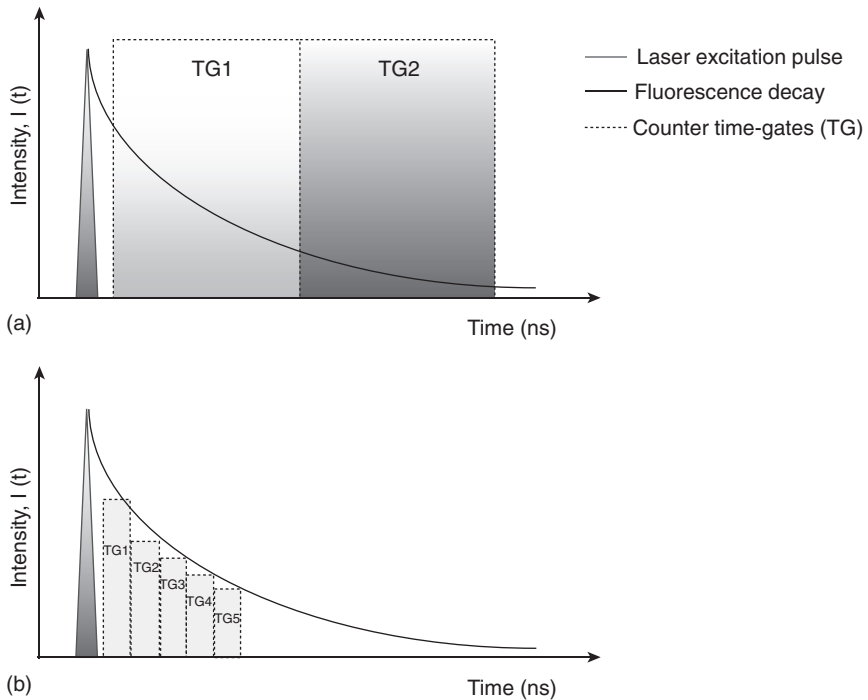
11.3 Time-gated FLIM principle. Fluorescence decay is captured using a series of two (a) or multiple time-gates (b).

electron multiplication, triggered by an initial electron-hole pair caused by an incident photon.

An alternative method of capturing fluorescence lifetime information in the time domain is to use gated detection (Fig. 11.3). This technique builds up a histogram of the fluorescence decay by gathering photons in a very narrow time-gate, the position of this time gate can then be shifted in order to generate decay data. Unlike TCSPC, time-gated FLIM can be used in high light intensity situations, with count rates only limited by the speed of the detector and the data acquisition hardware. However, it is not very photon efficient, rejecting all photons out with the region of interest. There are several ways to address this issue including sampling into multiple gates simultaneously or using a smaller number of wider gates; however, as the number of gates used is reduced the ability to resolve complex multi-exponential decays also diminishes.

Time-gated FLIM is typically achieved using a detector such as a gated image intensifier associated with a sensitive CCD or CMOS 2D detector. Such a system achieves electron multiplications between multi-channel plates situated in front of the imager. Time-gating is achieved by providing

ϕ = Phase

m = modulation factor = $\dfrac{B/A}{b/a}$
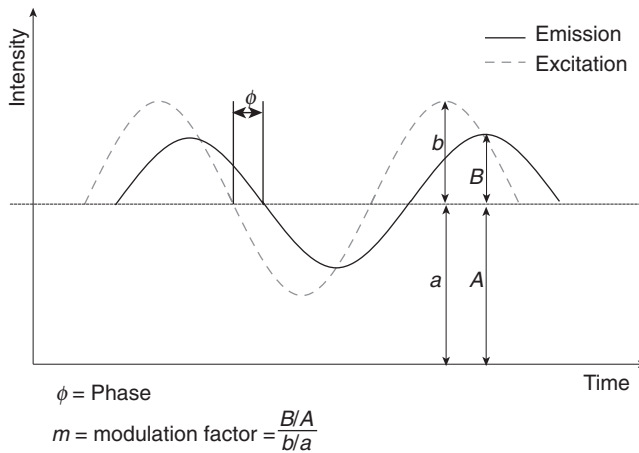
11.4 Frequency domain fluorescence lifetime principle.

a short electrical pulse which enables the multi-channel plates only during the period of interest.

Figure 11.4 shows the principle of frequency domain fluorescence lifetime estimation. A sample is excited through a sinusoidal illumination source (usually an LED). The fluorophore will emit at the same frequency as the excitation source, shifted in time by the excited-state lifetime. Two lifetime values, phase (τ_ϕ) and modulation (τ_m) can be derived from measurements of the phase shift and the modulation depth, respectively. The phase shift (ϕ) between the excitation and emission signals is a function of the excitation frequency (ω), and varies over a range 0 to 90 degrees across excitation frequencies. The modulation lifetime can be calculated from the difference between the emission and excitation amplitudes (Equation 11.7). The modulation parameter (m) is calculated from the average excitation intensity (I_{ex}), the excitation amplitude (A_{ex}), the average emission intensity (I_{em}), and the emission amplitude (A_{em}). The modulation parameter m varies over 0 to 1 across excitation frequencies according to Equation 11.6.

$$\tau_\phi = \frac{\tan(\phi)}{\omega} \tag{11.5}$$

$$m = \frac{A_{em} I_{ex}}{I_{em} A_{ex}} = \frac{1}{\sqrt{1 + \omega^2 \tau_m^2}} \tag{11.6}$$

$$\tau_m = \frac{1}{\omega} \sqrt{\frac{1}{m^2} - 1} \tag{11.7}$$

For fluorophores with mono-exponential lifetimes, the phase and modulation lifetimes will be identical. However, where the lifetime is composed of

multiple (n) decay components, the phase and modulation lifetimes require measurements performed at n different modulation frequencies (Lippitsch and Draxler, 1993).

11.2.4 Alternative FLIM techniques

A number of alternative methods exist for conducting lifetime measurements. These include the use of streak cameras, up-conversion methods and stroboscopic excitation.

Using a streak camera, very high timing resolution can be obtained (in the order of pico-seconds) (Krishnan *et al.*, 2003). Traditionally, streak cameras operate by sweeping incoming photons across the detection plane of the image sensor using high-voltage deflection plates. The resulting image contains temporal information in the axis across which the input photons were swept. From this, lifetime data can be recovered as the system also controls the speed and distance. Due to the requirement for high-voltage deflection plates, streak camera based lifetime analysis methods are not suitable for miniaturisation. However, recent work has seen the development of solid-state streak cameras, capable of high-frame rate operation (Kleinfelder *et al.*, 2005). In this work, temporal information is obtained by sampling the output of conventional integrating CMOS photodiodes on to on-chip capacitors. At present, this method can achieve a temporal resolution of 10 ns.

Fluorescence lifetime analysis using up-conversion methods offers the ultimate time-resolution performance, defined by the pulse width of a laser source (femto-seconds for a Ti:sapphire laser). The excitation light source is used as the input to an up-conversion crystal which only produces an up-converted version of the fluorescence emission during these laser input pulse. Through the use of a high-pass filter on the crystal output, only very short periods of the fluorescence decay can reach the detector. The time at which this short gate occurs can be altered through the use of an optical delay line. The need for two laser sources, an up-conversion crystal and an optical delay line (that provides just 1 ns of delay for every 30 cm of length) means that up-conversion is not an appropriate method for implementation in a micro-system for time-resolved fluorescence analysis.

Finally, stroboscopic excitation (Matthews *et al.*, 2005) greatly simplifies the design of the detection system, using the excitation source as the key system element for lifetime measurements.

11.2.5 Applications of FLIM

Fluorescence lifetime provides a convenient way to distinguish between different fluorophores in an assay if these have overlapping emission

spectra. Different biological and chemical phenomena can be observed by selecting a fluorophore that has a lifetime response that is sensitive to environmental effects. These include:

- pH – Many fluorophores demonstrate lifetime changes in response to varying pH. Examples of using FLIM as an indicator of pH changes can be found in Sanders *et al.* (1995) and Hanson *et al.* (2002).
- Ion concentration – Ionic quenching of fluorophores can be used to monitor changes in ion concentration, this is particularly important in the study of neuronal systems where changes in Ca+ or Cl– concentration can be observed. Probes such as Quin-2 demonstrate dramatic changes in lifetime in response to changes in local Ca2+ concentration (Lakowicz, 2006).
- Oxygen sensing – For many fluorophores, oxygen is an effective quencher. Therefore, fluorescence lifetime can be used to sense local changes in oxygen levels within a fluorophore's microenvironment. For instance, in Zhong *et al.* (2003) changes in the fluorescence lifetime of ruthenium based dyes were monitored in order to measure variations in oxygen levels within living cells.
- Förster resonance energy transfer (FRET) – The interaction and proximity of two complimentary fluorophores can be assessed using FRET. Lifetime changes occur as the distance between two fluorophore molecules (on an angstrom scale) varies. Energy transfer occurs between the two fluorophores and hence changes in the lifetime of the acceptor fluorophore.
- Viscosity – Fluorophores which demonstrate a high degree of internal flexibility (often referred to as molecular rotors) can be used to monitor changes in solvent viscosity.
- Explosives sensing – Thin-film conjugated polymers exhibit fluorescence lifetime changes in the presence of nitro-aromatic vapour. A handheld explosive sensing device using CMOS sensors has been demonstrated (Wang *et al.* 2011).

Many other FLIM applications exist including fluorophore aggregation, proximity to metal and tracking (Lakowicz, 2006).

11.3   Complementary metal-oxide-semiconductor (CMOS) detectors and pixels

11.3.1 Introduction

Active research on CMOS sensors for fluorescence lifetime started around 2005 spurred by the emergence of new detectors such as SPADs or the availability of low noise pinned-photodiodes from consumer CMOS image

sensors (Theuwissen, 2008). Prior development of very high speed image sensors based on gated CMOS photodiodes (Kleinfelder *et al.*, 2004, 2005) or CCDs with local transfer gates (Etoh *et al.*, 2003) demonstrated that images could be captured at MHz rates under high illumination fluxes. Lifetime imaging places the additional demands of high sensitivity due to very weak optical signals emerging from microscopic samples and decays in the nanosecond to a few 100 picoseconds range. There are strong similarities between the capture of fluorescence lifetime decays and the optical pulse round-trip delays for time of flight (TOF) ranging or 3D imaging. CCD or CMOS cameras for the latter function have been under development since the mid 1990s, only more recently being applied to FLIM (Esposito *et al.*, 2006).

CMOS sensors for FLIM have sought to increase the light collection efficiency of gated-image intensifier (GII) or the frame rate of scanning PMT based systems. The primary goal has been to improve imaging performance towards a more optimal tool for biomedical research. As the sensor will normally be placed within an already expensive and bulky microscope and laser system, the low cost and miniaturisation afforded by CMOS sensors compared to GII and photomultiplier tube (PMT) is of less significance.

Non-imaging applications of fluorescence lifetime are an important direction of research for CMOS devices, for example chemical sensors for oxygen or explosives and fluorescence sensors for DNA or protein analysis (Schwartz *et al.*, 2008a; Maruyama and Charbon, 2011). The full advantage of CMOS technology is leveraged here; small form factor, robustness, massively parallel sensing, low power operation and on-chip computation. Fully functional, autonomous fluorescence lifetime microsystems are emerging where the laser, microscope and PMT of FLIM systems are replaced by low cost LEDs, laser diodes, contact optics and microfluidic channels and a CMOS sensor chip (Khan *et al.*, 2009; Rae *et al.*, 2010). The following sections will look at the detectors and pixel performance of CMOS implementations of the three main lifetime sensing approaches: frequency, gating and TCSPC.

11.3.2 Frequency domain lifetime pixels

The recovery of the lifetime from the phase shift (φ) has been accomplished by techniques inspired by homodyne or direct conversion receivers (Razavi, 1997). A phase fluorometric oxygen sensor has been realised in 1.5 µm CMOS (Chodavarapu *et al.*, 2005) based on a large 16 × 16 phototransistor array measuring 720 µm × 720 µm (Fig. 11.5(a)). The choice of a phototransistor has been made to increase the current level from the detector at the expense of a relatively low bandwidth in the few hundred
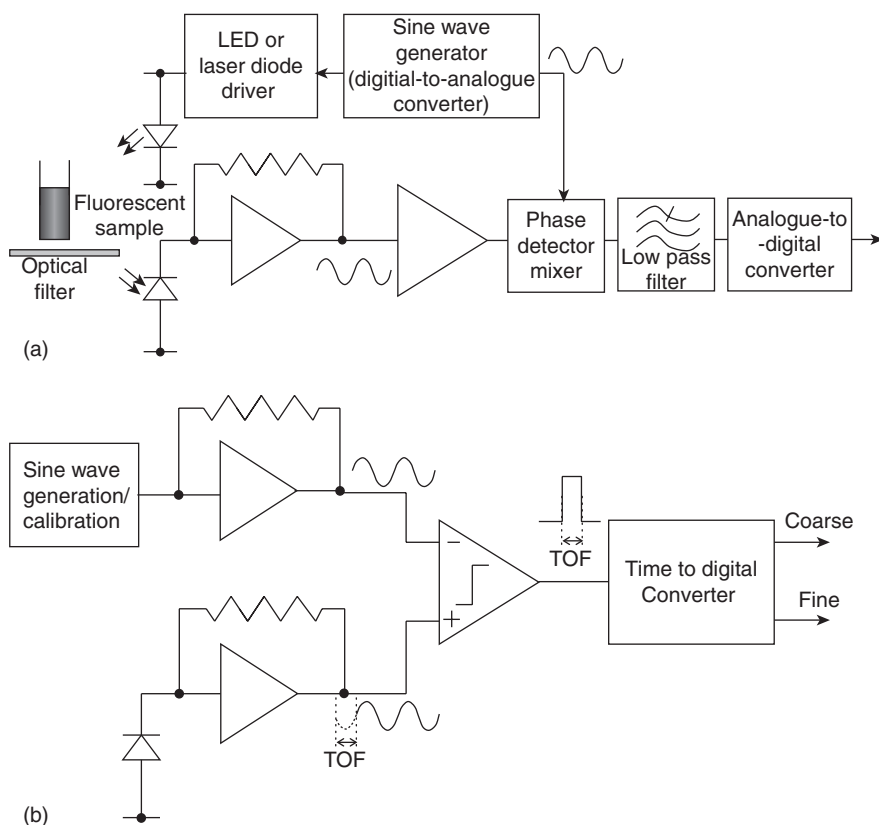
11.5 Phase-fluorescent sensor (a) phase demodulation pixel (b) delay digitisation pixel.

kHz range. This is acceptable for the few microsecond lifetime of the oxygen-sensitive ruthenium complex [Ru(dpp)3]2+ doped xerogel-film. The phototransistor peak responsivity of 675 nm has been chosen close to the emission of the xerogel of 595 nm. This sensor is the first to integrate the entire receiver consisting of current-to-voltage converter, amplifier, bandpass filter, and phase detector.

A similar system has been realised in a modern 65 nm CMOS process with direct digital phase measurement (Guo and Sonkusale, 2011a). They replace the phase detector by a comparator and time to digital converter (TDC), measuring phase directly as the time offset between the excitation signal and amplified, quantised fluorescence signal (Fig. 11.5(b)). The same team has re-used this architecture in the form of a 32 × 32 pixel image sensor for frequency domain FLIM (Guo and Sonkusale, 2011b). Passive pixels are employed with a 50 μm pitch and a 67% fill factor. The outputs of each pixel

are scanned to row-level transimpedance amplifiers (TIA) and comparators generating a zero crossing input to a global TDC. The TDC has a 110 ps temporal resolution over a 414 μs dynamic range. As only one pixel can be multiplexed to the TDC at any time, the frame rate of such an imager will be severely limited. It is unlikely that the TIA/TDC circuitry can be integrated at pixel level to allow this technique to be parallelised because of fill-factor restrictions.

A simpler demodulation approach which is amenable to parallel implementation has been presented in a recently developed CCD sensor for frequency domain FLIM (Zhao *et al.*, 2012). This is the first commercial solid-state camera development for scientific FLIM. It comprises 213 × 212 pixels at 17 μm pitch and 44% fill factor. The pixel integrates both phases of the modulated fluorescence simultaneously by directing accumulated photocharge via dual transfer gates from the photogate (PG) to storage gates (STG).

The operating principle of this imager is very similar to that of various time of flight image sensors which have been demonstrated in CMOS implementations (Lange and Seitz.*,* 2001; Oggier *et al.* 2004; Kawahito *et al.*, 2007) and demonstrated to be suitable for FLIM (Esposito, 2011). As shown in Fig. 11.6, fields under the transfer gates created by modulated voltages divert photo-generated carriers towards the storage gates where they are accumulated. The modulation voltages are synchronous with the emitter drive signals, creating a direct demodulation of the reflected, delayed light waveform at the two storage gates. The sensitivity of the modulated electron multiplied (MEM)-FLIM sensor compares well to an image intensifier solution with a duty cycle of about 50% when recording a single phase image. Fixed green fluorescent protein (GFP) cells, and GFP-actin stained live cells have been studied with this camera, demonstrating few nanoseconds lifetime resolvability at 25 MHz excitation frequency.

11.3.3 Time-gating pixels

Time gating is generally implemented using the two timing schemes shown in Fig. 11.7(a). The sliding window approach employs a single gate which is moved progressively across the lifetime decay, often at intervals set by the gate delay of an FPGA or on-chip delay-locked loop (a few 100 picoseconds). As the gates are overlapping by a small interval, the fluorescence decay characteristic can be reconstructed by differencing successive gated outputs (Patounakis *et al.*, 2006; Mosconi *et al.*, 2006). Many time gates must be accumulated with a penalty of the relatively slow rate of acquisition; however, only a single sampling channel is required. A faster and more photon efficient approach is shown in Fig. 11.7(b). Here, multiple non-overlapping time gates are used to sample the fluorescent decay. The most

$F = 0°$

$F = 180°$

Blocking gate   Storage gate

Transfer gate   Photo gate
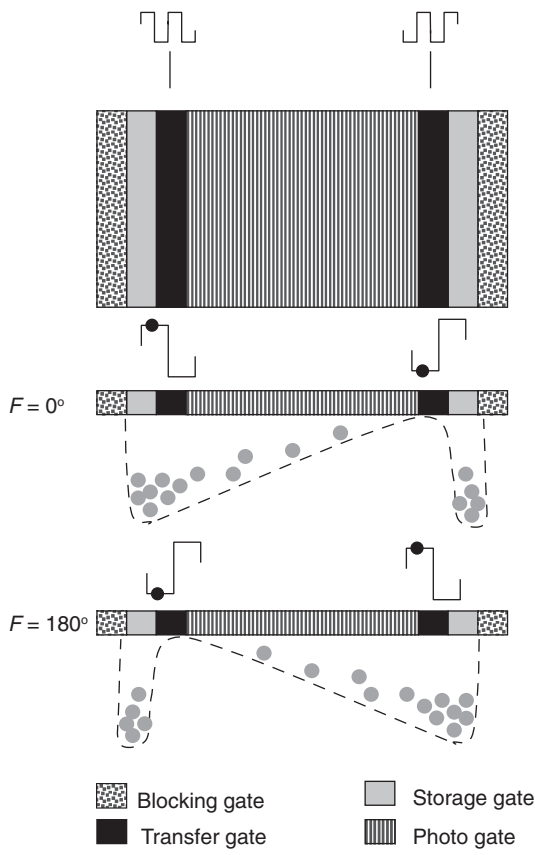
11.6 Modulated electron multiplied (MEM)-FLIM pixel and operating principle.

efficient scheme samples the fluorescence multiple times within each clock period, requiring multiple sampling channels and a larger pixel. A single channel may be used if the gates are accumulated in a sequential fashion, however photons falling out with the selected gate are then lost. Note that to equalise signal level in each gate, the duration can be chosen unequal. The simplest form of this scheme is called the two gate rapid lifetime determination (RLD) method (Ballew and Demas, 1989) which has been subsequently generalised to multiple gates (Grauw and Gerritsen, 2001).

An early implementation of gated image sensors for FLIM employed large 100 μm × 100 μm photodiodes in a 0.25 μm CMOS technology with a current mode readout (Patounakis *et al.*, 2006). The pixel shown in Fig. 11.8, samples the transient waveform induced by the photodiode through a gating signal TX on to a floating diffusion node. An innovative feature of this sensor was the improvement of signal to background ratio by draining
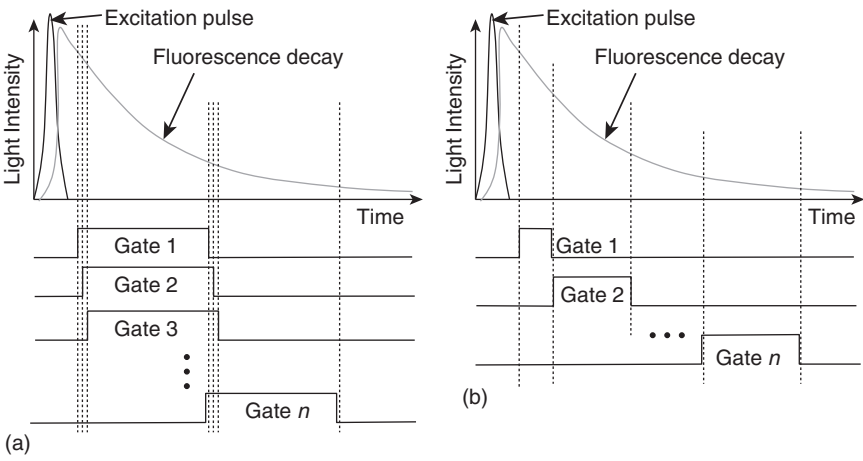
11.7 Gating schemes (a) sliding window (b) multiple gate.
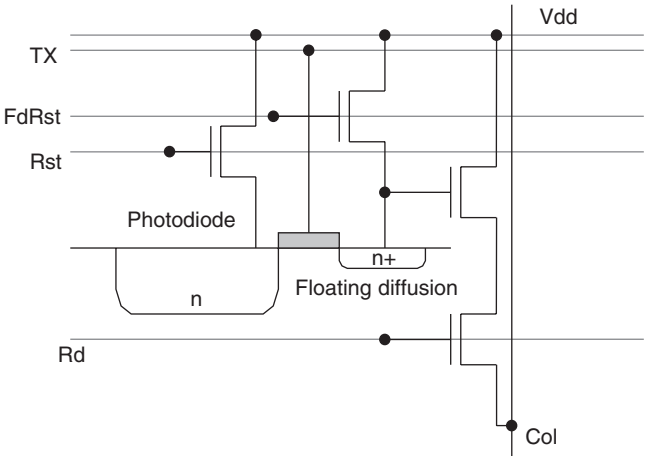
11.8 Gated photodiode.

carriers generated during the laser excitation through the reset transistor controlled by Rst. This reduces the requirements on optical filtering allowing contact sensing of DNA probes. The lack of gain in the photodiode and susceptibility to kT/C noise required long integration times (33 ms), and the relatively slow dynamics of the photodiode limit the technique to few nanosecond lifetimes.

Improved low light sensitivity has been achieved by employing pinned-photodiodes and correlated double sampling techniques from consumer CMOS imaging. Adapted pixel implementations were proposed to implement both sliding gate (Yoon *et al.*, 2009) and multi-gate (Bonjour

11.9 Gated pinned-photodiode implementations (a) multi-gate (b) sliding gate.

et al., 2011) schemes (Fig. 11.9). In both cases, an extra drain transfer gate has been added controlled by TXRst to reject the laser illumination and lessen the requirements on the excitation filter. The single output implementation uses a low leakage, pinned storage node to accumulate the signal charge and to perform correlated double sampling (CDS) (Fig. 11.9(a)). The two gate implementation (Fig. 11.9(b)) does not implement this storage node and so only performs double data sampling (DDS) to remove source follower offsets but not kT/C noise which dominates the overall noise (Bonjour *et al.*, 2011). The pixel is, however, more photon efficient and is capable of both frequency and time-domain FLIM. Fluorescence lifetimes down to 1 ns were resolved using pixels of 6.3 μm and 7.5 μm pitch and 256 × 256 resolution (Yoon *et al.*, 2009). The charge transfer speed of pinned-photodiodes is being improved through recent developments in TOF pixels with tapered doping levels or special transfer gate structures (Tubert *et al.*, 2009; Durini *et al.*, 2010). However, readout noise is still of concern at the extremely low light levels in typical FLIM experiments.

11.10 Digital gating pixel of a single photon avalanche diode.

SPADs provide both extremely high gain (>100 k for few micron diameter detectors) and fast response speed (~100 ps). They are excellent candidates for both short lifetimes and very low light imaging. Since the first detectors were proposed in CMOS around 2003 (Rochas *et al.*, 2003), researchers were quick to realise this potential and proposed a number of gated SPAD pixel realisations (Schwartz *et al.*, 2008b; Pancheri and Stoppa, 2008). A generic digital implementation is shown in Fig. 11.10. The gate input to the circuit both inhibits the SPAD by pulling the anode to ground whilst maintaining an output high state. Only if the SPAD triggers on a photon arrival when the quench p-type metal-oxide-semiconductor (PMOS) is enabled by setting gate = 0, will the pulse be transmitted to the output. The gate signal offers a way to modulate the sensitivity of the detector; however, it also possible to gate the collection of SPAD pulses by multiplexing the output to different counters or by selectively disabling the first toggle flip flop of ripple counters sharing a common input. These schemes have been successfully applied in a number of sensors; however, it has been difficult to achieve a small pixel pitch or high fill-factor because of the use of digital logic gates.

Two sensors composed of n-type metal-oxide-semiconductor (NMOS)-only, time-gated SPAD pixels of 25 µm pitch with fill-factors of 4.5% and 20.8%, respectively have been proposed recently in 0.35 µm high voltage (HV) CMOS (Maruyama and Charbon, 2011; Pancheri *et al.*, 2011) (Fig. 11.11). The absence of PMOS transistors improves fill-factor by avoiding large spacings between the SPAD n-well and adjacent transistor n-wells. The penalty is that static power consumption is introduced to implement
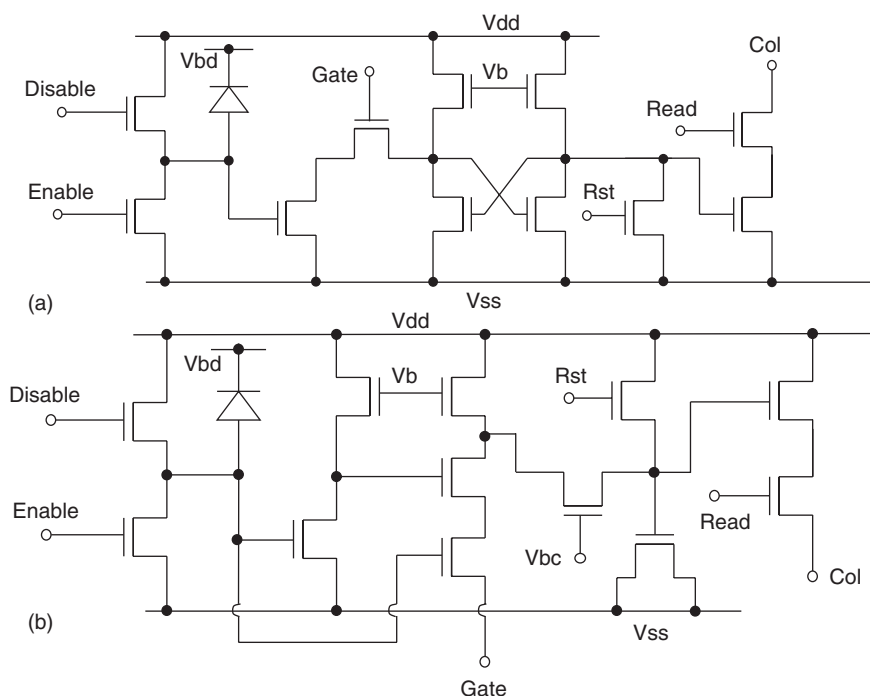
11.11 Gated analog NMOS-only SPAD pixels (a) SRAM pixel
(b) photon counting.

logic inversion. The pixel in Fig. 11.11(a) employs a single bit in-pixel static random access memory (SRAM) to indicate single photon arrivals during a gated access period. The memory is buffered to a column bus by a pull-down transistor. Bit-planes are read off the sensor using a sliding gate scheme and an FPGA receives the fluorescence lifetime data. SPAD gating has been used to achieve filterless contact fluorescence imaging of DNA hybridisation of few nanoseconds lifetime at a 200 ps resolution. This approach generates a large quantity of digital information since the pixel is unable to accumulate more than one photon at a time. The pixel in Fig. 11.11(b) addresses this issue by integrating photon arrivals as few millivolt discharge steps on a storage capacitor. A monostable composed of an NMOS-only inverter and not and (NAND) gate shortens the SPAD pulse to the enabled transistor allowing a current source transistor to draw a charge packet from a storage capacitor. The few mV per photon sensitivity of the pixel is above the thermal noise floor of the source follower and readout circuitry and allows single photon counting. A few hundred photons may be counted in this way. Gating is also possible by modulating the source voltage of the NAND. A minimum gate of a nanosecond is demonstrated

as well as photon shot noise limited performance. Fluorescence lifetime images of quantum dots were used to demonstrate the sensor. This technique shows enormous promise for future high resolution FLIM with SPADs.

11.3.4 Time-correlated single photon counting pixels

TCSPC requires single photon detection and at present can only be accomplished with SPAD detectors and pulsed laser sources. In this case, the pixels must be capable of resolving the times of arrival of single photons with respect to a laser synchronisation signal. The first implementations of time-resolved SPAD imagers employed off-chip timing circuits or multiplexed column-parallel time-to-digital conversion (Niclass *et al.*, 2005, 2008). These solutions were not sufficient for low-light imaging due to the loss of photons at unaddressed pixel sites. Fully-parallel time-resolved sensor architectures were researched in the MegaFrame EU project where a number of 50 μm pixel circuits were designed (Richardson *et al.*, 2009b; Stoppa *et al.*, 2009; Gersbach *et al.*, 2012). The small pitch, high throughput and moderate time resolution of these pixels were a significant step forward from existing large, power-hungry TDC circuits.

The MegaFrame pixels operate in reverse START–STOP mode whereby they are started by a single photon event from the SPAD and stopped by a synchronous clock from the pulsed laser. Two in-pixel TDC approaches were proposed (Richardson *et al.*, 2009b; Gersbach *et al.*, 2012) distinguished by adoption of an internal clock (IC-TDC) or an external clock (EC-TDC) to generate photon time of arrival estimates. The TDC-IC implemented an in-pixel gigahertz gated ring oscillator, clocking an n-bit ripple counter generating coarse photon arrival time estimates (Fig. 11.12(a)). The frozen internal state of the ring provides fine time estimates of single inverter delays. Two ring oscillator implementations were studied; a static approach based on a chain of tri-stateable inverters, and a dynamic approach based on a chain of inverters connected by pass-gates. There is a risk of metastability at moments where the ring oscillator clocks the ripple counter, which is avoided using hysteresis in the clock input stage. IC-TDCs consume power only when activated by an impinging photon due to the reverse START–STOP scheme, resulting in excellent power efficiency at low light levels.

The EC-TDC generates coarse time estimates by means of an n-bit counter, clocked by a globally distributed high frequency clock. (Fig. 11.12(b)). The propagation delay of the SPAD pulse through a buffer chain generates fine time estimates. The next rising edge of the STOP clock after a SPAD event freezes the state of inverter chain and the coarse counter. A thermometer decoder then converts the buffer chain state into a k-bit binary number. The power consumption of EC-TDCs is almost invariant
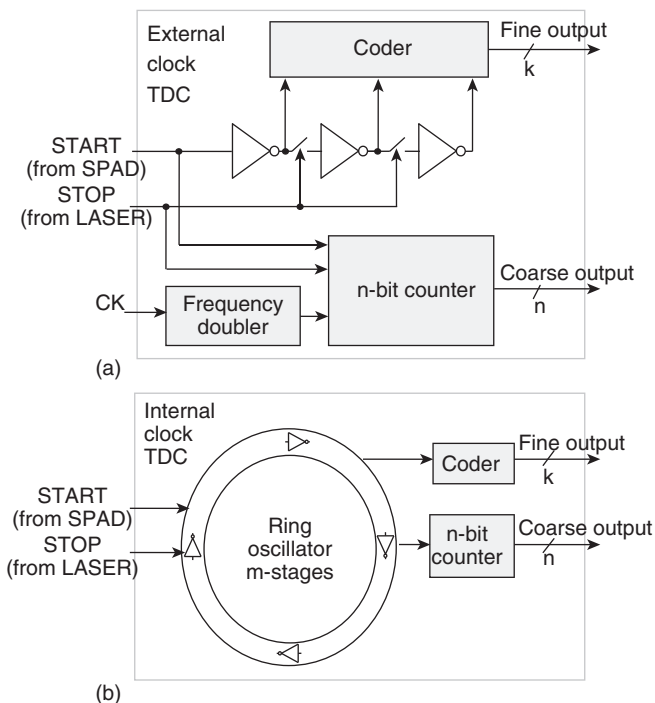
11.12 MegaFrame TDC pixel architectures (a) external clock (b) internal clock.

at all illumination levels. This architecture provides good global time resolution accuracy. Characterisation results from a 160×120 IC-TDC array have shown good TDC uniformity (Veerappan *et al.*, 2011).

11.4   FLIM system-on-chip

11.4.1 Introduction

FLIM, when combined with FRET techniques, can offer capability of imaging protein–protein interactions in living cells and facilitate scientists to uncover disease mechanisms (such as cancers) in early stages. The complexity, high cost and highly specialised knowledge required for the operation/acquisition has limited the prevalence of FLIM instrumentation. There have been efforts in integrating inexpensive illumination sources (such as LEDs), solid-state detectors capable of single-photon detection, and even lifetime imaging algorithms in a chip (Rae *et al.*, 2010; Tyndall *et al.*, 2012). These all solid-state technologies can significantly reduce the

size and cost with user-friendly interfaces facilitating the spreading of FLIM instrumentation for life and biomedical sciences.

Wide-field FLIM systems

In conventional wide-field FLIM systems, photons are captured on a CCD-based camera. This technology has limitations, including the fact that the capture of photons on an individual pixel results in an accumulation of charge which must be amplified. This leads to a background 'readout noise' that, in addition to stochastic 'dark noise' (which is reduced by the cooling of the device) and the 'shot noise' (quantum nature of light), results in a significant threshold below which photons cannot be detected with confidence. There are different types of CCD devices. Microchannel plate (MCP) based CCDs require a high voltage, usually 600–900 volts, applied to the MCP. EMCCDs contain several gain registers between the end of the shift register and the output amplifier. For EMCCDs, any dark current remaining will be multiplied up along the readout path. That is why EMCCDs must be cooled down to –100° C, significantly increasing their cost (~£25–30K). CCD-based systems mainly suffer from readout noise generated by the electronic readout circuitry, which includes charge transfer noise, amplifier noise and quantisation noise emerging from digitising signals. In low light situations, readout noise (especially for EMCCDs) can deteriorate the signal-to-noise ratio (SNR) significantly. Some pixel binning techniques are provided in some commercial devices to reduce, but do not eliminate, the readout noise. Moreover, they result in poor spatial resolution. Additionally, ICCDs and EMCCDs, the microchannel plate (MCP) and the electron multiplying (EM) register suffer from significant ageing effects.

Confocal scanning FLIM systems

In a conventional confocal scanning FLIM experiment, a PMT capable of single-photon detection is used in combination with a TCSPC module. The lifetimes are calculated, usually using iterative Marquardt–Levenberg algorithms on a pixel-by-pixel basis, in which the bottleneck to achieving high-speed imaging is the acquisition. For example, to avoid local heating and photobleaching, the pixel dwell time is set to be $15.25\,\mu s$ and the sample is scanned hundreds of times, say 300 times, to accumulate enough photon counts. The total dwell time in a pixel would be about 4.5 ms, within which the lifetime can be calculated using software with ease. To increase the imaging speed, multi-channel PMT systems, although instrumentally intensive, have been introduced and commercially available. They, however, are quite expensive and bulky, and usually not easy to operate. The number of channels is usually below 60. Image scanning is still required, and the lifetime calculations can be conducted in PCs.

CMOS miniaturised PMTs

Unlike the low optical gain CCD based or CMOS imaging sensors, a SPAD detector, when biased at above its breakdown voltage, can be triggered by a single photon that results in a self-sustaining avalanche multiplication process. The gain of SPADs is so high that a simple buffer can convert a detected signal into a digital pulse without using complicated/noisy front-end amplifiers. The latest SPADs can easily achieve a photon count rate of several MHz (1MHz = 1 million photons per second) with a dark count much less than 100 Hz, even without using a cooling system (Richardson *et al.*, 2009a). The diameter of the active area is 8 μm with a fill factor of 2%. A microlens array can be placed on top of the SPAD array to recover the light loss (Veerappan, 2011). The SNR of SPADs is nearly shot-noise limited, an inherent property of light (Equation 11.8). Thus, SPAD detectors have no readout noise with its dark current noise usually much lower than it photon count rate, so the signal to noise ratio is only Poisson noise limited capable of operating at a high frame rate.

$$SNR_{SPAD} = \frac{G\eta\phi\tau}{\sqrt{(N_{shot})^2}}. \tag{11.8}$$

With single photon sensitivity, SPADs are highly suitable for photon-starved applications such as single-molecule detection. In essence, SPADs are miniaturised PMTs, another type of single-photon detector. Unlike bulky PMTs, however, SPADs can be fabricated into large arrays offering a great deal of parallelism (see Table 11.1). Besides, the photon count rate of the latest developed SPADs can easily exceed tens of MHz. The high throughput data available from the new SPAD systems, however, poses a major challenge in the design of the readout architecture.

A more important factor is that the SPAD can provide a timing jitter or timing resolution of tens of picoseconds, a specification impossible to achieve with CCD-based or CMOS image sensors (CIS). Timing jitter mainly depends on the time a photo-generated carrier spends from the absorption point to entering the multiplication region. Without using cooling systems and high-voltage power supplies, SPAD-based devices offer low cost and high system integration. The latest CMOS SPAD arrays contain on-chip lifetime calculation processors that are capable of compressing high-throughput data and generating high-speed lifetime images. The array size of the current SPAD chips is 160×128 and the pitch ('pixel' size) is 50 μm due to in-pixel 10-bit 55 ps TDC that allow recording the time tag of each, individually detected photon. The deadtime of SPADs is in tens of nanoseconds and results in a photon count larger than MHz, which is much higher than CCD devices can achieve. Larger SPAD arrays with a smaller pitch are in development. Recent advances in CMOS technology, on the

Table 11.1 Comparison of photon-capture devices

	CCD	ICCD[a]	EMCCD	CMOS imaging sensor (CIS)	Multi-channel PMT	CMOS SPAD
QE[b]	50–90%	10–50%	50–90%	~60%	20–40%	30–40%
Cooling	−40 °C	−40 °C	−100 °C		−40 °C	No
Cost	5–15 K	20–30 K	20–30 K		>50 K	2–4 K[e]
Timing resolution	>> 1 µs (gating)	2–10 ns (gating)	~ µs (gating)		~ 100 ps	~ 100 ps
Frame rate (Hz; frames per second)	100	10[d]	10–100	100	Confocal scanning; image size dependent	500,000
Pixel pitch[c]	7 µm	10 µm	8–15 µm	7 µm	Not applicable but some are several mm, pending application	20–50 µm depending on the readout design
Array size	1024 × 1024	1024 × 1024	1024 × 1024	2560 × 2160	16-channel TCSPC (time-correlated single-photon counting) bin width = 80 ps	160 × 128 higher resolution SPAD arrays are in development
Imaging capability	Not suitable for lifetime imaging unless gated intensifiers are applied	Lifetime imaging with gating	Lifetime imaging with gating	Lifetime imaging with gating	Scanning confocal lifetime imaging	Widefield or confocal scanning with both TCSPC and gating
Single-photon detection capability	No	No	No	No	Yes	Yes

[a]Based on Hamamatsu's GaAsP.
[b]QE: quantum efficiency.
[c]Based upon high-end devices.
[d]PI-MAX3 'near' video rate frame rate.
[e]Current development costs. Market costs will be competitive.

other hand, offer an alternative to traditional PMT based systems. The latest CMOS SPAD-based cameras can provide high quantum efficiency, low noise, and low cost with high timing resolution comparable to PMTs. The channel number can easily exceed 20k. The bottleneck to real-time imaging in such systems shifts towards data readout and lifetime calculation. The data throughput in CMOS SPAD arrays is usually massive, challenging the limited bandwidth of readout circuitry. Lifetime imaging becomes very compute-intensive, and it is impossible to achieve real-time using iterative curve-fitting algorithms. It is desirable to have on-chip or on-FPGA imaging processors compressing or pre-processing the data before sending all raw data into PCs.

11.4.2 CMOS imaging sensors with frequency-domain lifetime imaging/sensing

In the past, high-speed wide-field FLIMs are demonstrated using MCP based image intensifiers. MCP based FLIM systems, however, are expensive and bulky usually requiring high-voltage power supplies, and their timing resolution or noise performances are not satisfactory. Some nanosecond lifetime systems were proposed by directly modulating the gain of sensors (Mitchell *et al.*, 2002), but the modulation frequencies are only hundreds of kHz. (Esposito *et al.*, 2006) demonstrated an all-solid-state 124×160 FLIM camera based on the integration of buried-channel charge-coupled devices (BCCD), CMOS active-pixel sensor (APS) readout architectures and low cost LEDs. The pixel size is $40 \times 55\,\mu m^2$ with an optical fill factor of 17%. The modulation frequency can be up to 20 MHz. Lifetime calculations are performed in the frequency domain by observing the phase delay and the luminescence signal demodulation at different relative phases (Gadella *et al.*, 1993). Such setups require analog front-end amplifiers to amplify the signals which usually introduce extra noise deteriorating the SNR. A longer acquisition time can be used to minimise the effects of readout noises (Guo and Sonkusale, 2011b) and a low frequency of 1.2 kHz for oxygen sensing. These frequency-domain approaches allow low cost LEDs to be employed as the light source.

11.4.3 Latest CMOS fluorescence lifetime imaging/sensing systems

Scientific or laboratory confocal scanning FLIM systems usually contain a single or multi-anode PMT connecting to a TDC with a timing resolution of several picoseconds inside the TCSPC module (Becker, 2012). The major criticism of TCSPC-based FLIM systems is low acquisition speed due to the need for scanning. A limited number of PMT plus TCSPC multi-channel systems have been commercially available to improve the imaging speed,

but they are bulky and expensive with image scanning still required. In usual laboratory setups, gated MCP based CCDs are used with wide field excitation to provide high-speed lifetime imaging. Such systems, however, are also expensive. The recent advance of CMOS technologies, on the other hand, can provide much cheaper solutions. Large scientific grade CIS arrays can be fabricated with on-chip amplifiers integrated in the same chip. With simple gating techniques, a 256×256 image sensor for wide-field FLIM applications in 0.18 μm CMOS technology with a pinned photodiode process option has been reported (Yoon *et al.*, 2009), where an in-pixel charge transfer mechanism was applied to calculate the fluorescence lifetimes. On-chip amplifiers are included to pre-process the signals. If the charge transfer lifetimes can be negligible as we expect from the devices shifting toward a more advanced process, the systems can be used to monitor lifetimes of nanoseconds. Huang *et al.* (2009) reported a FLIM system using CMOS APS arrays with a pixel pitch of 50 μm and a fill factor of 23%. Another 256×256 sensor array with a pitch of 6.3 μm and a fill factor of 14% fabricated in a standard CMOS process with a buried photodiode (BPD) option for FLIM applications was also reported (Bonjour *et al.*, 2011). Since CMOS photodiodes have no inherent gain, on-chip amplifiers are required which introduce extra readout noises. To achieve single-photon detection with a timing resolution of picoseconds, PMT-like single-photon detectors can be fabricated in low cost CMOS imaging process using avalanche photon diode structures. Latest CMOS SPAD detectors can offer comparable timing resolutions, noise performances and quantum efficiencies as the state-of-art PMTs (Richardson *et al.*, 2009a; Gersbach, 2009). The integrated chip of low noise SPAD plus TDC arrays can act as a miniaturised version of multichannel PMT plus TCSPC module (Richardson *et al.*, 2009b; Veerappan *et al.*, 2011), and video-rate fluorescence lifetime imaging has been successfully demonstrated (Li *et al.*, 2011). Compared with the multichannel PMT plus TCSPC systems, the latest SPAD-TCSPC arrays can offer a greater flexibility for both wide-field and multifocal multiphoton imaging. Gating rather than TCSPC techniques can be applied (Schwartz *et al.*, 2008b; Gola *et al.*, 2011; Pancheri and Stoppa, 2008; Pancheri *et al.*, 2011; Li *et al.*, 2012) to enhance the fill factor of SPAD arrays. On-chip fluorescence lifetime sensing and innovative circuitry have been integrated to avoid the pile-up effects and generate lifetimes in real-time (Tyndall *et al.*, 2012).

11.4.4 Time-domain fluorescence lifetime embedded algorithms

Unlike confocal scanning FLIM systems using iterative nonlinear least square methods (NLSM) to calculate lifetimes in a pixel-by-pixel manner,

imaging sensor arrays can be gated with different time delays, and the delayed snapshots can be recorded in one-shot (Elson *et al.*, 2004) or in parallel gated counters (Grauw and Gerritsen, 2001) providing real-time lifetime imaging. In the past, gating techniques are usually used in wide-field ICCD cameras by applying a gating signal to turn on/off the intensifiers. The circuitry generating the gating signals (having a voltage swing of hundreds of volts with a time resolution of nanoseconds) is usually a stand-alone component and brings an extra cost. In this respect, gated CMOS based cameras can provide a low cost, high-speed solution by integrating sensors and circuitry for clock generation and signal processing on a single chip (Yoon *et al.*, 2009; Bonjour *et al.*, 2011). Also, to study highly dynamic biological systems *in vivo* in real time such as following single molecule behaviour at replication forks requires fluorescent probes that photoswitch or photoconvert in combination with highly efficient detectors. Specifically, imaging integration times need to be in the milliseconds, or even an order of magnitude lower depending upon the velocity of the molecule, in order to avoid blurring. Limiting imaging integration times, however, causes problems since detector noise from CCD cameras is normally much higher than the signal generated when using video-rate imaging (even in total internal reflection fluorescence (TIRF) mode) (Lenn and Leake, 2012). Several pixels can be binned to reduce readout noises, but at the expense of the spatial resolution. To achieve low noise while high-speed imaging, SPAD-based FLIM systems have been successfully demonstrated in low cost CMOS imaging process (Maruyama and Charbon, 2011; Veerappan *et al.*, 2011). The frame rate a SPAD system can achieve is usually higher than hundreds of kHz. The high throughput signals pose significant challenges in the readout design, and on-FPGA or on-chip digital processing units are required for lifetime imaging.

Gating methods

Non-iterative, direct FLIM algorithms such as gating methods have been used in ICCDs and CMOS sensors with wide-field excitation to achieve real-time imaging. The simplest gating method uses only two time gates to generate an average lifetime (Ballew and Demas, 1989; Rae *et al.,* 2010), denoted as two-gate RLD, where the sensor array is gated with two different time delays, and the delayed intensity snapshots can be recorded in one-shot or sequentially. RLD can provide an optimised SNR when the gates are properly designed (Chan *et al.,* 2001; Li *et al.*, 2012) and the fluorescent emission is a mono-exponential decay. Gating techniques, however, are sensitive to background noise (Li *et al.*, 2009), and readout noises are usually difficult to calibrate if CCD based systems are used. A CMOS SPAD can convert detected photon events into digital pulses un-interfered from

readout noises. The photon counting rate of SPADs can be up to tens of MHz with a dark count less than 100 Hz at room temperature. Without readout noises deteriorating the SNR, the camera can operate at a frame rate higher than hundreds of kHz capable of capturing highly dynamic biological systems. Pancheri and Stoppa (2008) and Benetti *et al.* (2010) used four gated counters in parallel for each binned pixel in their SPAD arrays to collect photons. They used four non-identical gates and applied the NLSM to solve bi-exponential decays. In fact, end-users can design non-overlapping or overlapping gates in such systems and use fast direct algorithms proposed in Sharman and Periasamy (1999) to calculate life-times. Although direct algorithms do not provide results with a precision comparable to the NLSM, they are much quicker. With proper designed gates (number of gates larger than 4) and using NLSM, 'near-ideal' lifetime imaging is achievable (Grauw and Gerritsen, 2001), but only for single-exponential decays. Although not accurate in describing real biological behaviours, single-exponential decay models are useful to contrast different types of fluorophores. For diagnostic applications, obtaining high-speed lifetime contrast is probably more important than determining the absolute values of lifetimes (Elson *et al.*, 2004).

Consider a 128×128 gated SPAD array, if the array is working at a frame rate of 1 MHz with a laser pulse of 20 MHz used as the light source, there will be at most 20 photon events (5-bits) per pixel per frame (see Table 11.2). The data throughput is $64 \times 128 \times 2 \times 5 \times 1e6 = 82$ G bit/s (assuming two pads for a column). Digital input/output (I/O) pad can readout $64 \times 2 \times 5 \times 1e6 = 640$ Mbit/s in the latest CMOS technology, but high frame-rate readout does pose a significant challenge in data transfer. The data must be pre-processed before sent to a PC. For example, if the two-gate RLD is used, we can readout two counts N_1 and N_2 for each pixel per 1000 frames (1 ms). Photon count data are first readout off the chip, and 256 accumulators

Table 11.2 Data throughput and hardware required for 128×128 gated SPAD arrays

Readout frequency	Data throughput (Gbit/s)	Size of registers on FPGA (bit) 2-gate RLD	Memory required (bit)	High-speed link (FPGA-PC)
1 MHz	82 Gb/s	10	164 k	N/A
100 kHz	13 Gb/s	16	262 k	N/A
10 kHz	1.8 Gb/s	22	360 k	USB-3/PCIe
1 kHz	250 Mb/s	30	500 k	USB-2
100 Hz	30 Mb/s	36	590 k	USB, USB-2

are used on the FPGA to store these counts. The size for the N_1 and N_2 registers (usually on a FPGA) needs to be 15-bit. The data throughput becomes $128 \times 128 \times 2 \times 15 \times 1e3 = 480$ Mbit/s. A USB-2 link would be enough to accommodate such data throughput. The total memory required on the assembly is only $128 \times 128 \times 15 \times 2 = 480$ kbit. It is not a challenge for a multi-core PC to conduct 128×128 divisions (18-bit) in 0.01 second; although simple, gated cameras do not provide single-photon detection capability. The fluorescence decays are usually undersampled, and it is therefore challenging to extract the lifetimes and amplitude coefficients from multi-exponential decays (Becker, 2012).

Non-gating embedded algorithms

Gating systems usually contain a limited number of parallel gated counters due to hardware complexity. Grauw and Gerritsen (2001) suggested that 'near ideal' lifetime imaging is possible through optimally designed gating, but it is not easy to reach such optimisation as the distributions of fluorescent emission in most biological samples are usually unknown and are multi-exponential. For most biological FLIM experiments, it is desirable to record raw arrival time data for detailed scientific analysis (called raw mode hereafter) and also provide a high-speed preview mode. The camera can work in the preview mode for fast locating a region of interest, and switch to the raw mode with or without scanning for detailed observations. The latest CMOS SPAD array with in-pixel 10-bit 55 ps TDCs can offer such functionalities (Richardson et al., 2009b; Veerappan, 2011). SPAD arrays can contain in-pixel TDCs or column TDCs to generate time-resolved arrival time data and on-chip lifetimes (Tyndall, 2012) making them emulate PMT based TCSPCs while providing extra parallelism. Consider a 128×128 SPAD plus 10 bit TDC array with a laser pulse of 20 MHz as the light source, and suppose the imager works at a frame rate of 1 MHz. To avoid pile-up effects (Becker, 2012), there is at most a photon event per pixel per frame. Limited by pile-up effects, the photon count for SPAD-based TCSPC systems is less than the previous gated SPAD systems in strong illumination. The data throughput is $128/2 \times 10 \times 1e6 = 640$ M bit/s per pad (64 pixels per pad). Similar to the argument above, the data must be pre-processed before being sent to a PC. Instead of using gating algorithms, we proposed a new algorithm called the integration for extraction method (IEM) (Li et al., 2008; Li et al., 2009), and the calculated lifetimes are

$$\tau_{IEM} = \frac{h \cdot \sum_{j=0}^{M-1} (\bar{C}_j N_j)}{N_0 - N_{M-1}}, \qquad\qquad [11.9]$$
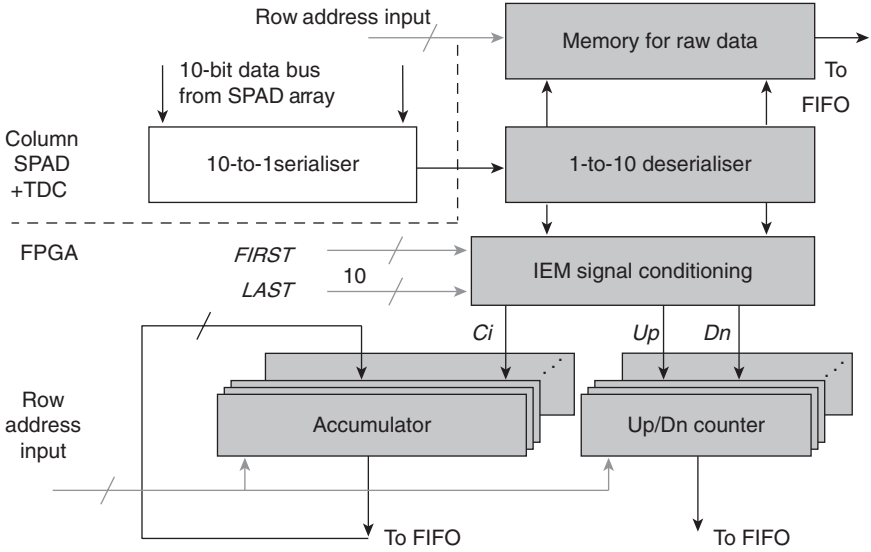
11.13 On-FPGA IEM for a column of CMOS SPAD plus TDC pixels.

where $\bar{C} = [1/3, 4/3, 2/3, \cdots, 4/3, 1/3]$ from the Simpson's integration rule, h is the TDC bin width, and N_j is the count number in the jth time bin. The IEM only requires simple arithmetic additions and divisions making it hardware-friendly. With this arrangement, the data throughput can be greatly reduced. From Equation 11.9, a simple accumulator and an up-down counter are required, and they can be easily implemented on-FPGA (Li et al., 2010, 2011) or on-chip allowing real-time lifetime imaging.

Figure 11.13 shows an on-FPGA IEM for a column of SPAD plus TDC pixels. Arrival time data are first transferred on to an FPGA through the serialiser, and the deserialiser can send the data to the FIFO directly or to the signal conditioning block where Ci, Up and Dn are generated to increment the accumulator and Up/Dn counter for the numerator and denominator of Equation 11.9. The inputs $FIRST$ and $LAST$ can be set by end users to define a measurement window.

The IEM provides some important features. Compared with RLD, IEM is much less insensitive to background noise (Li et al., 2009). Another important feature is that IEM can be used to determine the FRET efficiency without resolving complex histograms. Assume the fluorescence histogram is $f(t) = \sum_{i=1}^{k} a_i e^{-t/\tau_i}$, where a is the amplitude proportion and τ is the lifetime. Its amplitude weighted and intensity weighted lifetimes are $\tau_A = \left(\sum_{i=1} a_i \tau_i \right) / \left(\sum_{i=1} a_i \right)$ and $\tau_I = \left(\sum_{i=1} a_i \tau_i^2 \right) / \left(\sum_{i=1} a_i \tau_i \right)$, respectively. The

former is mostly used to calculate the FRET efficiency (Wu and Brand, 1994), whereas the latter is used for cases of dynamic quenching (Fišerová and Kubala, 2012; Sillen and Engleborghs, 1998). From the definition of the IEM (Li *et al.*, 2008),

$$\tau_{IEM} = \frac{h \cdot \sum_{j=0}^{M-1}\left(\bar{C}_j N_j\right)}{N_0 - N_{M-1}} \sim \frac{\int_0^T f(t)dt}{f(0)-f(T)} = \frac{\sum_i a_i \tau_i\left(1-e^{-T/\tau_i}\right)}{\sum_i a_i\left(1-e^{-T/\tau_i}\right)} \sim \tau_A = \frac{\sum_i a_i \tau_i}{\sum_i a_i}$$

[11.10]

where T is the measurement window.

The lifetime calculated by the IEM is equal to τ_A if T is much larger than the biggest lifetime component. Without resolving all a_j and τ_j from the decay, IEM can directly calculate the τ_A and therefore the efficiency of the FRET.

Similar to the IEM, another lifetime embedded algorithm centre of mass method (CMM) (Li *et al.*, 2010) can be implemented in our SPAD plus TDC systems. The lifetime is the average arrival time of all detected photons;

$$\tau_{CMM} = \frac{\int_0^T tf(t)dt}{\int_0^T f(t)dt} = \frac{\sum_i a_i \tau_i^2\left(1-e^{-T/\tau_i}-e^{-T/\tau_i}\cdot\frac{T}{\tau_i}\right)}{\sum_i a_i \tau_i\left(1-e^{-T/\tau_i}\right)} \sim \left(\frac{\sum_{i=1}^{N_c}\bar{D}_i}{N_c}+\frac{1}{2}\right)h$$

[11.11]

where $\bar{D}_i$ is the 10-bit TDC output of the ith captured photon.

CMM can be used in confocal TCSPC systems considering non-ideal instrument response (Won *et al.*, 2011). On-FPGA CMM (Fig. 11.14) is slightly more complicated than the on-FPGA IEM, but it is a 'near ideal' estimator if the background noises are correctly calibrated and the decay profile is single-exponential (Li *et al.*, 2010). On-FPGA and on-chip CMMs have been implemented on a 32×32 CMOS SPAD camera and a 32×32 silicon photon multiplier, respectively to show video-rate lifetime imaging (Li *et al.*, 2011) and high-speed sensing (Tyndall *et al.*, 2012). For multi-exponential decays, the CMM generates an intensity weighted lifetime if the measurement window T is much larger than the biggest lifetime component.

11.5   Future trends

The cost, miniaturisation and robustness benefits of CMOS fluorescence lifetime sensors have been clearly demonstrated by various researchers in new applications such as contact imaging for microarray or lab-on-chip or in portable sensing (Charbon, 2008). To replace PMT or GII in existing
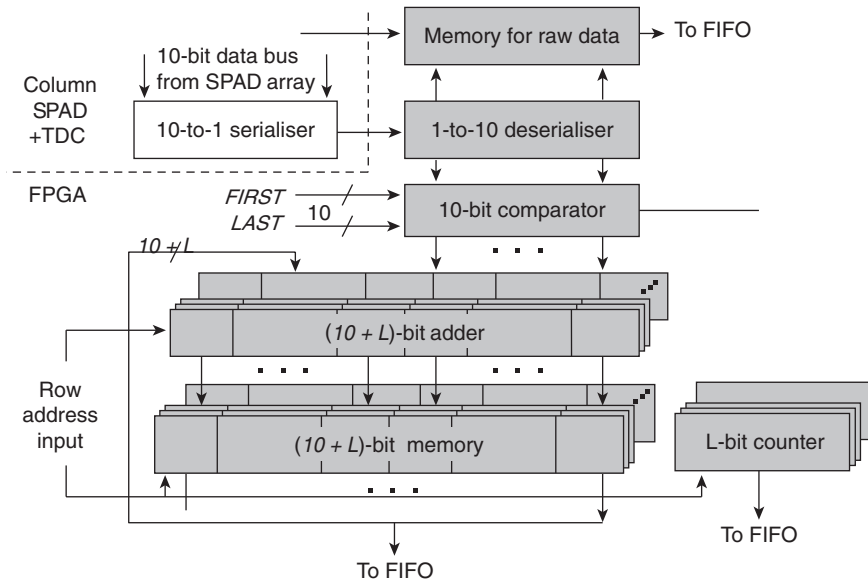
11.14 On-FPGA CMM for a column of CMOS SPAD plus TDC arrays.

applications such as microscopy, solid-state technologies must demonstrate higher sensitivity, spatial and time resolution, photon throughput and external quantum efficiency. Researchers have responded by innovating in detectors, pixel structures, readout techniques and on-chip photon processing. The first commercial imagers for frequency-domain lifetime imaging have recently been announced in CCD technology. In the same way that sCMOS imagers have demonstrated competitive low light imaging performance to EMCCD and intensifier technologies, CMOS can be expected to provide solutions for wide-field time-domain fluorescence lifetime. In particular, high resolution, high frame rate gated CMOS image sensors with microlenses and pixel sizes below 10 μm seem achievable goals. In this regard, new SPAD structures and pixel circuits promise a combination of speed and sensitivity, with unique capability to offer TCSPC single-photon imaging and auto-correlation. The convenience and low cost of solid-state technologies will undoubtedly transform fluorescence lifetime from an exotic laboratory procedure to an accessible and flexible sensing technique for the wider community of scientists and engineers.

11.6   Sources of further information and advice

A comprehensive textbook on fluorescence lifetime imaging techniques was published in 2005 by Wolfgang Becker entitled *Advanced Time-Correlated Single-Photon Counting Techniques* (Becker, 2005). Supplementary material

is found on the company website of Becker and Hickl (http://www.becker-hickl.com) and an excellent updated review article (Becker, 2012). This book covers all aspects of the technology and applications of time-resolved imaging; however, it was written just prior to the advent of CMOS sensor implementations. The standard reference in the wider area of fluorescence techniques is by J. R. Lakowicz entitled *Principles of Fluorescence Spectroscopy* (Lakowicz, 2006). A textbook edited by A. Theuwissen and P. Seitz entitled *Single Photon Imaging* (Theuwissen and Seitz, 2011) has appeared recently surveying all single photon imaging technologies including CMOS. Much useful information can be found on the websites of the various time resolved imaging equipment vendors, e.g., B&H, Picoquant, Andor, Lambert Instuments, LaVision, Hamamatsu, Olympus.

11.7   References

Ballew R.M. and Demas J.N. (1989), 'An error analysis of the rapid lifetime determination method for the evaluation of single exponential decays,' *Anal. Chem.*, 61, 30–33.

Becker W. (2005), *Advanced Time-correlated Single Photon Counting Techniques*, Berlin, Springer.

Becker W. (2012), 'Fluorescence lifetime imaging – techniques and applications', *J. Microscopy*, 247, 119–136.

Benetti M., Iori D., Pancheri L., Borghetti F., Pasquardini L., Lunelli L., Pederzolli C., Gonzo L., Dalla Betta G.-F. and Stoppa D. (2010) 'Highly parallel SPAD detector for time-resolved lab-on-chip', *Proc. SPIE*, 7723, 77231Q.

Bonjour L.-E., Singh A., Baechler T. and Kayal M. (2011), 'Novel imaging method and optimized demodulation pixels for wide-field fluorescence lifetime imaging microscopy', *IEEE Sensors*, 724–727.

Chan S.P., Fuller Z.J., Demas J.N. and DeGraff B.A. (2001). 'Optimized gating scheme for rapid lifetime determinations of single-exponential luminescence lifetimes', *Anal. Chem.*, 73(18), 4486–4490.

Charbon E. (2008), 'Towards large scale CMOS single-photon detector arrays for lab-on-chip applications', *J. Phys. D: Appl. Phys.*, 41, 094010.

Chodavarapu V.P., Shubin, D.O., Bukowski R.M., Titus, A.H., Cartwright A.N. and Bright F.V. (2005) 'CMOS mixed-signal phase detector for integrated chemical sensor systems', *IEEE Sensors*, 1068–1071.

Christenson M. and Sternberg, S. (2004), 'Fluorescence lifetime imaging microscopy (flim) lets biologists study cellular processes', *SPIE oemagazine*, 28–30.

Cubeddu R., Comelli D., D'Andrea C., Taroni,P. and Valentini G. (2002), 'Time-resolved fluorescence imaging in biology and medicine', *J. Physics D: Appl. Phys.*, 35, 61–76.

Durini D., Spickermann A., Mahdi R., Brockherde W., Vogt H., Grabmaier A. and Hosticka B.J. (2010) 'Lateral drift-field photodiode for low-noise, high-speed, large photoactive-area CMOS imaging applications', *Nuclear Instruments and Methods in Physics Research A*, 624, 470–475.

Elson D.S., Munro I., Requejo-Isidro J., McGinty J., Dunsby C., Galletly N., Stamp G.W., Neil M.A.A., Lever M.J., Kellett P.A., Dymoke-Bradshaw A., Hares J. and

French P.M.W. (2004), 'Real-time time-domain fluorescence lifetime imaging including single-shot acquisition with a segmented optical image intensifier', *New J. Phys.*, 6, 1–13.

Esposito A. (2011), 'Beyond range: innovating fluorescence microscopy', *J. Remote Sens.*, 4, 111–112.

Esposito A., Gerritsen H.C. and Oggier T. (2006), 'Innovating lifetime microscopy: a compact and simple tool for life sciences, screening, and diagnostics', *J. Biomed. Opt.*, 11, 034016.

Etoh T.G., Poggerman D., Kreider G., Mutoh H., Theuwissen A.J.P., Ruckelshausen A., Kondo Y., Maruno H., Takubo K., Soya H., Takehara K., Okinaka T. and Takano Y. (2003), 'An image sensor which captures 100 consecutive frames at 1,000,000 frame/s', *IEEE Trans. Electron Devices*, 50, 144–151.

Fišerová E. and Kubala M. (2012), 'Mean fluorescence lifetime and its error', *J. Luminesc.*, 132(8), 2059–2064;

Gadella T.W., Jovin T.M. and Clegg R.M. (1993), 'Fluorescence lifetime imaging microscopy (FLIM)-spatial-resolution of microstructures on the nanosecond time-scale', *Biophys. Chem.*, 48(2), 221–239.

Gersbach M., Richardson J., Mazaleyrat E., Hardillier S., Niclass C., Henderson R., Grant L. and Charbon E. (2009), 'A low-noise single-photon detector implemented in a 130 nm CMOS imaging process', *Solid State Electron.*, 53, 803–808.

Gersbach M., Maruyama Y., Trimananda R., Fishburn M. W., Stoppa D., Richardson J. A., Walker R., Henderson R. and Charbon E. (2012), 'A time-resolved, low-noise single-photon image sensor fabricated in deep-submicron CMOS technology', *IEEE J. Solid-State Circuits*, 47(6), 1394–1407.

Gola A., Pancheri L., Piemonte C. and Stoppa D. (2011), 'A SPAD-based hybrid system for time-gated fluorescence measurements', *Proc. SPIE*, 8033, 803315-803315-4.

Grauw C.J. and Gerritsen H.C. (2001), 'Multiple time-gate module for fluorescence lifetime imaging', *Appl. Spectrosc.*, 55, 670–678.

Guo J. and Sonkusale S. (2011a), 'A single chip fluorometer for fluorescence lifetime spectroscopy in 65 nm CMOS', *IEEE Sensors*, 766–769.

Guo J. and Sonkusale S. (2011b), 'A CMOS imager with digital phase readout for fluorescence lifetime imaging', *Proc. Eur. Solid-State Circs. Conf.*, 115–118.

Hanson K.M., Behne M.J., Barry N.P., Mauro T.M. and Gratton E. (2002) 'Two-photon fluorescence imaging of the skin stratum corneum pH gradient', *Biophys. J.*, 83, 1682–1690.

Huang T.-C.D., Sorgenfrei S., Gong P., Levicky R. and Shepard K. L. (2009), 'A 0.18 μm CMOS array sensor for integrated time-resolved fluorescence detection', *IEEE J. Solid-State Circ.*, 44(5), 1644–1654.

Kawahito S., Halin I.A., Ushinaga T., Sawada, T., Homma, M. and Maeda Y. (2007), 'A CMOS time-of-flight range image sensor with gates-on-field-oxide structure', *IEEE Sensors J.*, 7(12), 1578–1586.

Khan R., Chodavarapu V.P., Tripathi V.S. and Bright, F.V. (2009), 'Sensitivity-enhanced CMOS phase luminometry system using Xerogel-based sensors', *Biomedical Circuits and Systems, IEEE Transactions on*, 3(5), 304–311.

Kleinfelder S., Chen Y., Kwaitkowski K. and Shah A. (2004) 'High-speed CMOS image sensor circuits with *in-situ* frame storage', *IEEE Trans. Nucl. Sci.*, 51, 1648–1656.

Kleinfelder S., Kwiatkowski K., and Shah A. (2005), 'A solid state streak camera', *Proc. SPIE*, 5580, 235–243.

Krishnan R., Saitoh H., Terada H., Centonze V. and Herman B. (2003), 'Development of a multiphoton fluorescence lifetime imaging microscopy (FLIM) system using a streak camera', *Review of Scientific Instruments*, 74(5), 2714–2721.

Lakowicz J. (2006), *Principles of Fluorescence Spectroscopy*, New York, Springer.

Lange R. and Seitz P. (2001), 'Solid-state time-of-flight range camera', *IEEE J. Quantum Electron.*, 37(3), 390–397.

Lenn T. and Leake M. C. (2012), 'Experimental approaches for addressing fundamental biological questions in living, functioning cells with single molecule precision', *Open Biology*, 2, 120090.

Li D.-U., Bonnist E., Renshaw D. and Henderson R. (2008), 'On-chip, time-correlated, fluorescence lifetime extraction algorithms and error analysis', *J. Opt. Soc. Am. A.*, 25(5), 1190–1198.

Li D.-U., Walker R., Richardson J., Rae B., Buts A., Renshaw D. and Henderson R. (2009), 'Hardware implementation and calibration of background noise for an integration-based fluorescence lifetime sensing algorithm', *J. Opt. Soc. Am.*, 26(4), 804–814.

Li D.-U., Rae B., Andrews R., Arlt J. and Henderson R. (2010), 'Hardware implementation algorithm and error analysis of high-speed fluorescence lifetime sensing systems using center-of-mass method', *J. Biomed. Opt.*, 15(1), 017006–017006-10.

Li D.-U., Arlt. J., Tyndall D., Walker R., Stoppa D., Charbon E. and Henderson R. (2011), 'Video-rate fluorescence lifetime imaging camera with CMOS single-photon avalanche diode arrays and high-speed imaging algorithm', *J. Biomed. Opt.* 16(9), 096012–096012-12.

Li D.-U., Ameer-Beg S., Arlt J., Tyndall D., Walker R., Matthews D., Visitkul V., Richardson J. and Henderson R.K. (2012), 'Time-domain fluorescence lifetime imaging techniques suitable for solid-state imaging sensor arrays', *Sensors*, 12(5), 5650–5669.

Lippitsch M. E. and Draxler S. (1993), 'Luminescence decay time-based optical sensors: principles and problems', *Sensor and Actuators B*, 11, 97–101.

Maruyama Y. and Charbon E. (2011), 'An all-digital, time-gated 128 × 128 SPAD array for on-chip, filter-less fluorescence detection', *Proc. 16th Int. Solid-State Sensors, Actuators and Microsystems Conference*, 1180–1183.

Matthews D., Summers H., Njoh K., Errington R., Smith P., Ameer-Beg S. and Vojnovic B. (2005), 'Time-resolved fluorescence measurements using stroboscopic excitation', *Proc. of SPIE-OSA Biomedical Optics*, 5864, MD3.

Michalet X., Kapanidis A., Laurence T., Pinaud F., Doose S., Pflughoefft M. and Weiss S. (2003) 'The power and prospects of fluorescence microscopies and spectroscopies', *Annual Review of Biophysics and Biomolecular Structure*, 32, 161–82.

Mitchell A.C., Wall J.E., Murray J.G. and Morgan C.G. (2002), 'Direct modulation of the effective sensitivity of a CCD detector: a new approach to time-resolved fluorescence imaging', *J. Microsc.*, 206, 225–232.

Mosconi D., Stoppa D., Pancheri L., Gonzo L. and Simoni A. (2006), 'CMOS single-photon avalanche diode array for time-resolved fluorescence detection', in *Proc. 32nd European Solid-State Circuits Conference*, 564–567.

Niclass C., Rochas A., Besse P.A. and Charbon E. (2005), 'Design and characterization of a CMOS 3-D image sensor based on single photon avalanche diodes', *IEEE J. Solid-State Circuits*, 40, 1847–1854.

Niclass C., Favi C., Kluter T., Gersbach M. and Charbon E. (2008), 'A 128 × 128 single-photon image sensor with column-level 10-bit time-to-digital converter array', *IEEE J. Solid-State Circuits*, 43, 2977–2989.

Oggier T., Lehmann M., Kaufmann R., Schweizer M., Richter M., Metzler P., Lang G., Lustenberger F. and Blanc N. (2004), 'An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger),' *Proc. SPIE*, 5249, 534–545.

Pancheri L. and Stoppa D. (2008), 'A SPAD-based pixel linear array for high-speed time-gated fluorescence lifetime imaging', in *Proc. IEEE European Solid-State Circuits Conference*, 428–431.

Pancheri L., Massari N., Borghetti F. and Stoppa D. (2011), 'A 32 × 32 SPAD pixel array with nanosecond gating and analog readout', in *Proc. of 2011 International Image Sensor Workshop*, Hokkaido, Japan.

Patounakis G., Shepard K. and Levicky R. (2006), 'Active CMOS array sensor for time-resolved fluorescence detection', *IEEE J. Solid State Circuits*, 41, 2521–2530.

Pepperkok R., Squire A., Geley S. and Bastiaens P. I. H., (1999), 'Simultaneous detection of multiple green fluorescent proteins in live cells by fluorescence lifetime imaging microscopy', *Current Biology*, 9, 269–274.

Prasad P. N. (2003), *Introduction to Biophotonics*, New York, Wiley-Interscience.

Rae B.R., Yang J.B., Mckendry J., Gong Z., Renshaw D., Girkin J.M., Gu E., Dawson M.D. and Henderson, R.K. (2010), 'A vertically integrated CMOS microsystem for time-resolved fluorescence analysis', *IEEE Trans. Biomed. Circuits and Systems*, 4, 437–444.

Razavi B. (1997), 'Design considerations for direct-conversion receivers', *IEEE Trans. Circuits Syst. II*, 44, 428–435.

Richardson J.A., Grant L.A. and Henderson, R.K. (2009a), 'Low dark count single-photon avalanche diode structure compatible with standard nanometer scale CMOS technology', *IEEE Photon. Tech. Lett.*, 21(14), 1020–1022.

Richardson J., Walker R., Grant L., Stoppa D., Borghetti F., Charbon E., Gersbach M. and Henderson R.K. (2009b), 'A 32 × 32 50 ps resolution 10 bit time to digital converter array in 130 nm CMOS for time correlated imaging', in *Proc. IEEE Custom Integrated Circuits Conf.*, 77–80.

Rochas A., Gani M., Furrer B., Besse P.-A. and Popovic R. (2003), 'Single photon detector fabricated in a complementary metal-oxide-semiconductor high-voltage technology', *Rev. Sci. Instrum.*, 74, 3263–3270.

Sanders R., Draaijer A., Gerritsen H.C., Houpt P.M. and Levine Y.K (1995), 'Quantitative pH imaging in cells using confocal fluorescence lifetime imaging microscopy', *Anal. Biochem.* 227, 302–308.

Schwartz D.E., Gong P. and Shepard K.L. (2008a) 'Time-resolved Forster-resonance-energy-transfer DNA assay on an active CMOS microarray', *Biosensors and Bioelectronics*, 24, 383–390.

Schwartz D.E., Charbon E., Shephard K.L. (2008b), 'A single-photon avalanche diode array for fluorescence lifetime imaging microscopy', *IEEE J. Solid State Circuits*, 43, 2546–2557.

Sharman K.K. and Periasamy, A. (1999), 'Error analysis of the rapid lifetime determination method for double-exponential decays and new windowing schemes', *Anal. Chem.*, 71(5), 947–952.

Sillen A. and Engelborghs, Y. (1998), 'The correct use of average fluorescence parameters', *Photochem and Photobiol.*, 67(5), 475–486.

Stoppa D., Borghetti, F., Richardson J., Walker R., Grant L., Henderson R.K., Gersbach M. and Charbon E. (2009), 'A 32 × 32-pixel array with in-pixel photon counting and arrival time measurement in the analog domain', in *Proc. Eur. Solid-State Circuits Conf.*, 204–207.

Theuwissen A.J.P. (2008) 'CMOS image sensors; state-of-the-art', *J. Solid-State Electronics*, 52, 1401–1406.

Theuwissen A.J.P. and Seitz P. (eds) (2011), *Single-Photon Imaging*, New York, Springer.

Tubert C., Simony L., Roy F., Tournier A., Pinzelli L. and Magnan P. (2009), 'High Speed dual port pinned-photodiode for time-of-flight imaging', in *Proc. of 2009 International Image Sensor Workshop*, Bergen, Norway.

Turconi S., Bingham R.P., Haupts U. and Pope A. J. (2001), 'Developments in fluorescence lifetime based analysis for ultra-HTS', *Drug Discovery Today*, 6, 27–39.

Tyndall D., Rae B., Li D.-U., Richardson J., Arlt J. and Henderson R. (2012), 'A high-throughput time-resolved mini-silicon photomultiplier with embedded fluorescence lifetime estimation in 0.13μm CMOS', *IEEE Trans. Biomed. Circuits Syst.*, 6(6), 562–70.

Veerappan C., Richardson J., Walker R., Li D.-U., Fishburn M. W., Maruyama Y., Stoppa D., Borghetti F., Gersbach M., Henderson R.K. and E. Charbon (2011), 'A 160 × 128 single-photon image sensor with on-pixel 55 ps 10 b time-to-digital converter', *IEEE ISSCC Dig. Tech. Papers*, 312–314.

Verveer P.J., Squire A. and Bastiaens P.I.H. (2000), 'Global analysis of fluorescence lifetime imaging microscopy data', *Biophysical J.*, 78, 2127–2137.

Wang Y., Rae B.R., Henderson R.K., Gong Z., McKendry J., Gu E., Dawson M.D., Turnbull G.A. and Samuel I.D.W. (2011), 'Ultra-portable explosive sensor based on a CMOS fluorescence lifetime analysis microsystem', *AIP Advances*, 1, 032115.

Won Y.J., Moon S., Yang W., Kim D., Han W.-T. and Kim D. Y. (2011), 'High-speed confocal fluorescence lifetime imaging microscopy (FLIM) with the analog mean delay (AMD) method', *Opt. Express*, 19(4), 3396–3405.

Wu P. and Brand L. (1994), 'Resonance energy transfer: methods and applications', *Anal. Biochem.*, 218(1), 1–13.

Yoon H.-J., Itoh S. and Kawahito S. (2009), 'A CMOS image sensor with in-pixel two-stage charge transfer for fluorescence lifetime imaging', *IEEE Trans. Electron Devices*, 56, 214–221.

Zhao Q., Young I.T., Schelen B., Schouten R., van den Oever R., Leenen R., van Kuijk H., Peters I., Polderdijk F., Bosiers J., Jalink K., de Jong S., van Geest B. and Stoop K. (2012), 'MEM-FLIM: all-solid-state camera for fluorescence lifetime imaging', *SPIE Photonics West*, 82250N–82250N-11.

Zhong W., Urayama P. and Mycek M-A. (2003), 'Imaging fluorescence lifetime modulation of a ruthenium-based dye in living cells: the potential for oxygen sensing', *J. Physics D: Appl. Phys.*, 36, 168.

<div align="right">

12

</div>

Complementary metal-oxide-semiconductor (CMOS) X-ray sensors

A. STRUM and A. FENIGSTEIN, TowerJazz, USA

Abstract: CMOS image sensors (CIS) are taking an ever growing part in the X-ray sensors market, competing with a:Si flat panels and direct conversion sensors. CMOS image sensors are superior by means of power consumption, noise level, and in-pixel special functionality. This chapter reviews the history of CIS role in the X-ray sensing market. Performance parameters and figure of merits of X-ray sensors are presented, and the evolution of CIS for X-ray sensing is discussed. Requirements and solutions for different applications are also presented. Comparison to competing technologies is briefly discussed and the chapter concludes with future trends of these promising devices.

Key words: X-ray, complementary metal-oxide-semiconductor (CMOS) image sensors, pixel architecture, high X-ray detectivity (DQE), signal-to-noise ratio (SNR).

12.1 Introduction

Since the discovery of X-rays by Wilhelm Roentgen in November 1895, X-ray imaging has played a major role in medical as well as dental applications. The main use of X-ray was, and still is, radiography or still imaging, mainly for bone fractures, various dental applications and more recently, for mammography. For still X-ray imaging, like any other still imaging, the 'imager' used was a film or film plate; a sheet of celluloid covered by silver halide salts, sensitive to light. Unlike standard cameras that are equipped with lenses that project the image on the film, in X-ray photography there is no lens and therefore, the imaging is 'one to one' or 'contact imaging' meaning the image of an object is the actual size of the object. The ramification of this is that the film size, and later on the digital sensors sizes, are relatively large compared to those used in standard photography.

In the past, the process of taking an X-ray image was quite cumbersome and slow. Once the shot was taken, the film had to go to the lab to be developed and then it was returned to the doctor. During this time, the patient had to wait because there was a possibility that the image taken was not successful and another one would be required. We will see later on how different technologies were developed in order to address this problem

348

which placed a bottleneck on the amount of patients a radiologist or dentist could see in a day.

Similar to the film industry, a motion X-ray version was developed later, but obviously could not be film-based due to the large dimensions requirement mentioned above. The technology for motion X-ray imaging, or fluoroscopy that is still widely used, is based on an image intensifier tube (IIT) where an electron beam scans a phosphorus screen of a tube. The X-ray radiates on the external side of the tube and the scanning electron beam creates a visible image through another phosphorus layer at the output of the tube. In such a way, a high resolution and high frame rate video can be created in real time. Due to the high sensitivity of the tube and its inherent gain, much lower X-ray doses as compared to radiography can be used. This technology is widely used in cardiac (catheterization), vascular (angiography) and surgical applications. The major issues with IIT are image distortion and short lifetime of the tube.

12.2   Intra-oral and extra-oral dental X-ray imaging

A common use of X-ray imaging that surely every reader has experienced at least once, is dental imaging. Until recently, film was the common media for taking an image. This is a fairly easy procedure that the dentist can handle in the clinic, not only by taking the image, but also developing the film. However, the movement of standard photography from film to digital sensors was a major driver for X-ray technology to replace film with digital imaging. Moreover, this change created a steep reduction in demand for film and thus, raised the cost of film required for X-ray. In addition, environmental requirements made the chemical handling and disposal expensive. Other reasons included throughput at the dental clinic that could have been improved by the immediate result of digital sensor images compared to that of film which needs time for development as well as the fact that a digital image can be stored on a computer database and be easily retrieved. The first digital imaging used to replace film was CCD (charged coupled device), the same as in standard photography. Of course, silicon cannot be sensitive to X-ray radiation; the X-ray photon energy is much too high and silicon is basically transparent to X-ray light so it cannot absorb such a photon and translate it into an electrical signal. The way to solve this issue was to use a scintillation material that absorbs X-ray photons and emits visible photons. The scintillation material also has an inherent gain (1,000–10,000) which is the ratio between the number of emitted visible photons and the number of absorbed X-ray photons. The usual scintillation materials are CsI:Tl (thallium doped cesium iodine), with a small percentage of thallium that emits green photons and gadolinium oxide (GOS) which is lower quality but a much cheaper material. The

structure of the sensor is quite simple. It is built basically as a regular sensor but with much larger pixels, on top of which a scintillator is placed (usually with a fiberglass plate in between, to help channel the green photons into the pixels and prevent 'crosstalk'). This structure is packaged and sealed in a plastic (epoxy) package with a cable, usually a USB, that connects directly to the computer.

In time, digital still cameras moved from CCD technology to complementary metal-oxide-semiconductor (CMOS)- based sensors, and X-ray sensors have followed for the same reasons. The three major drivers were:

1. much lower power consumption of CMOS compared to CCD
2. higher data rate and frame rate that allows higher resolutions and
3. much higher yields and thus lower cost of CMOS versus CCD.

Today, many dental clinics have shifted from film to digital detectors, some to CCD and then to CMOS, and some directly to CMOS. There are two major sizes for a dental image sensor: size 1 and size 2 (smaller). The size 1 sensor is about $43\,mm \times 30\,mm$ and size 2 is about $32\,mm \times 24\,mm$.

In the extra oral dental market, especially in panoramic dental X-Ray, the same trend existed. Film was replaced by long CCD line sensors, usually using TDI (time delay and integration) architecture to get higher sensitivity. The main reason for the need for high sensitivity is that unlike intra-oral X-ray, where the X-ray tube is placed very close to the detector ($1–2\,cm$ proximity), in panoramic X-ray, the tube is usually placed further away and since the amount of X-ray photons per area is reversely proportional to the square of the distance, much less X-ray radiation will hit the sensor. This is the reason that CCD is only now starting to be replaced by CMOS. Until recently, CMOS sensitivity was not good enough and TDI architecture was much more natural for CCD.

A new area recently developed in the dental area is cephalography which is used mainly for dental surgery and dental CT. It provides a 3D image of the skull and is used for plastic and aesthetic dental surgery. The detectors are quite large, usually measuring 5×6 inches to allow the whole skull to be imaged with one shot.

12.3   Medical radiography, fluoroscopy and mammography

Medical radiography requires, by nature, large areas. The standard dimensions are 10×12 inches ($25 \times 30\,cm$) but can get as large as 17×17 inches ($43 \times 43\,cm$). For film, this never posed any problem. However, film although cheap, is very cumbersome and time consuming since it involves a development process using chemicals that are not environmentally friendly, and it takes a lot of both the patient's time as well as the doctor's.

Therefore, more than a decade ago several companies invested in a film type solution that is not disposable, a method called computed radiography (CR). The idea was to have a media very similar to film, covered by a photo sensitive phosphor. When an X-ray photon hits the phosphor molecule, it excites an electron to a high energy level. The electron gets stuck there, until at a later stage, the media is scanned by a laser beam (the 'printer'). When the beam hits an excited molecule, the electron falls back to its low energy level and a blue photon is emitted. A photo sensitive detector that is fully synchronized with the laser beam reads the level of blue photons. This process allows reading the image directly into the computer and multi-use of the 'film'. Still, the resolution of CR, similar to film is not great and the process is still time consuming.

The follow-on technology is based on TFT (thin film transistors) on glass that became commercially available due to the huge growth of LCD televisions. The technology uses amorphous silicon (a:Si) based photodiodes in a 1T (one transistor) pixel configuration. Then, on top of the structure of the pixel matrix, a Cs:I (cesium iodine) scintillator is deposited in order to translate the X-ray photons into visible light that is then detected by the photodiodes. There is no circuitry on the glass plate and therefore, every column pad and every row pad is wire bonded to adjacent electronics. The output is of course analog and is translated to digital by outside analog to digital converters (ADCs). This technology, although quite expensive compared to film or CR, has a lot of advantages such as high sensitivity to X-ray radiation allowing a much lower dose for shooting an image; high resolution, high dynamic range, and real digital imaging that can be even improved by follow-on image processing; and most importantly, an immediate result like regular commercial photography that saves enormous time for patients and doctors. This technology is not yet widely used mainly due to its high cost, but the adoption rate is high.

Fluoroscopy, or X-ray video, is being used widely in the medical field for chest X-ray, cardiac and catheterization, angiography and more. The most commonly used technology is IIT; its mode of operation was explained earlier. The advantages of this technology are mainly its high sensitivity and relatively low cost. The disadvantages are the high cost of ownership that includes the need for a high voltage stable power supply and the replacement of the tube every five to eight years as well as image distortion that is very difficult to correct. The developments in the flat panel display (FPD) area allowed this technology to also get into the fluoroscopy field. It took many years for the FPD technology to be able to perform well at a high frame rate, i.e. 60 frames per second which is required for cardiac applications. The major problems were related to the fact that the large dimensions of the panel introduced a very high capacitive load of the metal lines of the pixel columns combined with high resistance, creating practically delay

lines. But, with time, the problems were solved and now FPDs at sizes of 17×17 inches are widely used in operation rooms, gradually replacing the IIT systems.

Unlike standard radiography, mammography requires much higher resolution. The standard resolution of standard radiography FPDs varies between 130 and 180 micron pixels. For mammography, pixel sizes of 50 to 75 microns are required. FPD technology cannot support such small pixels and also, the dispersion of the cesium iodine scintillator deteriorates the modulation transfer function (MTF) of such fine resolution. Therefore, direct detection detectors using amorphous selenium were developed. These detectors are based on just a switching matrix of TFT that switches capacitors made between a metal plate on each individual pixel and the amorphous selenium. The X-ray photons hit the amorphous selenium material and create electrons that are being collected by the metal plates of the pixels. This technology is widely used now in all new mammography systems and is gradually replacing film.

12.4   CMOS image sensor (CIS)-based flat panel display (FPD) technology

The ever progressing CMOS technology that enables the highest performing image sensors in visible light is a natural candidate to dominate the X-ray indirect sensor field as well. Comparing the a:Si FPD to CIS technology, one can expect to have some significant advantages of CIS over the competing a:Si FPDs [1].

The maturity of CMOS technology offers:

- higher speed
- lower noise
- lower power consumption (and thus easier temperature control)
- better fill factor, especially for smaller pixels and
- no limits on pixel size.

On the other hand, CIS technology is limited to wafer sizes (8 or 12 inches in diameter) while a:Si FPDs enable sizes up to 43×43 cm panels quite easily. In addition, the cost per square centimeter is higher for CMOS technology. As described below, the area limitation can be solved by tiling of several CIS sensors; however, this will increase the cost. Nevertheless, as will be discussed later, for several applications the superior performance justifies this higher cost.

In general, CMOS flat panels are visible light sensors and as such, follow the visible light CIS pixel architectures. However, the characteristics of X-ray sensors, such as larger pixel dimensions (20 μm to 150 μm compared to less than 10 μm for most visible light), the demand for larger full well

capacity (100,000 to several millions of electrons compared to a few thousand electrons for visible light sensors), along with the requirement for immunity of the circuits against X-ray damage and the large die yields requirements, affect the considerations of pixel as well as overall architecture selection.

12.5   Pixel design considerations for CMOS-based FPDs

12.5.1 1T pixel architecture on a:Si TFT FPD

The simplest architecture of a pixel requires a photo diode and one transistor (see Fig. 12.1). Due to the complexity of making several high quality transistors within the pixel in TFT technology, this is the way most, if not all, a:Si FPDs are implemented. The operation of such a pixel is as follows. Pixels are reset line by line by applying high potential on the row-select lines and high potential for the V_o lines. After integrating photo-electrons on the diodes, the sensor is read line by line as well. Reading a pixel is carried out by setting the transistor to 'on' so it is connected to a charge amplifier that allows transfer of the negative charge to a capacitor in the readout circuitry where it is sampled. Double sampling is typically used; namely another reset operation followed by a sampling operation. Subtraction operation between the two signals gives a signal which does have less noise, but noise cannot be completely removed, especially the kTC noise which needs 'true correlated double sampling' in order for it to be removed. Since a high performing charge amplifier cannot be realized by TFT technology, all the circuits are implemented as stand-alone Si chips, usually one per column. This makes the flat panel boards crowded, bulky, heavy, and with rather high power consumption. All this can be applied of course to CMOS X-ray sensors, but will not benefit much from the advantages of the superior CMOS technology which allows embedding electronic circuits on the sensor itself.

12.5.2 Three transistors (3T) pixel architecture

3T pixels are simply what are known as 'active pixels'. The scheme of such pixels is shown in Fig. 12.2. The pixels are reset line by line (though in principle can be reset simultaneously for the whole array) by putting the reset transistor to the 'on' state. The diode integrates photo-electrons and thus the potential of the diode drops according to the diode's parasitic capacitance and to the amount of electrons being collected. This potential is being read through the unity gain amplifier realized by a transistor connected as 'source follower' (SF) through the row-select transistor switch
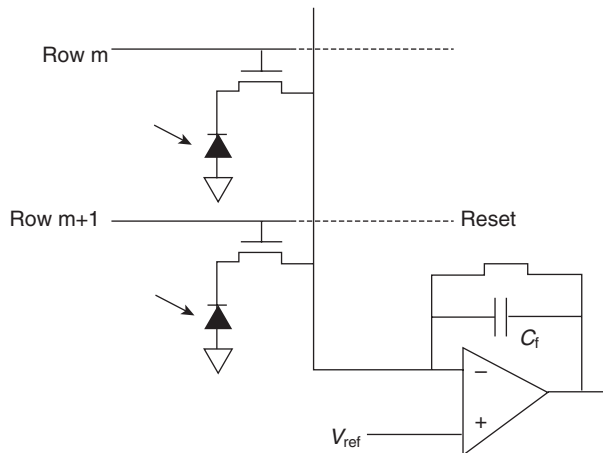
12.1 1T pixel architecture. Each pixel contains a photodiode and one transistor. The amplifier is located at the bottom of the column.
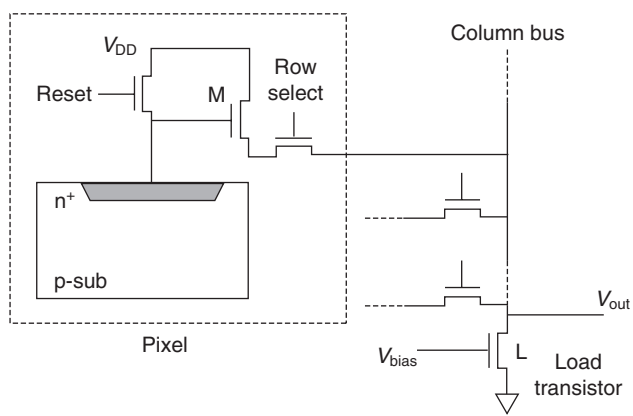
12.2 3T pixel architecture. Transistor 'M' is connected as a 'source follower' (SF) and serves as a unity gain amplifier.

to the column line. A 'double sampling' procedure is carried out by resetting the diode again and running another read operation. Subtracting the two samples gives a signal which is better cleared of noises (such as power supply noise), but again kTC noise cannot be eliminated due to the non correlating samplings (subtraction of the reset signal of the next cycle from the signal instead of subtraction of the corresponding reset signal). For visible image sensors like cell phone cameras and digital still cameras, the 3T pixel scheme was replaced by a four transistor (4T) scheme reviewed in the next paragraph.

As will be shown, the 4T pixel can eliminate the kTC noise as well. However, for indirect sensors, the full well capacity required is fairly high, capacitance of the sensing node is rather high and the resulting kTC noise is therefore rather low. For instance, an intraoral pixel which usually ranges between 15 µm to 20 µm typically requires full well of a few hundreds of thousands of electrons. This in turn needs a diode's capacitance of roughly 100 fF. The kTC noise calculated by $V_{n,kTC} = \sqrt{\dfrac{kT}{C}}$ is about 70 µV RMS, which is usually below the noise level of a state-of-the-art signal analog chain. For this reason, 3T pixels are very popular for CMOS indirect X-ray sensors. This is true for the rather small intraoral dental pixels, as well as for larger X-ray sensors. For the latter usage, the 4T scheme is even more problematic due to the long transfer time needed to fully evacuate the electrons from the photodiode during the read cycle of a 4T operation. The 3T scheme has low pixel complexity and a simpler implant scheme, thus having a good fill factor, fast read, lower cost process (lower photo mask layer count) and better yield.

12.5.3 Partially pinned and fully pinned photo diode: 4T pixel considerations

A 4T pixel, based on the concept of a fully pinned photodiode is currently the dominating technology in the visible camera market. This includes the huge market of cell phone cameras, digital still cameras and high end cameras for photography and other applications. The 4T pixel is depicted in Fig. 12.3.

In the heart of such a pixel, there is a specially optimized photodiode which is fully surrounded by grounded p-type implanted junctions. While being reset, the diode is fully evacuated from all of the electrons. After integrating photo-electrons, the collected electrons are fully transferred to a sense node known as 'floating diffusion' (FD) and then sampled. In this case, a true correlated double sampling can be obtained. The FD is reset
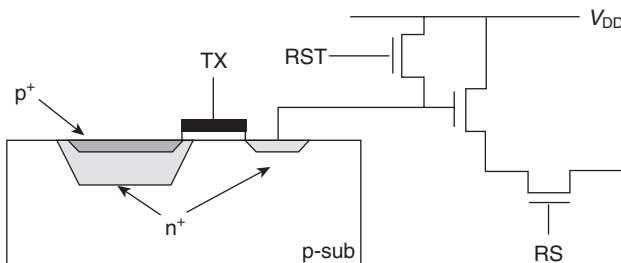
12.3 The 4T pixel architecture.

and sampled, and later on, a full charge transfer is carried out from the diode to the FD and the FD is sampled again. Subtraction of the two results in a signal which is clear of kTC noise. Another advantage of the pinned diode is the ability to achieve very low dark currents. Being 'encapsulated,' the photodiode is protected from carrier generation in surfaces and interfaces. A well optimized 4T pixel can achieve a noise floor of a few electrons.

As explained before, for indirect X-ray sensors, each X-ray photon is converted to hundreds of visible photons. Reducing the noise to a level of a few electrons is not required. However, the concept of a pinned diode can be used for other reasons. Although there is no need for a fully pinned photodiode, a 'partially pinned photo diode' (PPPD) can have the required impact without the penalty of the longer transfer time. The dark current reduction achieved by the abovementioned 'encapsulation' of the n-region of the photodiode by a grounded p-region is crucial for high performing image sensors, especially for long exposures. Another important feature of the pinned photodiode is the reduced photodiode capacitance. The photodiode capacitance while reversed biased (which is always the case for the photodiode) is due to junction capacitance, namely the charge modulation due to diode depletion region modulation. Any part of the photodiode which is fully depleted effectively has no capacitance because any additional voltage is not going to change the depletion width. This breaks the linkage between the diode area and its capacitance. Very large area photodiodes can have very low capacitance. The target capacitance can be achieved either by modifying the ratio between the fully depleted part of the photodiode and the non-pinned area (controlled by the implant doses and coverage of the pinning p+ area definition). Another way is to reduce the diode capacitance to minimum and add a well defined capacitor, which can be either a metal-oxide-semiconductor (MOS) or MiM capacitor to accurately define the capacitance. In this case, one can also enjoy better linearity than that of diode parasitic capacitance. These features are achievable for the simple 3T scheme. A remarkable kTC reduction (though not full cancellation) is achievable by using a 'charge amplifier pixel,' a slightly more sophisticated pixel scheme while using PPPD, as will be described below [2].

12.5.4 Binning

In any image sensor, there is a tradeoff between the signal-to-noise-ratio (SNR) and the resolution as will be elaborated in the next section. Smaller pixels enable higher resolution but collect less photons and thus have less signal. This trade-off is most pronounced for fluoro-radio panels which are used for both large dose, high quality still pictures ('radio') and low dose

video capture ('fluoro'). In the latter case, one would give up resolution in order to collect more photons to cope with the 'photon starved' nature of the fluoro mode. The way to achieve both is by the concept of binning. In binning mode, several pixels are read together as if they were one pixel. This can be carried out 'after the fact' by summing up the adjacent pixels' signals 'off chip'. Even better performance can be achieved if the binning is performed inherently in the pixel array. This may have a twofold advantage:

1. by employing 'analog' binning in the array, it improves the SNR by eliminating the noise coming from the signal sampling operation
2. one simultaneous sampling of several photodiodes together saves the need for multiple sampling followed by digital summing up 'off chip'.

Another advantage is the ability to have better frame rate since one has less pixel data to pump out of the sensor. Pixel level binning can be carried out in voltage mode, where the voltage output of the binned pixels is the average voltage output of the binned pixels. This can be realized, for instance by shorting the photodiodes of 3T pixels as described in the following scheme. For other pixel architectures the binning can be carried out in charge mode, where the charge of several diodes is summed up on a sensing node. In this case one does not get an average value output but rather the sum of the diodes' outputs. This can improve the SNR even further. An example of a scheme that can realize analog charge mode binning will be described in the next section.

12.5.5 Increasing functionality of CIS pixels

Due to the ever increasing density of circuits in the CMOS process, a lot of functionality can be added to the CIS pixels. Since X-ray pixels are relatively large, one can have several transistors in the pixel without losing too much of the fill factor. Too many transistors though will degrade not only the fill factor but probably the yield as well, so this should be used carefully. In the following, a few examples of improved functionality will be introduced. The first example is the so called 'charge amplifier' pixel. This pixel uses a charge amplifier implemented by one transistor connected as the common source. The electrical scheme of such a pixel is presented in Fig. 12.4. This pixel transfers the charge from the photo diode (though not the actual electrons as in the 4T pixel scheme) to the feedback capacitor. This pixel has several advantages. As can be clearly seen, this pixel enables analog charge mode binning. Enhanced linearity is achieved by transferring the charge from the non-linear diode to the linear feedback capacitor. Last but not least, using PPPD reduces significantly the kTC noise. While in reset state, the photodiode is almost fully pinned, thus it has very low capacitance and small
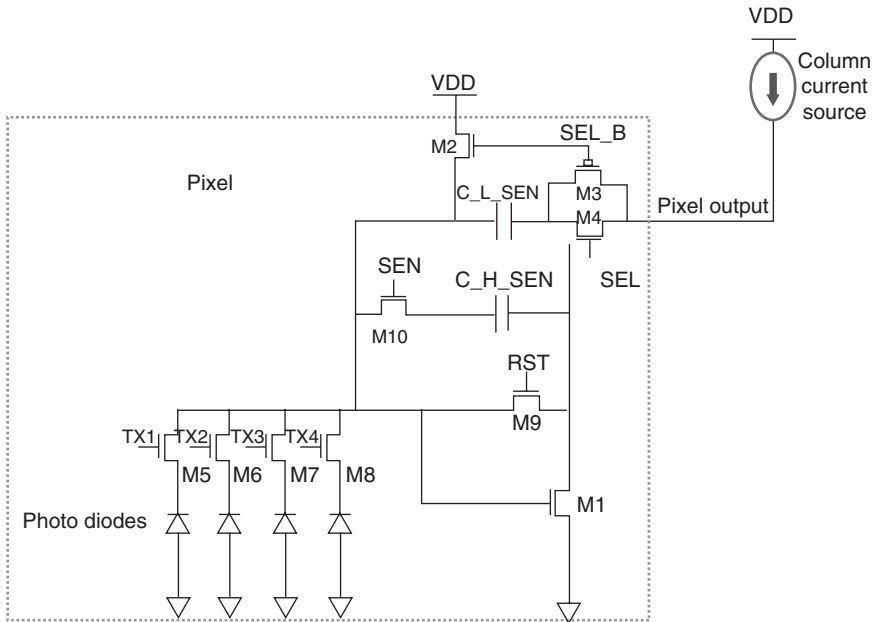
12.4 Charge amplifier pixel. The pixel is four shared by TX1 to TX4, which also enables binning by operating all the four together. The dual gain is enabled by the switch SEN which connects/disconnects the additional capacitor C_H_SEN.

charge noise. This does not come on the account of the full well capacity since during operation the voltage on the diode drops and more diode area becomes non-pinned [2]. Another option is the dual gain concept. Adding a capacitor with a switch can turn a pixel into a dual gain one. The conversion gain (CG), namely the output voltage per input charge (in µV/e units) is inversely proportional to the capacitance of the diode (3T pixel) or feedback capacitor (charge amplifier pixel). Below are the schemes of dual gain in a regular 3T pixel scheme (presented by Dalsa) [3] shown in Fig. 12.5 and for a charge amplifier pixel (in a C&T product) shown in Fig. 12.4 [4].

Dual gain mode enables switching between high gain that is fitted to low dose conditions, amplifying the signal above the noise level of the analog chain, and low CG that is used for large dose to avoid running out of the voltage swing of the sensor output (larger full well capacity). The charge amplifier dual gain pixel enables an outstanding high dynamic range (HDR) operation. After acquiring the image and storing the collected electrons in the diode one can read the same charge twice, starting with high CG and then switching the capacitor to the low CG and read again. This operation is described in [4] and the overall sensor performance is listed below.
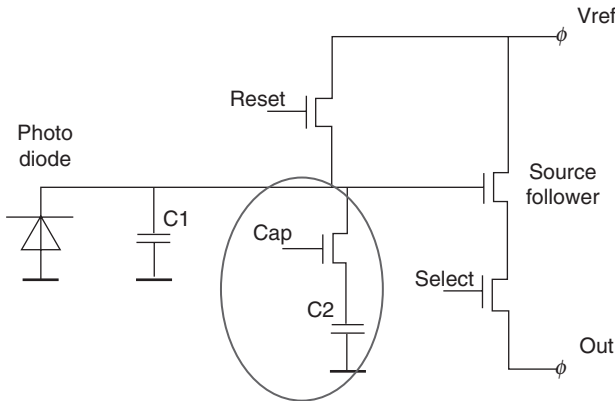
12.5 Dalsa's dual gain 3T pixel. Switch cap is used to add/remove the extra capacitance C2 to the 'floating diffusion' capacitance C1.

Another example of an even more advanced pixel functionality is the 'photon counting' pixel, proposed by Caeleste [5]. The number of simultaneously impinging visible photons gives information about the energy of the X-ray photon which enables what is called sometimes 'color' X-ray sensing. Those three are only examples of the wide range of opportunities for advanced X-ray sensing concepts opened by the in-pixel circuitry enabled by CIS technology.

12.5.6 On-chip ADC

The usage of crystalline Si for the sensor suggests the opportunity to have the ADC on-chip. This has a few advantages over the competing flat panel technology. The on-chip ADC is faster, has lower power consumption and is overall cheaper than off-chip (usually multiple) ADC(s). This helps of course to somewhat recover from the inherent higher cost of the crystalline Si sensors. On the other hand, having ADC on board complicates the design and thus reduces the area that is free for pixels, and is somewhat expected to lower the yield as will be discussed below.

12.5.7 Yield and cost considerations

X-ray sensors dimensions are much larger than common VLSI chips. The number of dies on wafers varies between a few tens (intra oral dental) to one die per wafer (medical). This places a huge challenge on yield expectations. In fact, using any standard wafer yield model will result in a yield close to zero. This has a large impact on the cost which is already rather high due to the Si area consumption of the sensors. Surprisingly enough, X-ray sensors, and even one die per wafer sensors can exceed 60% yield and be cost

effective, allowing them to compete quite well with TFT panels. One of the basic CIS features which enables this high yield is the fact that for CIS, losing a pixel, or even a lot of pixels, is usually permitted (with some limitations of course on clusters and total number of dead pixels). So, making sure that a possible defect will kill only one pixel and not the whole row/column (or sensor) is crucial. Since pixels are rather large, there is no need to stress the design to the minimum design rules of the process, especially for the back-end metal routing. Thus, extensive use of what is known as 'design for manufacturing' (DFM) can dramatically impact the yield. Another approach is to carry out smart 'thin' design, minimizing the non-pixel area (digital and analog) and the number of devices used. Non-pixel areas are usually denser and more prone to defects. Furthermore, the effect of a defect is usually catastrophic. Minimizing the number of devices used can significantly reduce the number of steps (especially the photolithography steps) in the process which reduces the cost, both, directly by means of process cost and indirectly through the resulting yield improvement (having less steps and thus less chance for defects).

Using a CMOS process for an X-ray sensor clearly limits the sensor size to wafer dimensions. This will be 200 mm or at most 300 mm in diameter (sensor diagonal). Certain medical applications, (e.g. chest, mammography and others) require larger sensors. The way to overcome this limitation is by tiling several 'one die' sensors to one big sensor, as shown in Fig. 12.6.
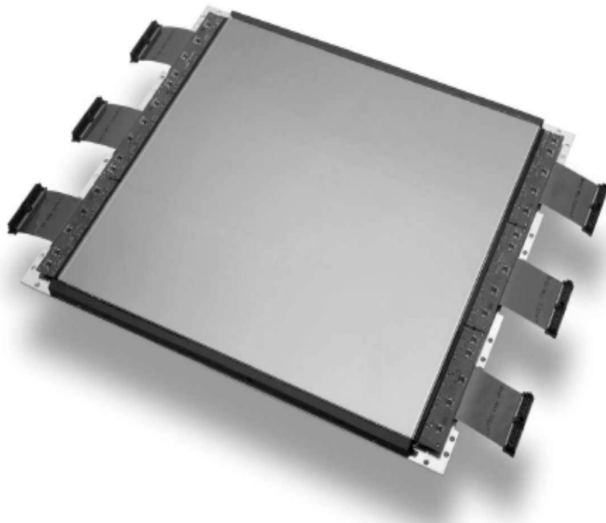
12.6 Large sensor composed of six 'butted' sensors. Ensuring frame continuity and loss of not more than one row/column for each butting line; pixels should cover the whole die to its edges excluding one edge (three side buttable).

Buffers   Logic   Shift register
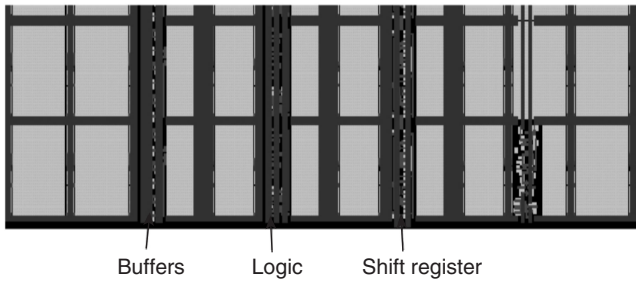
12.7 Array arrangement for three side buttable sensor with drivers and other circuitry located between the columns.

The accurate butting of several dies is a challenge by itself which will not be discussed here. There are some considerations that should be handled in the single sensor level. First, in order to enable connecting several sensors, the single sensor should be buttable, namely, having pixels all the way to the edge of the sensor. For $2 \times n$ tiling, the sensor should be three side buttable, which means that all the circuitry should be on one side of the sensor and all the other sides should have pixels all the way to the edge. Such a three side buttable sensor poses several challenges to the sensor design and manufacturing. Cutting the Si close to the pixel can have a bad effect on performance due to stresses and damage associated with the sowing, because pixels are very sensitive to even small damages or stresses. Another issue is the seam between adjacent tiles, which has a large impact on the resulting picture. One way to cope with this is to give up this row/column and correct with software algorithms. A more sophisticated, yet common approach is to have smaller pixels on the edges in such a way that the overall length of the two edge pixels and the seam between them will add up to one row/column width. Finally, having all the digital part in one side puts a challenge on the routing of the orthogonal direction which is usually the row drivers direction. This becomes even more complicated due to the fact that such large sensors are made using photo lithography with 'stitching' which requires building the array from repeating parts. The way to deal with this is either by pushing circuitry inside the array [3] as shown in Fig. 12.7, or by a smart stitching scheme [6].

12.6   Key parameters for X-ray sensors

X-ray sensors parameters or figures of merit are similar of course to any image sensors though there is a specific figure of merit, namely detectable quantum efficiency (DQE), which is used in this field more than any other imaging field. In the following section these figures of merit will be discussed along with their typical numbers and the trade-offs between them.

12.6.1 SNR

SNR is the one parameter which quantizes our ability to resolve the required image from the noise it is embedded in. This of course depends on the object, the illumination (or dose) and the sensor quality, and thus it cannot serve as a figure of merit of a sensor. SNR of ~3 is considered 'visible,' namely an image can be seen on the background of the noise if the image signal level is at least about 3x higher than the noise level. The signal itself has Poissonic noise distribution known as 'shot noise'. The RMS noise of the photon number equals the square root of the average dose hitting the pixel. It is most important to note that in the case of an indirect sensor where one has large gain from X-ray to visible photons (at least several hundred) coming from the scintillator, the SNR will be determined by the X-ray photon shot noise. The scintillator gain by itself is not a deterministic process and thus, adds even more noise which is independent of the CIS being used. So the largest possible SNR of a perfect indirect sensor is the square root of the number of the X-ray photons. For low dose, the shot noise will become smaller than the noisefloor of the pixel and the analog chain. Assuming that noise sources are independent, the total SNR for aspecific illumination level (or dose) will be calculated by $SNR = \dfrac{V_{sig}}{\sqrt{\sum\limits_i^N V_n^2}}$ where V_1, V_2 ... V_N are noise sources which can be signal dependent like the shot noise or signal independent like the analog chain noise floor.

12.6.2 Resolution, MTF and DQE

Resolution usually refers to the number of distinct pixels in the sensor. The ability to distinguish between two small adjacent features in the scene naturally depends upon pixel dimensions but also on the amount of crosstalk between pixels which degrades the effective resolution. The measure for total resolution is the MTF which is defined as the response of the sensor to different spatial frequencies. The MTF is a function, not a single number that represents per each spatial frequency, measured in line-pairs-per-millimeter (lp/mm), the ratio of the signal with respect to the signal at zero spatial frequency. The MTF can be measured directly by radiating an X-ray signal through a 'sine target' or 'bar target' as shown in Fig. 12.8.

It can be shown that for the simplest assumption on pixel response, namely uniform response at any point in the pixel and no response to light outside of the pixel, the MTF is: $MTF = \sin\left(\dfrac{\pi f}{2 f_n}\right) \bigg/ \dfrac{\pi f}{2 f_n}$
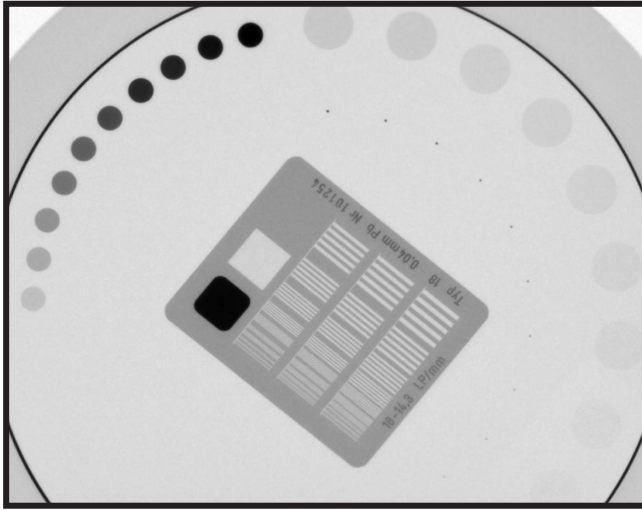
12.8 X-ray phantom for MTF evaluation.

where f, is the signal spatial frequency and f_n is the so called Nyquist frequency; $f_n \equiv 1/2p$ where p is the pixel pitch. Sometimes MTF is reported by the signal drop at the Nyquist frequency which is roughly 50% for ideal geometric MTF. The MTF drops further when there is crosstalk between pixels. This crosstalk can be due to several mechanisms: optical crosstalk (light getting to the 'wrong' pixel), diffusion crosstalk (photoelectrons diffused to an adjacent pixel) and electrical crosstalk (interference between output signals). Due to the rather large X-ray pixel geometry, diffusion crosstalk usually can be minimized and for a well designed X-ray sensor, electrical crosstalk is avoided as well. Thus, for an indirect X-ray sensor, the major crosstalk contribution is optical crosstalk usually due to the scintillator. An example of MTF taken for Trixell's 154 μm pixel FPD is shown in the left curve in Fig. 12.9.

In this case, Nyquist frequency is about 3 lp/mm. The theoretical geometric MTF is shown for comparison. The MTF gives information only on resolution but not on SNR. It is quite obvious that having low SNR affects the ability to resolve high spatial frequencies, but it is not reflected in the MTF curve. A rather new concept, called DQE, proposed by Albert Rose [7] , takes both SNR and resolution into the same framework. The DQE is defined by:

$$DQE \equiv \frac{SNR_{out}^2}{SNR_{in}^2}$$

for a perfect sensor $DQE = 1$. The DQE is presented as a function of the spatial frequency. The low pass filter nature of the MTF reduces the signal for high spatial frequencies and the DQE is degraded. The DQE for the
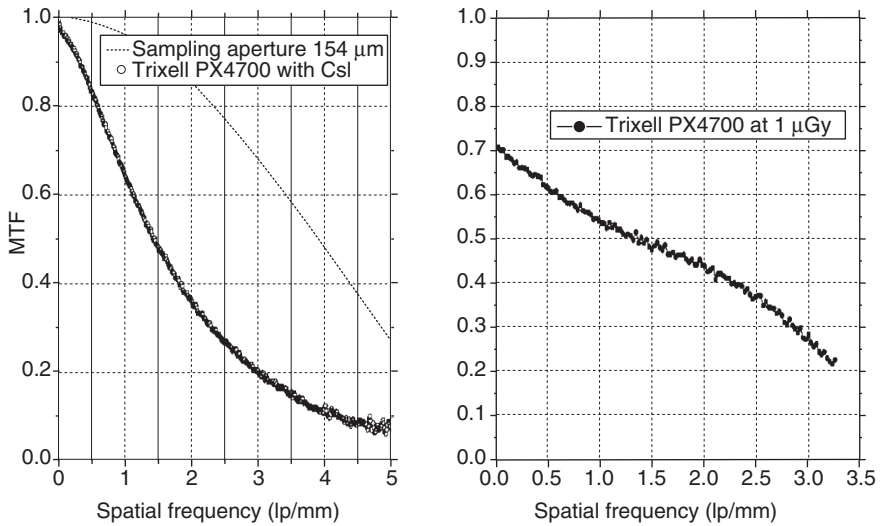
12.9 An example of typical MTF and DQE curves for a medical X-ray sensor (Trixell's PX4700 flat panel sensor).

same Trixell's sensor is shown above. As stated before, the SNR for low doses is dominated by the pixel and analog chain noisefloor. In this case, the DQE will be dose dependent, because the SNR_{out} is not linearly dependent on SNR_{in}. The convention is to show DQE for a high enough dose to be significantly higher than the noisefloor such that the DQE becomes independent of the dose. This curve is indeed more informative but still cannot give the full picture of the sensor quality. The other two sensor parameters that shed more light on the low dose performance and its trade-offs are the noise equivalent dose (NED) and the dynamic range (DR) of the sensor.

12.6.3 NED and DR

NED provides information about the lowest dose which is still resolvable. NED is defined by the X-ray dose which will result in an SNR value of 1. This definition takes into account the major performance features dictating the low signal sensitivity of the sensor, namely the overall efficiency of the sensor to convert X-ray photons to an electrical signal (the 'signal'), and the suppression of noises disturbing the sensing (the 'noise'). It is clear, for instance, that scintillator performance has a direct impact on the NED, as well as the pixel dimensions and of course the pixel and circuitry quality. The parameter that completes the picture is the DR. The DR reflects the ability of a sensor to have high illuminated and low illuminated details within the same frame. For a very sensitive sensor having small DR, the

illuminated parts of the frame will be saturated and will show no details. The DR is formally defined as the largest possible signal divided by the minimum resolvable signal – in dB units. This can be one of the following equivalent expressions:

$$20 \cdot \log_{10}\left(\frac{\text{Maxdose}}{\text{NED}}\right) = 20 \cdot \log_{10}\left(\frac{\text{VoltageSwing}}{\text{Voltage} - \text{noisefloor}}\right)$$

$$= 20 \cdot \log_{10}\left(\frac{\text{FullWell}}{\text{noisefloor} - \text{in} - \text{electrons}}\right)$$

There are a lot of other important features of a sensor which should be optimized according to the specific X-ray sensing application. Some of these parameters are sensor frame rate (FR) in units of frames per second (FPS). This can be very low, down to 1–3 FPS for basic mammography, larger for fluoroscopy where it is usually similar to a video standard of 30 FPS, and can go up even higher to hundreds of FPS for dental CT. Another varying requirement is the linearity of the response. Depending on the application and the algorithms being used, it can vary from a 'forgiving' application that can live with even 5% of non-linearity, to others that need linearity of better than 0.1% (as in CT applications).

12.7   X-ray sensors: types and requirements

Different X-ray sensors and their specifications are listed in Table 12.1. In Table 12.2 and Table 12.3 there is a list of several sensors of various vendors

Table 12.1 Different X-ray applications and the typical requirements for each application

Application	Size (cm²)	Frame rate (fps)	Typical Energy (keV)	Resolution (at patient) (lp/mm)
Radiography	43 × 43	0.05–2	40–60	3–4
Angiography	30 × 40	2–30	33	1–2
Full-field mammography	24 × 30	0.05–2	17–25	11
Mammography, CT tomosynthesis	S: 18 × 24 L: 24 × 30	5	–	10–15
Mammography, biopsy	5 × 10	0.05–2	17–25	11–20
Cardio-vascular	20 × 20	15–60	40–60	1–2
Fluoroscopy	43 × 43	30	–	–
CBCT	S: 6 × 6 M: 13 × 13 L: 20 × 25	30	80–90	–
Dental Panoramic	15 × 0.8	300	80–90	–

Table 12.2 Examples of different X-ray sensors from different vendors and their main parameters (part 1)

Manufacturer	Reference	Wafer size	Image size	Pixel size (um)	Resolution (Mp)	FF	FPS	Pixel type	Full well
Hamamatsu	8	8	12 × 12	50 × 50	5.8	79	2	–	2.2 M
Hamamatsu	9	12	22 × 18	50 × 50	15.5	76	1	–	
Vatech	10	8	12 × 14	200 × 200	0.5	67	30	4T	10 M
Vatech	11	12	23 × 17	70 × 70	8.5	75	3	3T	17 M
Rutherford	12	8	14 × 12	50 × 50	–	–	–	–	–
Canon	13	12	20 × 20	160 × 160	1.6	–	100	Smart	77 k
U. Lincoln	14	8	13 × 13	50 × 50	6.5	70	30	3T	280 k
T-DALSA 1	15	8	14 × 8	33.5 × 33.5	10	84	8.7	3T	650 k
T-DALSA 2	16	8	13 × 13	100 × 100	1.7	85	45	3T	500 k
Thales	4	8	12 × 15	150 × 150	0.75	80	30	Smart	5.2 M

Table 12.3 Examples of different X-ray sensors from different vendors and their main parameters (part 2)

Manufacturer	Reference	DR	QE (%)	Noise floor	Output	Comments
Hamamatsu	8	–	–	1100e-	Analog	–
Hamamatsu	9	–	–		Analog	–
Vatech	10	84 dB	55	910e-	Analog	–
Vatech	11	84 dB	45		Analog	–
Rutherford	12	–	–	30e-	Analog	3-side buttable
Canon	13	75 dB	–	13e-	Analog	In pixel circuitry PGA, S/H
U. Lincoln	14	65 dB	45	140e-	Analog	2 × 2 charge binning 2-side buttable
T-DALSA 1	15	71 dB	50	175e-	Analog	Dose sensing binning 3-side buttable
T-DALSA 2	16	70 dB	50	135e-	Analog	Dose sensing binning 3-side buttable
Thales	12	96 dB	65	87e-	Digital	2 × 2 binning 3-side buttable, real HDR

Table 12.4 Parameters of SiX HD sensor of Thales

Sensitivity	8.47 lsb/nGy
QE × fill factor	65%
CG (high/low)	2.32/0.29 µV/e
Noise	2.19 lsb, 147 e^-
NED	0.16 nGy
Maximum Dose	11 µGy
MTF (Nyquist)	15%
DQE(0)	71%
Average dark current	1.9 pA/cm^2, 2800 e^-/sec per pixel
Lag	0.12%

with their main parameters and features. This list is changing fast due to growing interest in CIS sensors for medical X-ray. As an example of a state-of-the-art sensor, more details on a Thales sensor are given in Table 12.4. The table shows the sensitivity and noise parameter of the sensor in its HDR mode (lsb stands for least significant bit).
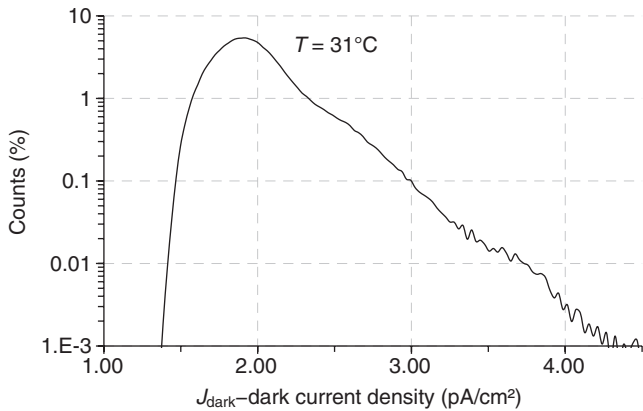
12.10 Dark current distribution for Thales sensor.
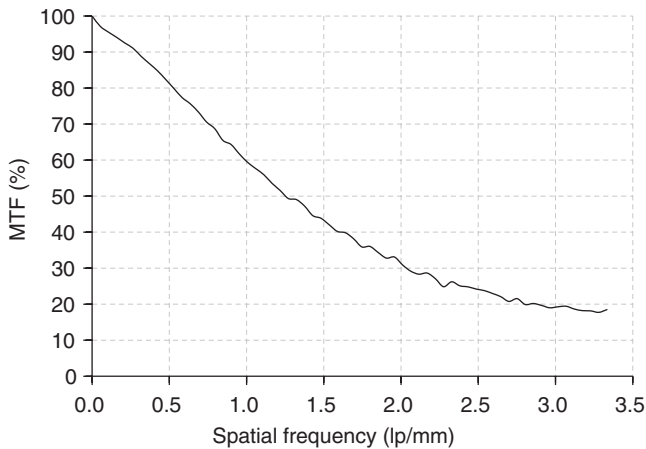
12.11 MTF curve for Thales SiX HD sensor, pixel size is 150 microns.

The dark current distribution is shown in Fig. 12.10. This parameter is usually measured by the dark current per unit area (pA/cm^2). The actual dark current per pixel is reported in (e$^-$/sec). The MTF and the DQE curves of the Thales HD sensor is shown in Figs. 12.11 and 12.12, respectively. The SNR versus radiation dose is presented in Fig. 12.13. The 10 dB/decade, namely noise which is proportional to the square root of the dose is a 'pure' shot noise of the X-ray photons.

12.8   Direct X-ray sensors

CIS X-ray sensors were presented in the context of application of image sensors to X-ray sensing. This naturally led to the so called 'indirect' X-ray
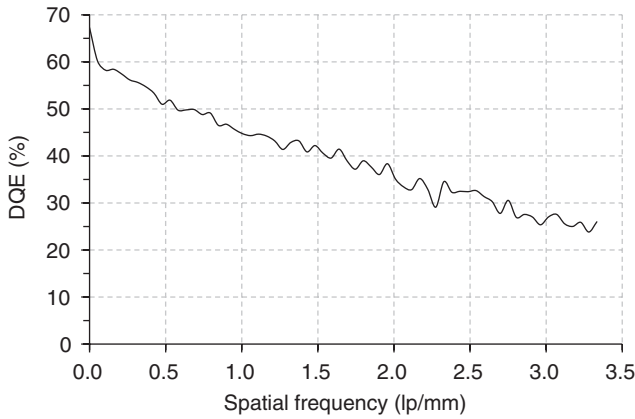
12.12 DQE curve of Thales SiX HD sensor.
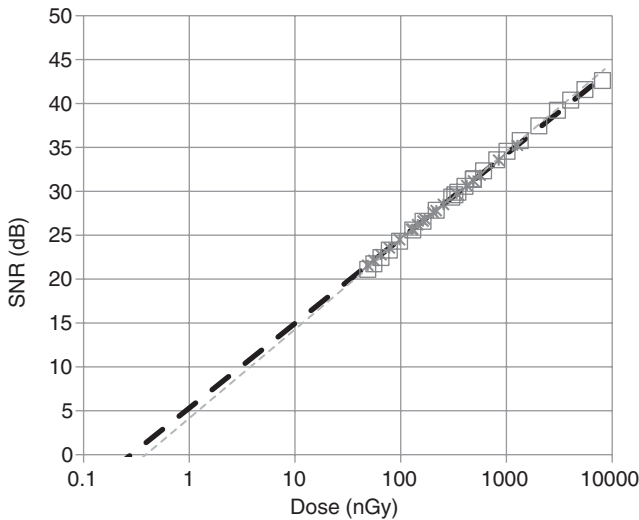
12.13 SNR in dB vs. radiation dose.

sensing which is based on the conversion of X-ray photons to visible photons. The visible photons in turn are sensed by the CIS photodiode. Another approach to X-ray sensing is direct conversion of X-ray photons to electrical charge. In most cases, these direct sensors use X-ray photoconductors [17]. The simplest possible pixel in this case will have, in addition to the X-ray sensitive photoconductor, a capacitor and an addressing transistor as shown in Fig. 12.14.

The capacitor is charged to a certain voltage, and then discharged by a photocurrent produced by the X-ray photons. A lot of crystals can be used
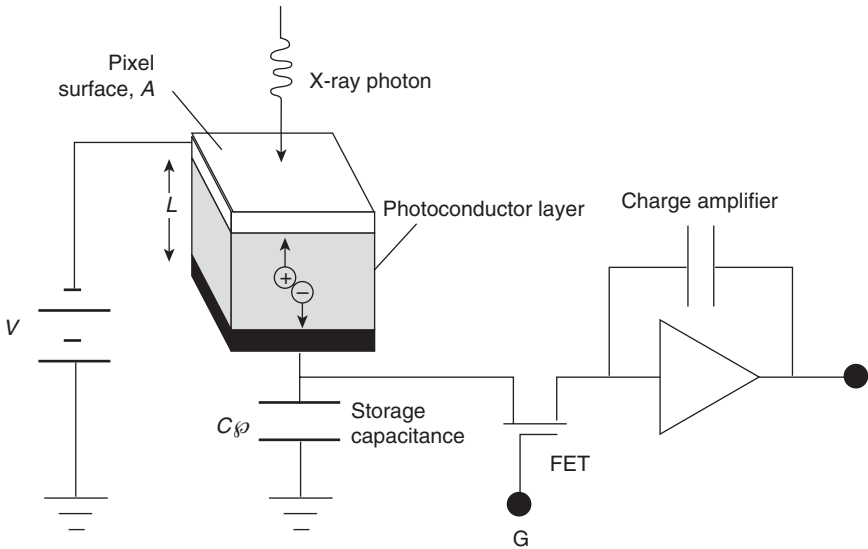
12.14 Direct conversion X-ray pixel. The photoconductive element, the capacitor and the addressing FET are within the pixel, while the amplifier is for the whole column and can be off chip.

as an X-ray photoconductor, but the one which is already commercially available is based on amorphous selenium (α-Se). The operation of such a photoconductor requires rather high voltages; a few kVolts in the case of α-Se. The main requirements from such a photoconductor are good quantum efficiency, X-ray sensitivity and low noise. In addition there are other requirements like uniformity, reliability and low defect count. The main noise source in such a pixel is the dark current of the photoconductor. This current is in the range of nA/cm2 for most of the candidate materials excluding α-Se which is in the range of tens of nA/cm2. The most significant advantage of direct sensors over indirect sensors is related to MTF. The scinitillator used to convert X-ray photons to visible photons emits these visible photons in all directions which smears the signal on several pixels. The most advanced scintillators, like needle (column) shaped CsI(Tl) improves dramatically the MTF by wave guiding the visible light into the underlying pixel [18]. But, even this improved MTF is still inferior to the geometrical ideal MTF as can be seen in Fig. 12.11 where the MTF at Nyquist frequency (3.3 lp/mm) drops much below the theoretical 64% value. In contrast, the MTF of direct conversion X-ray sensors is very close to the theoretical value.

Comparing the direct conversion photoconductive sensors to indirect CMOS X-ray sensors, the noise performance and the quantum efficiency (QE) are still much better in the latter. However, the better MTF is

important especially in applications where small pixels and high resolution are required [19]. Beyond the inferior noise performance, direct conversion sensors need improvement in their reliability (mainly due to the high voltage required) and the defect density level. Thus, currently indirect conversion sensors are still the mostly used sensors. Yet, a lot of work is invested in this field and only the future will tell which of these approaches will win.

12.9   Conclusion and future trends

CIS sensors for X-ray applications is a fast growing and developing niche of the medical X-ray field. In the case of intra oral dental, CIS sensors are almost completely replacing CCD sensors, and in medical applications, increasingly competing with the other players like the a:Si panels. Though inherently suffering from costly VLSI Si process and wafer dimensions limits, CIS sensors are winning in almost all other fronts. CIS has better basic X-ray sensor performance parameters like QE, MTF, speed, noisefloor, and power consumption. However, the most promising aspect of CIS for X-ray applications is the open gate for a whole world of sophisticated functions offered by the availability of modern in-pixel VLSI circuitry. The super high dynamic range, and the 'colored' X-ray sensor examples can give small hints of the torrent of new innovative ideas which are still to come. Careful design for yield, and good tiling technology can partially compensate for the CIS drawbacks mentioned above and make wide use of the superior CIS technology. Another emerging technology which can have an impact on X-ray sensors is the integration of single photon avalanche photodiodes (SPADs) into the CMOS sensors process. SPADs are superb photon counting devices with excellent time resolution; features which can find their use in future advanced X-ray sensors.

12.10   References

1. G. Zentai, 'Comparison of CMOS and a-Si flat panel imagers for X-ray imaging,' in *2011 IEEE International Conference on Imaging Systems and Techniques (IST)*, pp. 194–200, 2011.
2. A. Lahav, T. Leitner, R. Reshef and A. Fenigstein, 'Enhanced X-RAY CMOS sensor panel for radio and fluoro application using a low noise charge amplifier pixel with a partially pinned PD,' *IISW 2011*.
3. J. Bosiers, L. Korthout, and I. Peters, 'Medical X-ray imaging using wafer-scale CMOS imagers,' *6th Fraunhofer IMS Workshop on CMOS Imaging*, June, pp. 12–13, 2012.
4. A. Fenigstein, R. Reshef, S. Alfassi and G. Yehudian, 'Single-exposure high dynamic range CMOS image sensorpixel with internal charge amplifier', not yet published.

5. B. Dierickx, B. Dupont, A. Defernez and N. Ahmed, 'Color X-ray photon counting image sensing,' *IISW 2011*.

6. E. Sarig, R. Reshef, S. Alfassi and D. Cohen, 'Horizontal row drivers for CMOS image sensor with tiling on three edges,' Google Patents, 2010.

7. A. Rose, 'The sensitivity performance of the human eye on an absolute scale,' *JOSA*, 38(2), 196–208, 1948.

8. H. Mori, R. Kyuushima, K. Fujita and M. Honda, 'High resolution and high sensitivity CMOS panel sensors for X-ray imaging,' *2001 IEEE Nuclear Science Symposium Conference Record (Cat. No.01CH37310)*, vol. 1, pp. 29–33, 2002.

9. K. Fujita, H. Mori, R. Kyuushima, M. Honda, K. Yamamoto and A. C. I. Sensor, 'High resolution large formatted CMOS flat panel sensors for X-ray,' pp. 2114–2118, 2004.

10. S. K. Heo, S. K. Park, S. H. Hwang, D. A. Im, J. Kosonen, T. W. Kim, S. Yun and H. K. Kim, 'Development of a large-area CMOS-based detector for real-time X-ray imaging,' in *Proc. SPIE Medical Imaging*, 76223, 2010.

11. T. W. S. Keoa, J. Kosonena, S.H. Hwanga, H. K. K. Kima and S. Yunb, '12-inch-wafer-scale CMOS active-pixel sensor for digital mammography(empty).pdf,' *Physics of Medical Imaging, Proc. of SPIE*, 7961, 2011.

12. R. Turchetta, N. Guerrini and I. Sedgwick, 'Large area CMOS image sensors,' *Image Sensors Europe*, Jan. 2011.

13. Y. Yamashita, H. Takahashi, S. Kikuchi, K. Ota, M. Fujita, S. Hirayama, T. Kanou, S. Hashimoto, G. Momma and S. Inoue, 'A 300 mm wafer-size CMOS image sensor with in-pixel voltage-gain amplifier and column-level differential readout circuitry,' in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International*, pp. 408–410, 2011.

14. M. Esposito, T. Anaxagoras, A. Fant, K. Wells, A. Konstantinidis, J. P. F. Osmond, P. M. Evans, R. D. Speller and N. M. Allinson, 'DynAMITe: a wafer scale sensor for biomedical applications,' *Journal of Instrumentation*, 6(12), C12064, 2011.

15. L. Korthout, D. Verbugt, J. Timpert, A. Mierop, W. de Haan, W. Maes, J. de Meulmeester, W. Muhammad, B. Dillen, H. Stoldt *et al.*, '*A wafer-scale CMOS APS imager for medical X-ray applications*,' in *Proc. of 2009 International Image Sensor Workshop*, vol. 70, pp. 22–28, 2009.

16. http://www.teledynedalsa.com/ls/products/cameradetail.aspx?partNumber=Xineos-1313-EO-CL.

17. S. O. Kasap and J. A. Rowlands, 'Direct-conversion flat-panel X-ray image sensors for digital radiography,' *Proceedings of the IEEE*, 90(4), 591–604, 2002.

18. M. Nikl, 'Scintillation detectors for X-rays,' *Measurement Science and Technology*, 17(4), R37, 2006.

19. E. Samei and M. J. Flynn, 'An experimental comparison of detector performance for direct and indirect digital radiography systems,' *Medical Physics*, 30, 608, 2003.

<div align="right">

13

</div>

Complementary metal-oxide-semiconductor (CMOS) and charge coupled device (CCD) image sensors in high-definition TV imaging

P. C E N T E N, Grass Valley, Nederland BV, The Netherlands

Abstract: The particulars of the image diagonal, signal-to-noise ratio, artifacts, modulation transfer function and aliasing in broadcast cameras in relation to imagers and imaging are explained. Examples of advances in broadcast imagers are given and the transition from charge coupled device (CCD) to complementary metal-oxide-semiconductor (CMOS) imagers is described.

Key words: broadcast camera, CCD imager, CMOS imager.

13.1   Introduction

Broadcast cameras are synonymous with having the latest technology inside. Imagers must have the lowest noise, the highest sensitivity, the lowest smear, the best highlight handling and the lowest aliasing (but not too low!). The target is faithful reproduction. The mainstream broadcast camera has a color-splitter with red, green and blue (RGB) optical outputs and in each channel a monochrome imager: the 3-imager camera. Many of the performance parameters relate to visual experiments and just noticeable differences (JND). The thresholds can be traced back to the Weber–Fechner law and the experiments of Blackwell (1946). It is important to be aware that parameters, when taken at face value, are not what they seem to be. As an example, when people say they use an imager with a ⅔-inch image-diagonal they actually mean that the light sensitive part of the imager within the scanning format is 11.0 mm and not 16.93 mm.

Cinematography is the domain for large format single imager cameras. Imagers are intrinsic monochrome. The colors are defined through, for example RGB-stripe filters (Genesis from Sony and Panavision) on top of the imager, or RGB-Bayer patterned color filters (ALEXA from ARRI, EPIC from RED, ORIGIN from DALSA), or RGB-diagonal (Panavision Inc.). By introducing 2048 (2k) sensitive pixels per line, cinematography has made a small deviation from the HDTV scanning format which counts 1920 sensitive pixels per line. The number of vertical scanning lines is 1080 for both HDTV and cinematography. The 3-imager ⅔-inch HDTV cameras,

like Sony's CineAlta (Thorpe, 2000b) and Grass Valley's VIPER (van Rooy *et al.*, 2002), were the first to be used in digital cinematography and set an example for the large image format cinematography cameras (Thorpe 2013).

13.2   Broadcast camera performance

In assessing broadcast cameras there are more than 120 performance parameters that matter. Some are measurable and others are subjective evaluations based on observation under special viewing conditions. Many of the parameters are related to imager design and imager yield, like: leaking pixels, blemishes, reflections and flare. Others relate to Moore's law and improve year on year like: read-noise, sensitivity, maximum charge that can be collected in a pixel (Q_{max}) and darkcurrent induced fixed pattern noise (FPN). Many aspects of the performance of a broadcast camera are defined in standards, e.g. REC. 709. This covers aspects such as colorimetry, transfer curve, white point, interfacing (HD-SDI), number of scanning lines, number of pixels, frame rate, aspect ratio, definition of luminance (Y), color-difference signals (CrCb), conversion of RGB to YCrCb and back.

A key method for measuring and comparing camera performance is the 'linear mode'. Imagine your camera is used in a shootout next to 5 others. It generates beautiful pictures. Then instead of putting numbers to what one sees you are asked: 'to switch off gamma, set the matrix into 1:1, switch contours off, color temperature at 3200 K, exposure nominal, switch the camera to 0 dB gain and maybe lift the black a little bit'. Then some camera numbers, like the rms-noise in black, the lens *f*-number to reach 100% output and MTF at 27 MHz are measured. Clearly this is not the operational mode of the camera. The parameters are measured in a state in which a camera will never be used. The method was introduced many years ago for comparing and measuring the camera in an objective fashion and is known as the linear mode.

The reason for switching the camera into a state which is not the operational state stems from the fact that the gamma curve in a camera is non-linear. The differential gain for small signals, like noise, is large for small output levels. At a higher output level the differential gain for small signals goes down. Switching off the gamma removes that ambiguity. The present noise measurement is measuring the readnoise only. The photon induced shotnoise is not measured (Janesick, 1987). This kind of noise is visible in the gray exposed parts.

Cameras in broadcast have a long life cycle. Once a camera is introduced in the market it will be available for some seven years. On average service parts need to be available for five to seven years. As a consequence the same imagers must be available for more than 10–14 years. Many of

the imager specifications relate to non visibility of artifacts. They stem from days when video cameras were equipped with cathode ray tubes (e.g. Plumbicon) where image quality was quantified through panelist observations arriving at the notion of just noticeable differences (JNDs).

Humans are better at detecting changes in contrast rather than absolute luminance values. For a given set of viewing circumstances, a JND of contrast is the threshold luminance difference between two areas that is required to just detect that difference. Detection (perception) accuracy is dependent upon size of the object or pattern (spatial frequency) and the time span that it is visible (temporal frequency). The threshold values can be traced back to the Weber–Fechner law and Blackwell's (1946) data on contrast perception.

Some of the artifacts can be categorized as offset errors and others as gain errors. With offset errors the pixel signal is the summation of the wanted signal and an offset. This is the case with leaking pixels and FPN. Next are the gain errors caused by differences in pixel sensitivity or photo response non uniformity (PRNU). Spatially there are column artifacts, row artifacts and large area artifacts.

Another offset error is the dark current induced fixed pattern noise. When the FPN is 16 dB below the readnoise it is camouflaged by the readnoise and not visible. The 50% detection limit for visibility of FPN is at 10 dB below the readnoise. A gain error, or the photo random non uniformity, is the difference between pixels and should be less than 1% of the local value. Since the eye is even more sensitive for changes in large areas, like with column stripes, the threshold value is at 0.06% of nominal exposure.

A broadcast video camera has a 0 dB setpoint and around that point the gain can be changed from, for example −6 dB to +12 dB and beyond. The 0 dB is defined through the sensitivity and the noise of the camera. In SDTV the noise in 0 dB is around 60 dB in luminance (Y) in a 5 MHz bandwidth and was perceptually noise free. At the beginning of HDTV the noise in luminance(Y) was around 54 dB in 30 MHz bandwidth and has recently reached 60 dB. The 0 dB is defined as: given a scene illumination of 2000 lux, 3200 K and 89.9% scene reflection what is the f-number to reach the given noise numbers. At present they are in the range of f/8-f/12 for 1080i60 or 720p60 to reach 60 dB in Y.

Before the light hits the imager it has to pass a number of glass elements which are part of the optical chain in the camera. For proper colorimetry the infra red (IR) filter and the colorsplitter for creating the three primary colors RGB are needed. The horizontal and vertical optical low pass filters (H-OLP and V-OLP) consist of a set of birefringence crystals to reduce the aliasing. Next a retardation plate is used to change the linear polarization of light (caused by specular reflections on a shiny surface) into circular

polarization again. The lens creates an image on the imager surface. Broadcast lenses have different focal planes for R and G and B to compensate for the optical path differences in the colorsplitter. As a consequence, when using broadcast lenses together with a single imager the green will be in focus and R and B out of focus. The colorsplitter often supports down to $f/1.4$, and lenses are used from $f/1.4$ upwards. To cover large illumination levels neutral density (ND) filters can be switched in the optical path. For artistry, special effect filters can be used too.

13.3   Modulation transfer function (MTF), aliasing and resolution

The modulation transfer function (MTF) relates to the sinodal transfer function of a camera system for spatial frequencies. In practice the MTF is measured with a square wave instead.

This measurement is also known as the contrast transfer function (CTF). In practice MTF is determined by normalizing the 100 kHz square wave camera output to 100% and measuring the output at 27 MHz relative to the 100 kHz. The 27 MHz at the camera output is a sine wave since the higher harmonics of 27 MHz do not pass the 30 MHz regulatory bandpass filter; practically gaining about 1.2 in MTF compared to sine wave values. This can be understood by doing a Fourier decomposition of a square wave. Given that the square wave has an amplitude of 1.00, then the amplitude of the fundamental sine wave is 1.27 (= $4/\pi$). Square wave test charts, as opposed to sinodal, are cheaper to produce and give higher MTF numbers. Where the pixel aperture determines the MTF of the imager, the repetition of the pixel determines the aliasing. MTF and aliasing together give rise to the resolution.

Having an asymmetric pixel is no problem as long as the pixel is repeated. An odd/even mirroring of the pixel layout results in aliasing at subharmonics. A practical method to judge MTF and aliasing of a camera system is the ZONE-chart (Drewery, 1979). The ZONE-chart (Fig. 13.1) is a two dimensional frequency sweep and is very helpful in showing the frequency response of the camera system and its artifacts at a glance.

To achieve proper aliasing performance optical low pass filters are applied. Often they operate for horizontal and vertical spatial frequencies and have a single or double notch at the sampling frequency. Only in frame transfer charge coupled device (FT-CCD) (Theuwissen et al., 1991, Theuwissen, 1995) was it possible to engineer the vertical MTF such that one could interchange MTF and aliasing (Centen, 1995; Stoldt et al., 1996) through filtering in the charge domain.

The MTF is the frequency response of the camera system to the excitation of an optical sine wave normalized to a low frequency, ideally 0 Hz. The
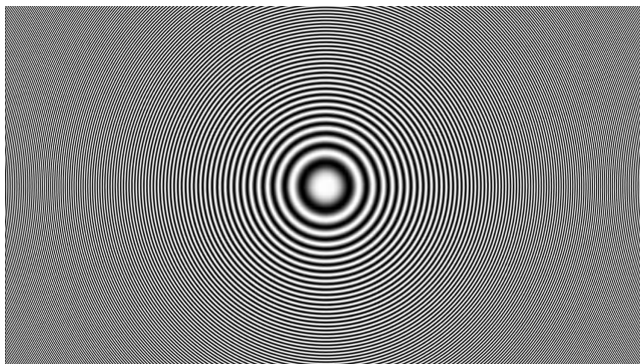
13.1 Zone chart; a two dimensional spatial frequency sweep.

(a)   (b)
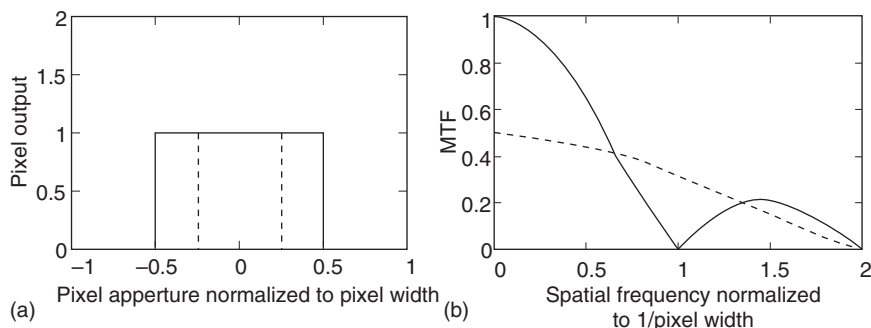
13.2 Solid trace: (a) a pixel aperture of 100% and (b) the MTF of the pixel. Dotted trace idem for aperture 50%.

MTF of a pixel is a $\sin(x)/x$ function with zero transfer slightly above the inverse of the pixel width. Figure 13.2 depicts the effect pixel aperture (100% and 50%) has on MTF. In the monochrome case and with μ-lens the pixels are co-sited and aperture will be close to 100% of pixel width and MTF at Nyquist will be in the range of 64%. In the Bayer'd case the pixels belonging to one color are separated with a pixel of the other color and aperture will be 50% or less. The horizontal sample frequency for R or B or G is half the monochrome case.

The MTF of broadcast lenses is very well described with the diffraction limit equation

$$\mathrm{MTF}(f_s) := 1 - 1.22 \cdot F \cdot \lambda \cdot f_s \tag{13.1}$$

where F is the *f*-number of the lens, λ the optical wavelength and f_s the spatial frequency. Broadcast lenses create an image as if at infinity, and hence, one has no need for the μ-lens shift to be applied to the imager.
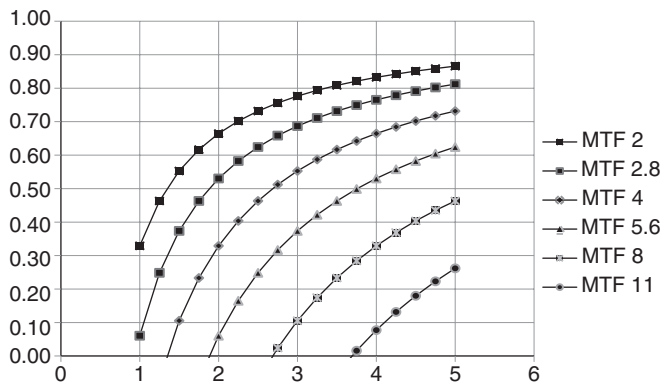
13.3 Y-axis: MTF at Nyquist (= 1/(2 × pixel_pitch)), *X*-axis: the pixel pitch in μm. Curves have *f*-number as parameter. λ = 550 nm.
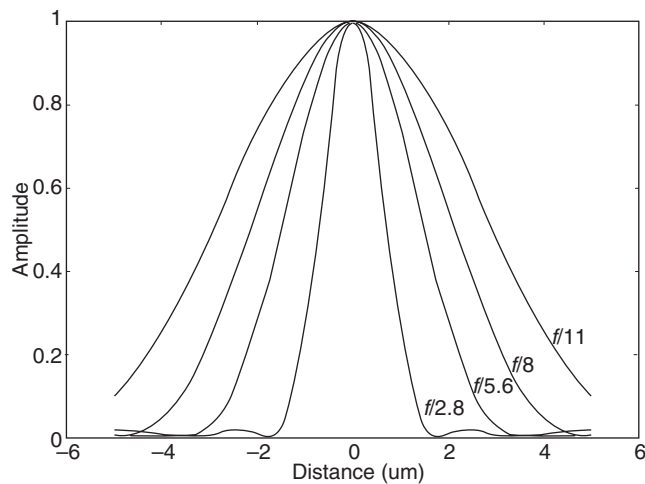
13.4 Airy disk for several f/numbers at 520 nm.

In Fig. 13.3 the diffraction limited lens MTF is plotted against the pixel-width with the *f*-number as a parameter. The MTF is at the Nyquist frequency. At *f*/11 a 5 μm pixel still has a MTF of slightly more than 25% and at *f*/5.6 even a 2.5 μm pixel can resolve with 25% at Nyquist. In Fig. 13.4 the pixel profile (Airy disc) at different *f*-numbers is given. The widening of the pixel profile for the higher *f*-numbers is obvious.

The MTF of a camera system is determined by the optics: lens, optical low pass filter, the pixel aperture, and the electrical processing: zero order hold, bandwidth limiting, low pass filtering, aperture correction. The camera
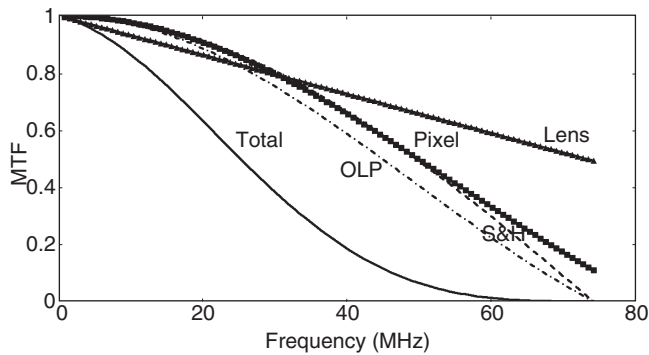
13.5 MTF curves of the electrical and optical elements in the camera. Lens at *f*/4 and 520 nm.

MTF will often be defined at *f*-numbers in the range of *f*/4–*f*/5.6 because at higher *f*-numbers the MTF drops due to increased diffraction and at smaller *f*-numbers chromatic aberration sets in. In Fig. 13.5 the MTF of the individual elements in a camera chain are plotted. The horizontal axis is in MHz and the system clock is 74.25 MHz with Nyquist at 37.125 MHz. In a ⅔-inch imager (5 µm pixel) the 74.25 MHz frequency corresponds to 200 lp/mm and Nyquist is at 100 lp/mm. The optical low pass filter has a cosine shape, the pixel a $\sin(x)/x$ just as the sample and hold.

In a charge coupled device (CCD) camera one will often find a 30 MHz low pass filter in the signal chain to prevent clock feed through and noise aliasing. In a complementary metal-oxide-semiconductor (CMOS) imager the processing is on a column basis and there is no need for an additional 30 MHz low pass filter and hence video signals have a higher spectral content.

13.4   Aliasing and optical low pass filtering

The imager is a two dimensional spatial sampling device. In general the scene that is viewed does not comply to the Nyquist criteria: i.e., no spatial frequency content above half the sampling frequency. There are no optical (brick-wall) filters that help fulfil the Nyquist's criteria. The next best solution is engineering optical filters that make use of the properties of human vision. The eye is very sensitive for low spatial frequencies and much less for high spatial frequencies. Frequencies close to the sample frequency will fold-back, or alias, to low frequencies and as a consequence are very visible.

The aperture of the pixel, which has a $\sin(x)/x$ shaped MTF, assists in reducing the aliasing. Only when the aperture is 100% the dip of the $\sin(x)/x$
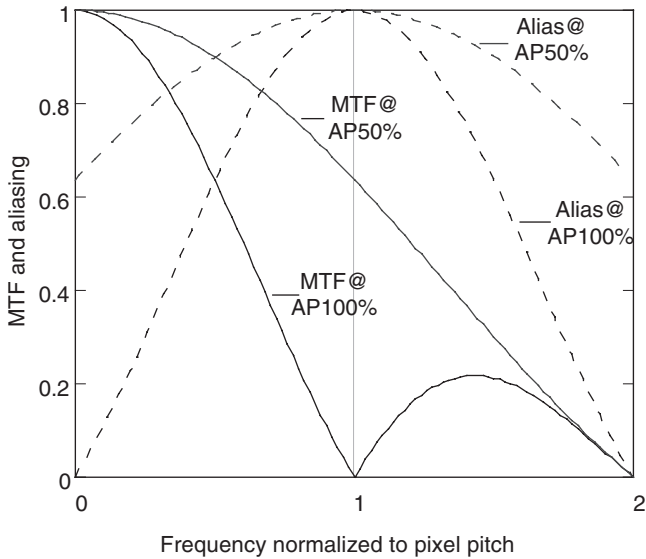
13.6 The solid curves are the MTF for a pixel aperture of AP = 100% and a pixel aperture of AP = 50%. The dotted curves are the corresponding aliasing. For a pixel aperture AP = 100% the MTF@ AP100% is 0% at the sample frequency (f_x = 1) and Alias@AP100% at f_x = 0 Hz is 0. For pixel aperture AP = 50%, the MTF@AP50% is 64% at the sample frequency and Alias@AP = 50% at f_x = 0 is about 64%. The vertical axis is the amplitude and horizontal axis the spatial frequency normalized to the inverse of the pixel pitch (sample frequency). Hence Nyquist is at f_x = 0.5. AP = pixel aperture.

is at the sample frequency. But the aperture of the pixel is always less than 100% and the dip in the MTF is above the optical sample frequency and alias suppression is insufficient. Figure 13.6 shows the MTF (solid) curve and its first alias (dotted) curve for a pixel with a 100% aperture (notch at f_s = 1) and one with a 50% aperture (notch at f_s = 2). The alias for the pixel with 100% aperture goes to zero for values near 0 Hz and the pixel with 50% aperture has 64% alias at 0 Hz.

With a bifringent crystal one can engineer an optical low pass filter. The MTF of the optical low pass filter is cosine shaped. The thickness of the crystal determines the frequency at which the transfer is zero. In engineering a notch at the sample frequency, the aliasing for frequencies close to the sample frequency will be much reduced and its visibility too. In a three-imager camera the R, G and B imagers have the same spatial sample frequency. The optical low pass filter has the same curve for the three colors.

In a Bayer'd, single imager camera (Thorpe, 2013), the sample frequency for R and B and G is the same. Aliasing in R and B is the same but for G
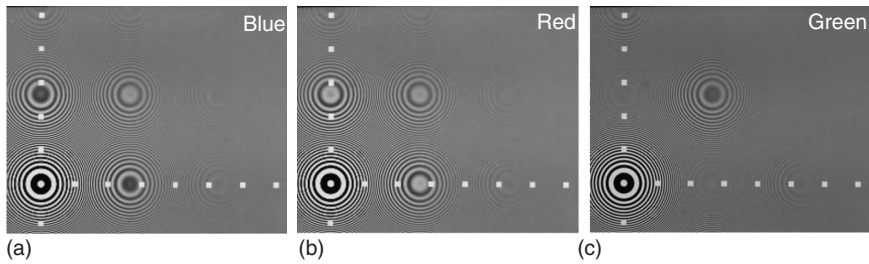
13.7 The RGB outputs of an imager with Bayer pattern when viewing a two dimensional frequency (zone chart). Note the phase-reversal for the vertical and horizontal aliases in R and B. Also note that in green the aliases have vanished where they still occur in R and B.

it is different because of the quincunx pattern (Fig. 13.7). The red and blue pixels are in a rectangular lattice and the aliasing has that pattern too. Green is in a quincunx lattice. Quincunx can be regarded as the addition of two rectangular lattices of which one has an offset to the other in the vertical and the horizontal direction. The addition of those two lattices causes the pure horizontal and the pure vertical aliasing at the sample frequency to be canceled (Fig. 13.7). In only green can it be used to effectively double the Nyquist frequency for pure vertical and horizontal frequencies; the doubling is compared to the Nyquist frequency for the red and blue lattice. The aliasing at the diagonal in green is not cancelled and is at a same position as the red and blue. (Figs. 13.7 and 13.8)

In Fig. 13.7 the red, green and blue output of a Bayer'd imager, when exited with a zone chart (Fig. 13.1), are shown. In the left lower corner the origin ($f_x = 0$ and $f_y = 0$) of the spatial sweep is situated. Along the vertical axis the first aliasing is clearly visible and weakly visible in the second aliasing. The same holds for the aliasing on the horizontal axis. In red and blue the phase of the first aliasing are opposite to each other. From a sampling point of view the red lattice is the blue lattice shifted one pixel in the vertical and horizontal direction. That shift causes the alias to change phase.

Figure 13.7 shows the outputs of RGB together when viewing a zone chart: (a) shows the blue pixel output, (b) the red pixel output and (c) the green pixel output. Figure 13.7 (a) and (b) show the expected aliases on a rectangular lattice. The green aliasing pattern is in quincuncx (Fig. 13.7(c)). Purely vertical and horizontal patterns can be resolved at double resolution compared with (a) and (b) (no alias at the same positions as (a) and (b)). Only at the diagonal they coexist.

Figure 13.8 shows the zone chart output of a single imager camera. The aliasing related to twice the sample frequency of the lattice is gray and all the other, sub-harmonics of those are colored. This figure shows the dilemma
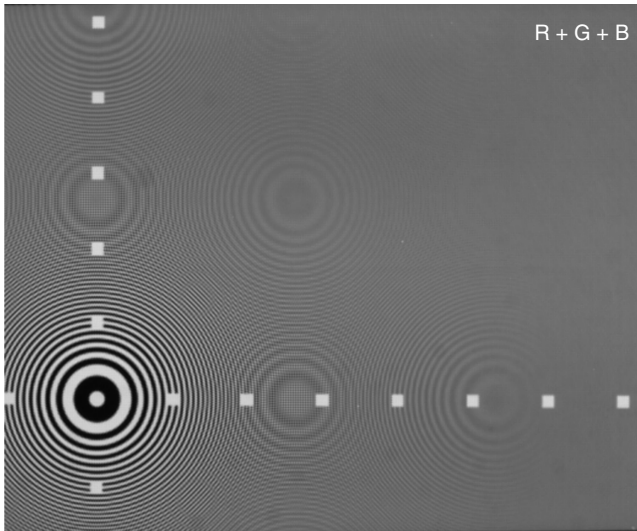
13.8 Two dimensional frequency sweep. The RGB outputs of an imager with Bayer pattern (Fig. 13.7) are added to obtain the luminance signal. The aliasing in horizontal and vertical direction occur at half the monochrome lattice resolution and at the monochrome lattice resolution.

one faces in engineering the dips for the optical low filter. Either one reduces the aliasing in R and B and at the same time reducing MTF in G, or one reduces the aliasing in G and is left with a high amount of colored aliasing at sub-harmonics.

When the signal is optically oversampled it is possible, in some cases, to apply proper optical filtering through the weighted addition of pixels. A nice example is the creation of an interlaced image from a progressive image through the addition of two neighboring pixels. Now the vertical MTF is cosine shaped with a notch at exactly the field sample frequency (Thorpe, 2000a). In general, combining the pixel signals will exhibit optical filtering properties but often the dips do not match sample frequencies or filter already far below the Nyquist frequency removing too much of the wanted signal.

Another example of optical low filtering in the charge domain became known as dynamic pixel management. The pixel is shifted vertically, in sub-pixel steps, in an up-down pattern. The method allows creating the dips at the sample frequency and shape the MTF in a rather detailed way. It is the only method known that has no need for an additional vertical optical low pass filter. (Centen *et al.*, 1995; Stoldt *et al.*, 1996).

In cinematography there is a multitude of sampling structures found: rhombic pixels (Toyama *et al.*, 2011, F65), diagonally striped colorfilters
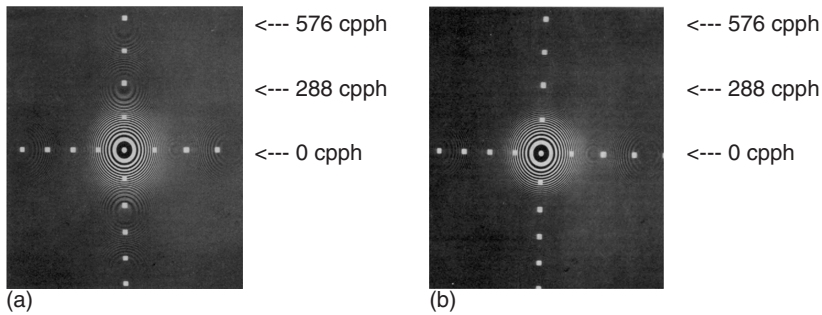
13.9 The effect of low pass filtering: (a) without optical low pass filtering and (b) with optical low pass filtering.

(Panavison) and column striped (Sony, Genesis). Rhombic shaped pixels have twice the sensitivity of their rectangular counterparts at the cost of halving the resolution at the diagonal. In Genesis one applied the vertical RGB-stripe filters, which is still the best from an aliasing reduction point of view. The diagonal type from (Panavison) is less straightforward since it has a dominant diagonal which generates different responses in the left angled diagonal compared to the right angled one.

The effect an optical low pass filter has on aliasing is shown in Fig. 13.9. At 288 cpph there is aliasing in (a) and it vanishes in (b). On the horizontal axis the aliasing is visible in both figures showing that there is sufficient spatial content.

A numerical example can be seen in a HDTV imager with 5 um pixel. Assume a diffraction limited lens at *f*/4, and an optical low pass (OLP) filter with a notch at 200 lp/mm for aliasing suppression. At 27 MHz (equals $27/74.25 \times 200 = 73$ lp/mm or $27/74.25 \times 1920 \times 2 \times 9/16 = 785$ TV lines (TVL)) the MTF values for the separate elements are:

$$\text{lens at } f/4 = 1 - 1.22 * 4 * (520 * 10^{-9}) * (73 * 10^{3}) = 82\% @ 520 \text{ nm}$$

$$\text{OLP} = \cos\left(\frac{\pi}{2} * \frac{73}{200}\right) = 84\%$$

$$\text{pixel} = \frac{\sin\left(\pi * \dfrac{73}{200}\right)}{\pi * \dfrac{73}{200}} = 80\%$$

$$\text{S\&H} = \frac{\sin\left(\pi * \dfrac{27}{74.25}\right)}{\pi * \dfrac{27}{74.25}} = 80\%$$

Sinodal MTF at 27 MHz of all elements = 43% or square wave = 52%.

The video camera is just an observer and cannot control the scene it is viewing. Hence all possible spatial frequencies will occur. There will be fine details in the scene (high spatial frequencies) that will violate Nyquist and there will be aliasing. In the end there is no MTF without aliasing. By proper engineering of the dips of the optical low pass filters one can have both MTF at higher frequencies and reduced low frequent aliasing.

With the well known Retma charts, resolution was often expressed in TVL equivalent. Here one judges where the wedge pattern starts to vanish (no MTF) or to distort (aliasing). That point at the chart then was regarded as the resolution and was expressed in TVL.

Expressing the horizontal MTF in TVL had some ambiguity. The horizontal resolution in TVL would change when the aspect ratio changes even when the number of pixels per line would not. Given an 8 Mpixel Bayer imager, then for every row there are 4 k pixels/row and in every even row 2 kpixels/row are red and 2 kpixels/row are green and in every odd row 2 kpixels/row are green and 2 kpixels/row are blue. The sample frequency relates to the 2 kpixels/row and the optical filtering too. Hence the 8 Mpixel imager will perform as if it was HDTV: 1920 × 1080 (Thorpe, 2013). This is why the three-imager HDTV video camera performed so well in a cinematographically environment.

13.5   Opto-electrical matching and other parameters

An important aspect of a broadcast camera is the opto-electrical matching between cameras. After matching, the cameras on a set generate the same look on a display. That implies the same opto-electrical transfercurve, colorimetry, angle of view and that the cameras must capture images at the same time instant. The latter puts a design constraint on individual imagers and the camera system as a whole.

Also mechanical stability and flatness of the die in the package is important. In a three-imager camera, imagers are aligned to the color splitter within sub-pixel accuracy (µm range). It must stay within sub-pixel accuracy during its life time of some five to seven years. In three-imager cameras the opto-electrical matching between imagers must be high as well. If not, a gray curve like the gamma test chart can turn colored at certain exposure values.

Next, the image diagonal is an important parameter. In contrast to the consumer market it is not free to choose. A high end camera will have imagers with an image diagonal of ⅔-inch. On one hand the ⅔-inch is dictated through the huge amount of expensive exchangeable lenses that customers have in stock. On the other hand it is because these are high performance lenses with respect to chromatic aberration, zoom, backfocus and MTF.

Table 13.1 Image diagonals in broadcast

Format	Image diagonal
1″	16.0 mm
2/3″	11.0 mm
1/2″	8.0 mm

When people say the imager is ⅔-inch they do not mean ⅔ × 25.4 = 16.9 mm but 11.0 mm, 1-inch is not 25.4 mm but 16.0 mm. It is a metric from the past, stemming from the tubes (Table 13.1). There are two flavors of ⅔-inch: in broadcast the image diagonal is 11.0 mm and in industrial vision it is 12.0 mm. Either one uses 16 mm/inch or 18 mm/inch for the conversion from inch to mm. The ⅔-inch lenses for broadcast support an image diagonal of 11.0 mm; outside the 11.0 mm diagonal vignetting increases rapidly – this is because of the drop in light-output at the edges of the lens.

The image diagonal relates to angle of view, MTF, depth-of-field (DOF) and aperture induced diffraction limit of the lenses and sensitivity. Assuming an equal scanning format than when changing the image diagonal from ⅔-inch to ½-inch the pixel area shrinks by a factor two. The ⅓-inch imager has a four times smaller pixel area. When the ⅔-inch imager would have $f/11$ for a given exposure then the ½-inch imager would need $f/8$ to obtain that same exposure and the ⅓-inch imager $f/5.6$. A 4k, ⁴⁄₃-inch single-imager camera and a three-imager ⅔-inch HDTV camera will have pixels with the same area and therefore the same sensitivity.

For indoor studio-use or outdoor-use the broadcast camera has several color-temperature presets. These presets are to obtain proper colorimetry (Lang, 2001; Poynton, 2003) under different types of illumination. The 3200 K preset is for scenes which are illuminated with tungsten type illuminators and 5600 K for outdoor use.

The sensitivity of a camera is defined and measured at a color temperature of 3200 K. Figure 13.10 shows the photonflux of a blackbody radiator at 3200 K, 5600 K and 20 000 K. The unit of the vertical axis is in the number of photons per (pixel) area in μm^2, per nm of wavelength and per lux.sec. Along the horizontal axis the optical wavelength is given in nm. The calculation of lux is through the CIE1931 photopic eye-weight curve ($V(\lambda)$), which has a maximum at 555 nm. As a consequence the curves cross each other close to the 555 nm wavelength.

A 3200 K radiator has its energy in the near-IR and not so much in blue (400 nm–480 nm). The blue channel receives almost three times lesser photons than the red or the green channel. That is why the blue channel is noisier. The luminance (Y) is the weighted sum of R, G and B. The weight
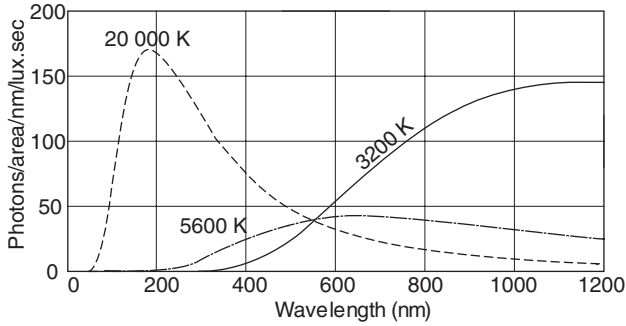
13.10 Photon flux of a blackbody radiator at 3200 K, 5600 K and 20 000 K. Horizontal axis has the optical wavelength in nm. Vertical axis has the number of photons per area in μm2 per nm optical wavelength per lux.sec.

for blue is 0.0722 (eqn [13.6]) and its contribution to the signal-to-noise in Y is small. The formula for calculating the number of photons, in green, impinging on a pixel for a color temperature T is:

$$\text{Photons } G(T) := \frac{A_{\text{cell}} \cdot T_{\text{int}}}{\eta \cdot h \cdot c} \cdot \frac{E_{\text{v}} \cdot \rho}{4 \cdot F^2} \cdot \frac{\int_{400}^{750} P(\lambda, T) \cdot \tau_{IR}(\lambda) \cdot \tau_{\text{lens}}(\lambda) \cdot \tau_{\text{KssG}}(\lambda) \cdot \lambda \, d\lambda}{\int_{400}^{750} V(\lambda) \cdot P(\lambda, T) \, d\lambda}$$

[13.2]

where P is the blackbody radiator with powerdensity in Watt/m³, the transmisison curves τ_{IR} for the IR filter, τ_{lens} the transmisson curve for the lens and τ_{KssG} the transmission curve of the green channel of the colorsplitter and finally V the CIE-1931 photopic luminosity function with a maximum of 1 at 555 nm. η is a constant with value 683 lm/watt, c the speed of light and planck's constant h. A_{cell} is the image cell area which is 25 μm² for a HDTV pixel in a ⅔-inch imager and T_{int} the integration time of the pixel of ¹⁄₆₀ sec. The scene illumination is E_{v} often equal to 2000 lux, ρ the reflection equal to 89.9% for the standardized measurement condition and finally F the f/number of the lens, e.g. f/11.

Dividing eqn 13.2 by eqn 13.3 one arives at the photon density in #photons/μm²/lux.sec for a given color channel.

$$\xi := A_{\text{cell}} \cdot T_{\text{int}} \cdot \frac{E_{\text{v}} \cdot \rho}{4 \cdot F^2}$$

[13.3]

In Table 13.2 the photon densities in RGB and monochrome are tabulated for a few blackbody radiator temperatures. To calculate the number of

Table 13.2 Photon density as a function of black body radiator temperature

Photon density	Number of photons/um²/lux.sec			
Color-temperature (K)	Red	Green	Blue	White + IR
2000	2790	1510	220	6057
3200	2069	1955	719	6076
5600	1587	2248	1635	6891
6500	1506	2295	1893	7173
20000	1218	2446	3277	8786

Table 13.3 Number of photons and electrons in a HDTV pixel of 5 μm as a function of master-gain for 100% video-signal at 2000 lux, 89.9%, f/11, 3200 K, QE = 60% and exposure of $\frac{1}{60}$ sec

	Red	Green	Blue
0 dB	3206 photons	3029 photons	1114 photons
	1924 electrons	1818 electrons	668 electrons
+6 dB	962 electrons	909 electrons	334 electrons
+12 dB	481 electrons	454 electrons	167 electrons

photons in the red, green and blue channel use is made of the color-transmittance curve from Martinez–Verdu *et al.* (2002). The concatenated transmissions in the pass-band of the lens, IR filter, color splitter is assumed at 80%.

The photon density in red and blue change strongly and opposite with each other with color temperature, whereas green is realtively constant. Assuming an overall conversion efficiency (QE_{eff}) of 60%,

$$QE_{\text{eff}} = QE * Aperture * (1 - refelction) * \mu lens\_gain \qquad [13.4]$$

then the charge accumulated in a 5 μm green pixel under nominal illumination conditions of 2000 lux, f/11, 89.9%, 3200 K is $G = 1818$ electrons. With an overexposure margin of 500% the maximum charge handling capacity, Q_{max}, must be larger than: $Q_{\text{max}} \geq 9.1$ kel. The number of signal electrons in Table 13.3 are given for the RGB channel at three mastergain levels of: 0 dB, +6 dB and +12 dB. With these number of signal electrons very nice images are made!

In everyday operation the imager must withstand very high light levels; for instance when shooting a scene in which an arc or a bright sun is in the background with an illumination of some 100000 lux or more. At these lux levels the photon flux increases up to 10^9 photons/μm²/sec. This amounts to 16 *f*-stops more photons being generated than under nominal conditions.
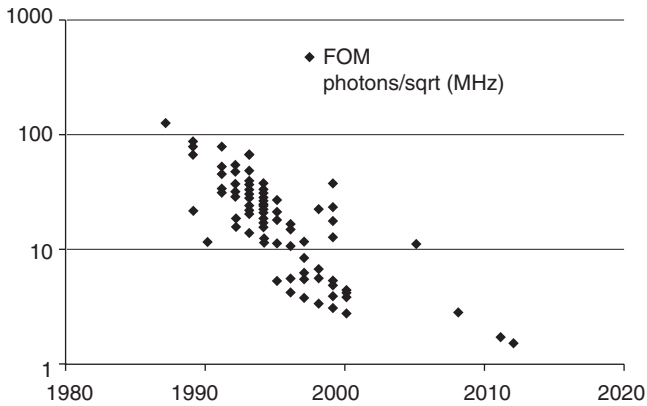
13.11 Noise and sensitivity improvements in broadcast cameras.

In Fig. 13.11 the read-noise per unit of bandwidth divided by the quantum efficiency is plotted for broadcast cameras. One can calculate this figure of merit (FOM) from eqn 13.5 by using the fact that the signal-to-noise in green is about 1.5 dB less than in Y. All the other numbers can be found in any brochure of a broadcast camera.

$$\text{FOM} = \frac{A_{\text{cell}} * T_{\text{int}} * E_{\text{v}} * \rho * 2.25}{4 * F^2 * \sqrt{BW} * 10^{\frac{|dBY|-1.5}{20}}} \tag{13.5}$$

FOM in photons/(unit bandwidth), A_{cell} is pixel area in μm^2, T_{int} is integration time in ms, E_{v} is scene illumination in lux, ρ is scene reflection, for broadcast standardized to 89.9%, F is aperture of the lens, BW: videobandwidth in MHz, dBY is the signal to noise in luminance channel in dB and 2.25 is a number with dimension in photons/μm^2/ms/lux.

Inspection of Fig. 13.11 reveals that the performance of imagers has increased by 1.5 dB/year for almost 30 years now. The better performance (lower numbers) in recent years has been due to the arrival of CMOS imagers.

13.6   Standards for describing the performance of broadcast cameras

ITU and SMPTE are two bodies that standardize many aspects in broadcasting including the camera and its interfacing. The camera transfer curve, known as gamma, colorimetry and frame rates for HDTV is determined through ITU-R BT. 709. The SMPTE 274M standard describes a 1920 × 1080 matrix of light sensitive pixels. Vertically there are a fixed

number of 1125 TV-lines. The number of horizontal clock counts is 2200 in i60, p30, i60/1.001 and p30/1.001. To accommodate lower framerates like p25 the horizontal blanking is increased proportionally.

The SMPTE 296M describes a 1280×720 matrix of light sensitive pixels. Vertically there is a fixed number of 750 TV-lines. The number of horizontal clock counts is 1650 in p60 and p60/1.001. To accommodate lower framerates like p25 the horizontal blanking is increased proportionally.

Typical framerate values for 50 Hz systems are: 25p, 50i and 50p and for 60 Hz systems: 23.976p, 24p, 29.97p, 30p, 59.94i, 60i, 59.94p and 60p. The system clock frequencies for 50 Hz systems are 74.25 MHz or twice the value for 1080p50: 148.5 MHz and for 60 Hz systems they are 74.25 MHz or 74.25/1.001 MHz or twice the value for 1080p: 148.5 MHz or 148.5/1.001 MHz. The 1.5 Gbps serial HD-SDI interface relates to the 74.xx MHz system clock and the 3 Gbps to the 148.yy MHz system clock.

When the imager complies to the same clock and pixel numbers, one does not have to translate from one clock domain to another and no additional artifacts are generated in terms of aliasing, cross color or resolution drop caused by scaling. Also the pixel can have its maximum area for the given resolution paving the way for an optimal performance of the imager.

13.7 Charge coupled device (CCD) and complementary metal-oxide-semiconductor (CMOS) image sensors used in broadcast cameras

In broadcast there are two aspect ratios: 4:3 and 16:9. HDTV and 16:9 are used synonymously. The VIPER camera, a cinematography camera based on a ⅔-inch HDTV camera, has an aspect ratio of 2.37:1, which is close to the Academy aspect ratio (standardized by the Academy of Motion Picture Arts and Sciences). The imager scanning was in 1920×1080. In cinematography one finds a slight deviation from the 16:9 (1.778:1) aspect ratio because they opted for 2048 pixels during the active line time while maintaining the 1080 vertical scanning lines (2048×1080). The aspect ratio is then slightly widened to 1.896:1. Imagers used in broadcast cameras are either the frame transfer type CCD, the interline CCD or more recently the CMOS imager in rolling shutter and the CMOS imager in global shutter mode:

- **FT: frame transfer** (Thorpe *et al.*, 1988; Theuwissen *et al.*, 1991; Theuwissen, 1995). Imager consists of an image area and a separate storage area below the image area. Chip area is about twice the image area. During the vertical blanking the charge image generated in the image area is transported in some tens of μs into the storage area. Then during each horizontal blanking interval a charge line is shifted in the

horizontal register. During the active line time the charge packets are shifted through the horizontal register to the output. Conversion of charge into voltage is at the floating diffusion. Using a mechanical shutter enables smearless image capture.

- **FIT: frame interline transfer** (Thorpe *et al.*, 1988; Harada *et al.*, 1992; Theuwissen, 1995). Idem as FT with the exception that the image area has two spatially separated functional areas: conversion of photons into electrons and storage. FIT achieves better smear suppression numbers than interline transfer (IT). With the advent of the pinned photo diode (Teranishi *et al.*, 1982), the imager reached very low dark current levels.
- **IT: interline transfer** (Thorpe *et al.*, 1988; Morimoto *et al.*, 1994; Theuwissen, 1995; Asano *et al.*, 2011). The chip is slightly larger than the image area. In recent years FIT has been overtaken by IT because smear suppression reached −120 dB and beyond. This is regarded as sufficient for broadcast.
- **CMOS: rolling shutter** (Kozlowski *et al.*, 2005; Centen *et al.*, 2007; Nitta *et al.*, 2006). Sufficient performance could be obtained with these imagers, even surpassing CCDs from SNR point of view. The pitfall was the rolling shutter distortion. For 50 frames and beyond it is more of an academic discussion. At 24 frames is is very visible and annoying.
- **CMOS: global shutter** (Centen *et al.*, 2013). With the introduction of the CMOS global shutter imager the CCD got its successor in broadcast.

CCD imagers for HDTV and SDTV scanning (HD-DPM) are designed to scan dedicated in 1080i/p (Blankevoort *et al.*, 1994; Morimoto *et al.*, 1994; Theuwissen *et al.*, 1991; Kozlowski *et al.*, 2005; Centen *et al.*, 2007; Asano *et al.*, 2011) or 720p (Spitzer *et al.*, 1996; Honda *et al.*, 2005). Imagers supporting 1080p can be readout in 720p using region of interest scanning (Kozlowski *et al.*, 2005). The penalty is a reduction of the 11 mm image diagonal down to 7.3 mm, which is slightly less than ½-inch format. This causes a substantial change in viewing angle. Another method is deriving 720p by downscaling from 1080p. In triple speed applications the HDTV-imager reaches a pixel output rate of 223 Mpixel/sec. The imager reported in Honda *et al.* (2005) supports up to 720p96 at 118 Mpixel/sec.

Building on DALSA's technology that employs thin transparent membrane gate electrodes and tungsten strapping (Peek 1999; Stoldt *et al.*, 1996; Bosiers *et al.*, 2002) a novel imager was designed that supports the 1080i, 1080p, 720p native scanning formats at constant image diagonal for single speed (Centen *et al.*, 2001) and triple speed (Centen *et al.*, 2009). The imager became known as the HD-DPM (high definition-dynamic pixel management) imager. Use is made of the simple numerical relationship between the scanning formats. By choosing 4320 pixels vertically

Table 13.4 Possible scanning formats at constant image diagonal, with 4320 vertical gates

16:9 scanning formats	Number of gates/pixel	In 12-phase clocking schema
1080p	4 (= 4320/1080)	4-phase clocking 1 pixel
720p	6 (= 4320/720)	6-phase clocking 1 pixel
1080i	8 (= 4320/(1080/2))	4-phase clocking 2 sub-pixels
480p	9 (= 4320/480)	3-phase clocking 3 sub-pixels
576i	15 (= 4320/(576/2))	3-phase clocking 5 sub-pixels
480i	18 (= 4320/(480/2))	6-phase clocking 3 sub-pixels
1080p in 2.37:1 aspect ratio	3 (= 4320*(3/4) / 1080)	3-phase clocking 1 pixel

(Table 13.4) the SDTV scanning formats (575i, 480i, 480p) could also be derived at that same image diagonal of 11 mm.

The vertical pixel profile is shaped through the application of a special pulse pattern to the row-gates in the image area. This enables shifting the pixel up and down in a predetermined fashion in sub-pixel shifts, effectively behaving as a charge domain anti-alias filter. By changing the pulse pattern, for driving row-gates, the pixel profile is adapted to the scanning format. A similar concept was used in a previous SDTV-imager (Centen *et al.*, 1995; Stoldt *et al.*, 1996) for changing the aspect ratio from 4:3 to 16:9. An SDTV high-speed camera system employing these 3× SDTV-imagers was reported by Moelands *et al.* (1998).

Figure 13.12 shows a schematic representation of the HD-DPM imager. The image area has 4464 gates (V) and 2040 columns (H). Only 4320 (V) × 1920 (H) are used to define the scanning format; the remaining positions are used for over scan and black reference pixels. The 4464 gates are connected through a 12-phase interconnect scheme. The height of one gate is 1.25 µm and the column width is 5 µm. In 1080i180 the two horizontal registers are clocked at 111.375 MHz each and the vertical transport frequency is 9.28 MHz.

The storage area has a fixed number of storage sites and includes the over scan and black reference pixels 2040 (H) × 1102 (V), and has a four-phase interconnect scheme. During the horizontal line blanking a (charge) TV line from the storage area is distributed over the two readout registers with the four-phases driven independently.

During the active line time, when horizontal transport is at 2 × 111.375 MHz, the horizontal register is driven in quasi-two phase mode.
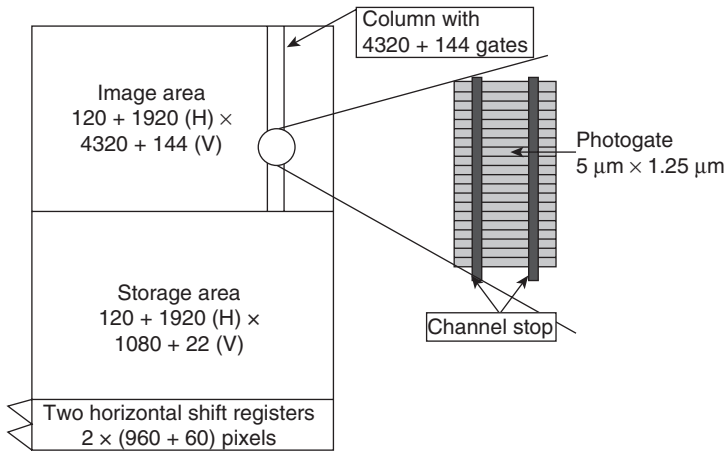
13.12 Schematic representation of the HD-DPM imager.

Additional tungsten straps on the horizontal gates reduce series resistance, maximizing transport speed. The layout of the two horizontal registers forces interleave between the pixels of top and bottom channel at the two outputs. Figure 13.13 denotes the possible vertical scanning formats in 16:9 aspect ratios by binning several patterns.

The first column in Table 13.4 describes the broadcast scanning formats. The 4320 gates per column enable 1080p, 1080i and 720p but also the NTSC (480p, 480i) and PAL (576i) scanning standards. Last but not least the cinemascope aspect ratio of 2.37:1 in 1080p was also supported. The second column shows the number of gates to meet the required native scanning format and the third column shows how to achieve the scanning format, given a 12-phase clocking and interconnect scheme (see also Fig. 13.13). Note that in 1080i the image area is scanned in 1080p and only in the horizontal registers (Fig. 13.12) the interlaced image is generated by adding two TV- lines of a 1080p image together. The active image width (H) is $1920 \times 5\,\mu m$ and the active image height (V) is $4320 \times 1.25\,\mu m$. This results in an image diagonal of 11.0 mm for all scanning formats as required for a ⅔-inch optical format.

In an old, but still very valid paper (Wong, 1996), it is argued that the ratio between lithographic feature size and pixel size is about 20 for complex type pixels. Given a ⅔-inch HDTV CMOS imager, with 5 μm pixels one needs a feature size of 0.25 μm or less. Loose *et al*. (2001) published the first ⅔-inch HDTV CMOS imager with soft-reset 3T-pixels in 0.25 μm technology.

A CCD needs a complicated power supply to generate all the voltages needed for vertical and horizontal transport, storage and reading; Often with complicated pulse-patterns, like changing the state of the gates in each
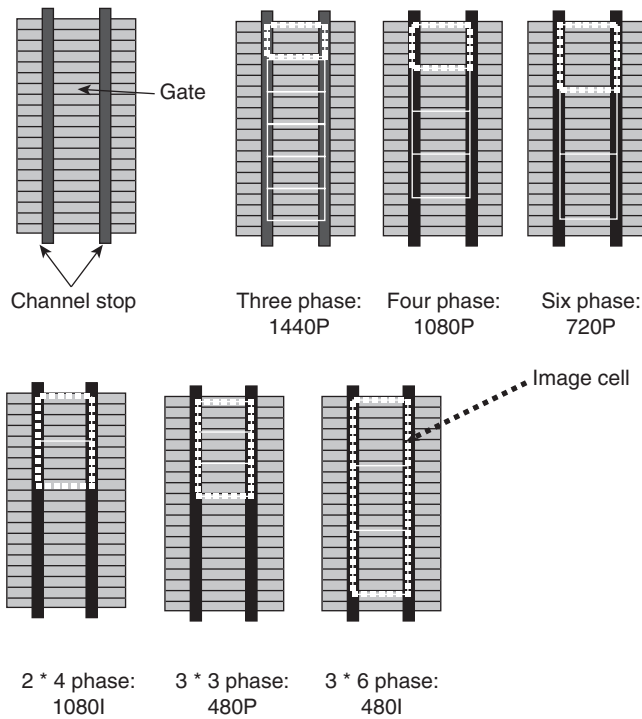
13.13 Various binning patterns.

line blanking to reduce FPN through pumping (surface inversion) or to enable optical filtering in the charge domain. External to the CCD are the analog pre-amplifiers and correlated double sampling (CDS). Next are the filtering for suppression of carrier feedtrough and bandwidth limiting to prevent sampling of noise outside the usable baseband.

The change from CCD to CMOS gave a reduction in power consumption of real estate (PCB) and of electronics, and the ability to integrate the full analog signal chain with the inclusion of the analog to digital converter (ADC) on-chip. Another attractive advantage of CMOS is the ability to shape the charge to voltage conversion of the pixel through the use of the sub-threshold region of MOS transistors. It is a region known for its logarithmic response (Fig. 13.14). Another attractive feature is the multiple non-destructive reads to reduce read-noise. This improves signal-to-noise ratio (SNR) further by the square root of the number of reads (Chen *et al.*, 2012).

With the arrival of 180 nm CMOS processes for imaging, the read-noise, sensitivity and Q_{max} are such that CMOS ⅔-inch HDTV imagers became
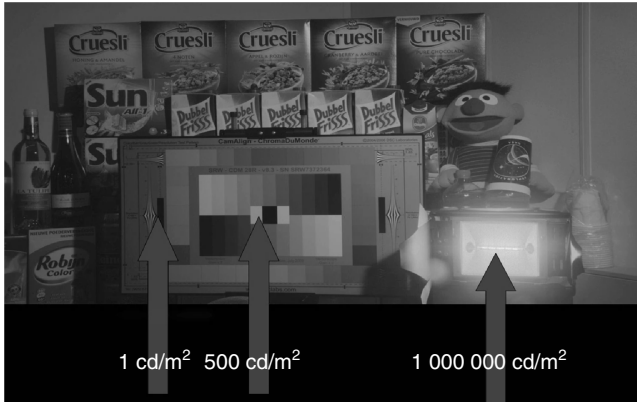
13.14 What a CMOS pixel can do with highlights (Jeroen Rotte, Grass Valley, Nederland BV).

viable for broadcast (Loose *et al.*, 2001; Kozlowski *et al.*, 2005; Centen *et al.*, 2007) and in related fields like industrial vision (Wang *et al.*, 2010; Meynants *et al.*, 2011). Grass Valley was the first to successfully introduce a ⅔-inch three-imager broadcast camera for use in live productions (LDX) with a global shutter cmos imager (Centen *et al.*, 2013).

13.8   Signal-to-noise ratio (SNR)

In the telecom world SNR means: 'measure the signal level, measure the rms-noise that goes with it and calculate the ratio between signal and rms-noise'. In broadcast the SNR is determined by the noise when there is no light falling on the imager (capped lens) and then to divide it by 700 mV (nominal output level).

Usually when the SNR is used, the SNR in luminance (Y) is meant, the REC. 709 luminance definition for HDTV is:

$$Y = 0.2126 * R + 0.7152 * G + 0.0722 * B \qquad [13.6]$$

The signal-to-noise in Y is mainly determined by the signal-to-noise in the green channel. In the 3200 K case a rule of thumb for the SNR in Y is:

$$SNY = SNG + 1.5 dB \qquad [13.7]$$

In SDTV the signal-to-noise in Y is stabilized at 60 dB–62 dB with *f*/number in the range of *f*/10–*f*/14. For HDTV, the SNY started at 54 dB with sensitivity at *f*/8 as the bare minimum for acceptable image quality. Due to improved sensitivity, lower read-noise and noise reducers the signal-to-noise moved up to 60 dB and sensitivity in the range of *f*/8-*f*/11.

The image area is surrounded with additional light sensitive and non-light sensitive pixels. The additional light sensitive pixels are used for run-in of the filters and the others for black reference purpose. The horizontal black pixels are often used to restore the black reference level. More often than not the restoration process generates low frequency noise known as line-noise. Due to the low frequency nature already, low noise levels are very visible. An alternative is the use of the black reference lines at the top or the bottom of the image allowing a longer integration time for reducing the noise. In both cases care must be taken to prevent stray light into the black reference pixels and also preventing charge spilling from neighboring light sensitive pixels into the black pixels. The illuminated pixels can be overexposed easily with 16 *f*-stops.

The classical parameters for noise optimization of the on-chip amplifier are the width, the length and the bias current of the source follower (Hynecek, 1984; Centen, 1991; Fasoli and Sampietro, 1996). Intuitively one thinks that high bandwidth and low noise are contradictory. Fortunately that is not the case when the oxide thickness of the MOS transistors can be chosen as an additional design parameter independent of the CCD performance determining properties.

Based on data from Wong (1996), a relation between minimum channel length versus oxide thickness can be plotted (Fig. 13.15). The straight line is given by $L_{\min} = 0.035*d_{\mathrm{ox}}$ (L in μm and d in nm).
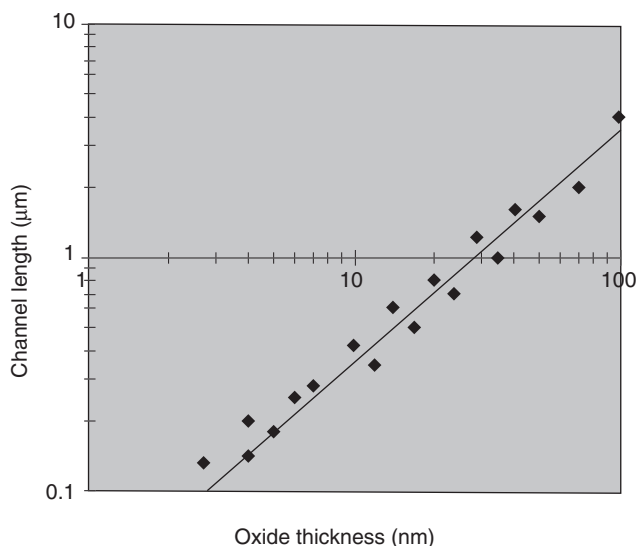
13.15 Minimum channel length versus oxide thickness.

From Centen (1991) we know that the noise electron density (NED), the noise power at a given frequency divided by the conversion gain of the on-chip amplifier) is written as:

$$\text{NED} = \frac{4kT}{g_\text{m}} * \left(\frac{C_\text{tot}}{q}\right)^2 \qquad\qquad [13.8]$$

and the bandwidth during charge sensing as:

$$F_\text{3dB} \equiv \frac{g_\text{m}}{2 * \pi * C_\text{load}} \frac{C_\text{fd}}{C_\text{tot}} \qquad\qquad [13.9]$$

with g_m the transconductance of the source follower, C_tot the sum of all the capacitance connected to the floating diffusion and C_fd known as the detection node capacitance. The capacitance at the source, loading the source follower, is C_load. Note that the 3 dB bandwidth of the source follower stage is reduced during charge sensing with the ratio C_fd/C_tot, this is due to the positive feedback from source to gate (floating diffusion).

In the idealized case, where the parasitic capacitances are neglected, the total capacitance reduces to $C_\text{tot} = \frac{2}{3} \times C_\text{g} + C_\text{fd}$, with $C_\text{g} = W.L.C_\text{ox}$ where the source follower has gate capacitance C_g, channel length L and channel width W. The capacitance per unit area is C_ox.

A first order approximation for the transconductance g_m, using for the drain-source current I_ds the product of channel width and a maximum current density $I_\text{ds} = W * J_\text{x}$, with maximum current density at $J_\text{x} = 10\,\text{A/m}$, and simplifying, is:

$$g_\text{m} = \sqrt{2 * \frac{W}{L} * \frac{\varepsilon}{d_\text{ox}} * u_\text{n} * I_\text{ds}} \approx \frac{C_\text{g}}{L} * \sqrt{\frac{d_\text{ox}}{L}} \approx \frac{C_\text{g}}{d_\text{ox}}. \qquad\qquad [13.10]$$

A minimum for the NED is at $C_\text{tot} = 2 \times C_\text{fd}$, or equivalently the gate capacitance of the source follower should be $C_\text{g} = 1.5 \times C_\text{fd}$.

Often the floating diffusion (C_fd) is determined by other design parameters that relate to the CCD channel and the CCD pixel. The total capacitance of the detection node C_tot is then fixed through the latter equation and so is the source follower gate capacitance. Referring back to the NED equation the NED decreases with increased transconductance (g_m). In the 3 dB bandwidth equation the capacitance ratio C_tot/C_fd is 2 and the bandwidth (F_3dB) increases with increased transconductance of the source follower. Halving oxide thickness d_ox and channel length L doubles the transconductance g_m, doubles the bandwidth and reduces the noise by 3 dB. The outlined approach was applied to the first stage of the three-stage on-chip amplifier of a triple-speed HDTV imager (Centen et al., 2009).
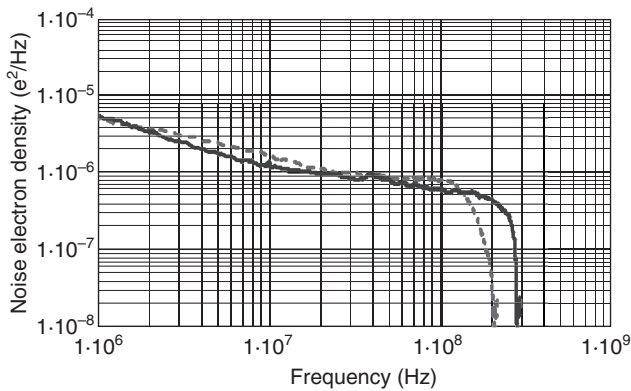
13.16 Static noise measurement of the on-chip amplifier. Dotted curve prior amplifier; solid curve the new amplifier.

Figure 13.16 shows the measured noise spectrum (solid line) of the new three-stage CCD on-chip amplifier in comparison with the prior one (dotted line). Parameter extraction revealed a $1/f$ corner frequency of 16 MHz with an exponent of 0.83 and a 3 dB bandwidth of 241 MHz. This includes bandwidth limitation of the external instrumentation amplifier. The new amplifier has a NED = $0.75e^2$/MHz at 37 MHz and at 112 MHz, NED = $0.59e^2$/MHz. After CDS the new amplifier delivered a noise level of 8e in 30 MHz bandwidth.

In Table 13.5 the state-of-the-art amplifier noise performance is given using reference to the reset frequency (Bosiers *et al.*, 2006). Noise levels in the range of 4.6 to 2 electrons are reported (Kozlowski *et al.*, 2005; Krymski *et al.*, 2003; Takahashi *et al.*, 2007) but only for CMOS imagers and only at a high gain setting of the column amplifiers. The high column gain reduces the saturation level at the same time. In prior papers sub-electron noise levels were reported for CCDs but always for bandwidths less than 1 MHz.

In interpreting the noise levels one must also take into account that in CMOS imagers each column has its own CDS circuit. The readout time of a CMOS pixel is on a μs time scale, while the CCD pixels are on an ns time scale and hence the noise is generated in a much larger bandwidth.

Given a CCD imager designed for single (1×) and triple speed (3×) operation. Compared with 1×, the sensitivity in 3× is reduced because the integration time is reduced from $\frac{1}{60}$ sec to $\frac{1}{180}$ sec and the noise will go up due to the increased video bandwidth from 30 MHz to 90 MHz. The SNR in 3× is reduced by 14.3 dB, of which 9.5 dB in sensitivity loss and 4.8 dB in noise increase. In a CMOS imager the sensitivity reduces the same 9.5 dB as in a CCD imager. Often the noise of the pixel is dominant and determined

Table 13.5 State-of-the-art amplifier noise performance

Type	Ref	Reset frequency	Amplifier type	Conversion gain	Noise after CDS/ $\sqrt{}$resetfrequency
FF-CCD	Burke *et al.* (1997)	100 kHz	Source follower	20 µV/e	6.3 e/$\sqrt{}$MHz
CMOS	Krymski *et al.* (2003)	50 kHz	Source follower +gain	60 µV/e	6.3 e/$\sqrt{}$MHz
FF-CCD	Draijer *et al.* (2005)	25 MHz	Source follower	40 µV/e	2.8 e/$\sqrt{}$MHz
CMOS	Kozlowski *et al.* (2005)	104 kHz	Source follower +gain	–	46 e/$\sqrt{}$MHz
CMOS	Yoshihara *et al.* (2006)	156 kHz	Source follower +gain	40 µV/e	17.7 e/$\sqrt{}$MHz
CMOS	Takahashi *et al.* (2007)	156 kHz	Source follower +gain	75 µV/e	11.6 e/$\sqrt{}$MHz
CMOS	Cho *et al.* (2007)	625 kHz	Source follower +gain	101 µV/e	10.4 e/$\sqrt{}$MHz
FT-CCD	This amplifier	111 MHz	Source follower	18 µV/e	1.3 e/$\sqrt{}$MHz

by the kTC-noise at the column (with C the column capacitor). The kTC is the same in 1× and in 3× operation. This is where one of the other advantages for CMOS imagers in broadcast cameras and cinematography lies: CMOS imagers can have intrinsic better SNRs

Recently 3 × 2.2 Mpixel CMOS imagers have been used in ⅔-inch broadcast cameras (Kozlowski *et al.*, 2005; Centen *et al.*, 2007); they are rolling shutter type. Imagers in broadcast need to be low noise, high speed, high sensitivity, high Q_{max} and low dark current. The imager here has an adjustable integration time and a global shutter (Centen *et al.*, 2013). The exposure control and global shuttering is realized with a 5T pixel.

In Fig. 13.17 several functional blocks are drawn. The vertical scanning is controlled either through an internal timing generator (FIT: flexible industrial token generator) that connects through a multiplexer (MUX) to the flexible vertical token registers (FVT). Alternatively, the timing is carried out externally and applied through the MUX to the FVTs. The scanning format is 1920 (H) × 1080 (V). For industrial vision the internal timing is used and the external timing is for broadcast camera performance. To prevent skew related timing problems the pixel array is driven from one side.
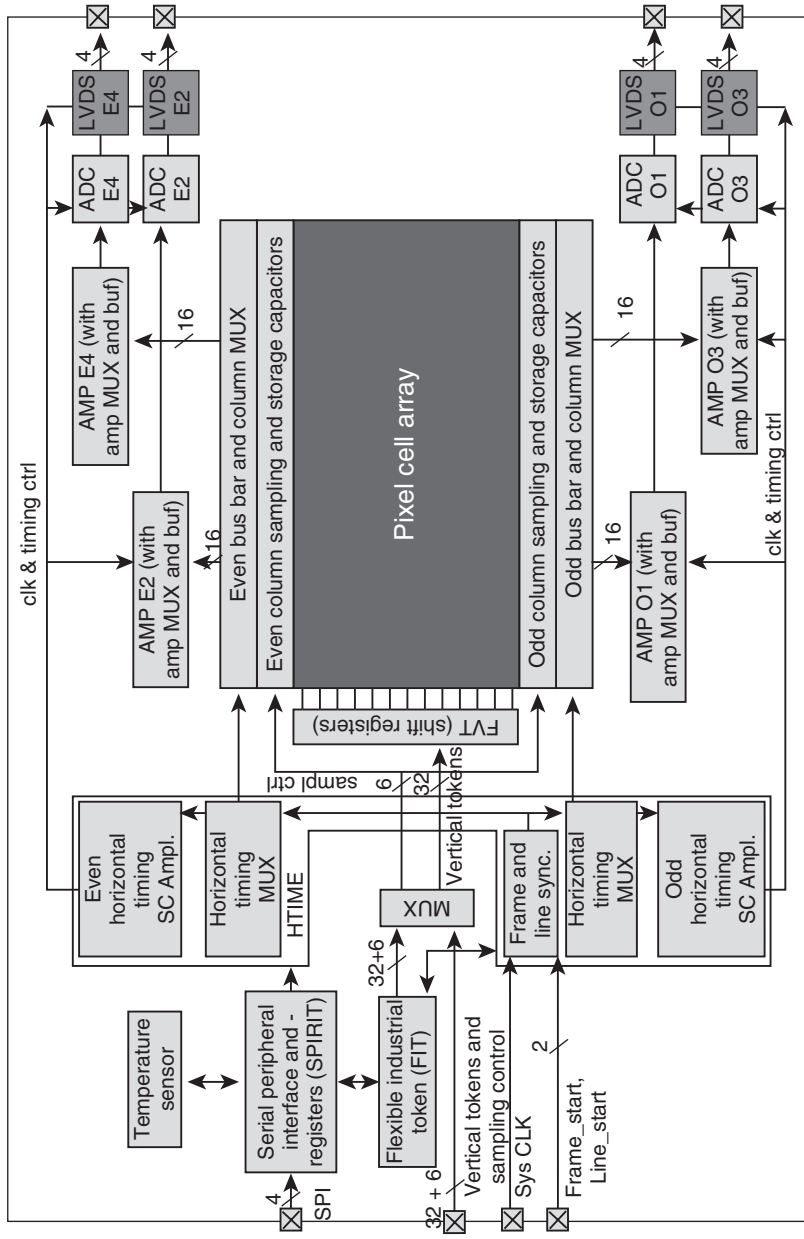
13.17 Block diagram CMOS imager (Steffen Lehr, Viimagic).

In rolling shutter mode, analog double sampling can reduce kTC noise, with 4T- or 5T-pixels. In global shutter mode one needs either an intermediate storage node (Sakakibara et al., 2012; Solhusvik et al., 2011) for kTC reduction, or external digital double sampling (Centen et al., 2007) and is applied to the imager. The input to digital double sampling (DDS) is in 1920 (H) × 1080 (V) at 120 frames/sec and the kTC noise reduced output is in 1920 (H) × 1080 (V) at 60 frames/sec. The 4e read-noise is achieved at a moderate analog gain of 2×. It allows, simultaneously, for a large pixel output swing and low noise in black, a requirement in broadcast. Normally low noise numbers require a high analog gain (Sakakibara et al., 2012; Solhusvik et al., 2011) reducing the maximum output swing that can be reached. The imager depicted in Fig. 13.17 is fabricated in 0.18 um 1P4M technology, with lightshield, and can support 1920 (H) × 1080 (V) up to 240 frames/sec, which is 120 frames/sec after DDS.

The 5T-pixels (Fig. 13.18) are placed in a rectangular lattice. The imager has a horizontal timing generator (HTIME) and supports region-of-interest (ROI) readout. The ROI window width is programmable in 66 blocks of 32 columns and the rows can be chosen arbitrarily. The vertical scanning of the image array can be controlled with a built-in token generator (FIT + MUX + FVT) and for demanding applications all vertical scanning tokens can be generated externally and applied to the shift registers through the multiplexer (MUX + FVT). To support high frame rates the odd columns are readout at the bottom side of the pixel array and the even columns at the top, simultaneously.
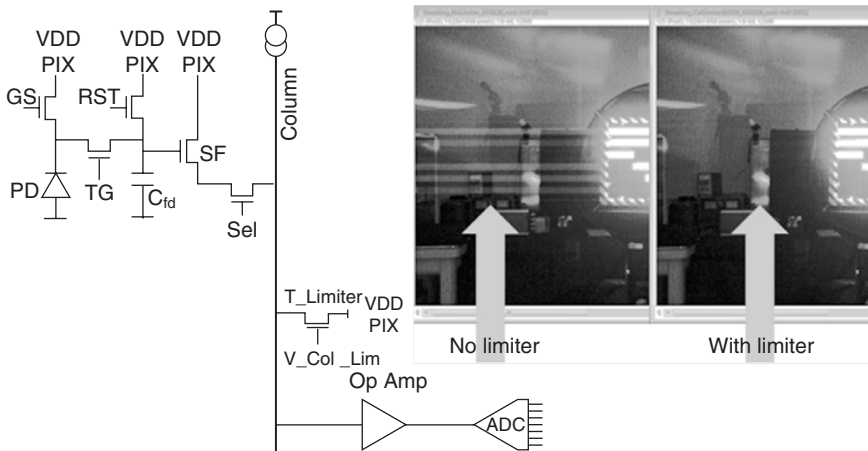
13.18 5T-pixel with limiter transistor; T_Limiter, at the column to reduce streaking (Jeroen Rotte, Grass Valley, Nederland BV).

After sampling at the column capacitor, the signal is multiplexed onto one of the 16 switch-capacitor-amplifiers (SCA). There are 4×16 SCAs and the output of each set of 16-SCAs is routed through an analog multiplexer to one of the ADC inputs. There are four ADCs and each one of them is connected to a four-lane low voltage differential signalling (LVDS) transmission.

The odd and even rows of the imager array can be addressed independently. The same holds for the odd and even columns. Reading the odd rows and odd columns on the bottom side and the even rows and even columns at the top side, the pixels are read in zig-zag (quincunx) fashion at 1920 × 1080Q480 which is 1920 × 1080Q240 after DDS, doubling the frame rate, without the need for a rhombic shaped pixel (Toyama *et al.*, 2011).

To reduce image artifacts (Purcell *et al.*, 2009) at highlights, each column has a n-type metal-oxide-semiconductor (NMOS) transistor, with an adjustable gate voltage. It reduces the voltage swing at the input of the analog gain stages and the input of the ADC, preventing overload. The additional (limiter) transistor is shown in Fig. 13.18 in relation to the 5T-pixel. Each column has a transistor, T_Limiter, that limits the negative going signal excursions through an adjustable voltage at the gate (V_Col_Lim). The source of T_Limiter is connected to the column and the drain to VDD_PIX = 3.3V. When the column voltage drops below the gate voltage minus the threshold, transistor T_Limiter takes over. The NMOS limiter prevents the current-source at the column entering the linear region and for the switched-capacitor-amplifiers to be overloaded. An overload condition often generates an image artifact know as LF-streaking (smearing) (Purcell *et al.*, 2009). The images show the difference with and without the limiter operation. The four analog to digital convertors each have a four lane LVDS output. In Fig. 13.19 an eye diagram is measured of an LVDS
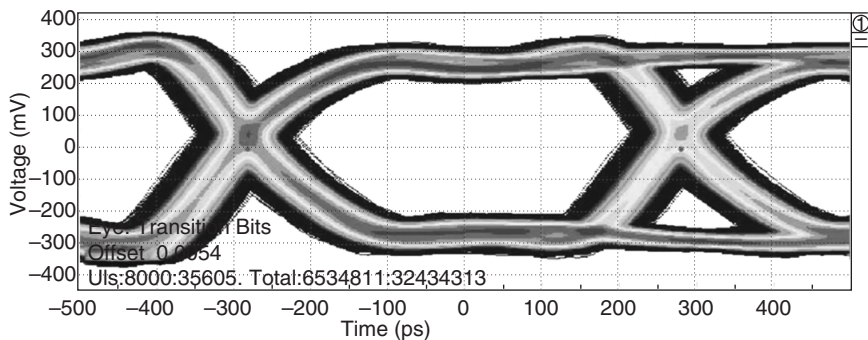
13.19 Eye diagram of an LVDS lane at 1.782 Gbps (Ruud van Ree, Grass Valley, Nederland BV).
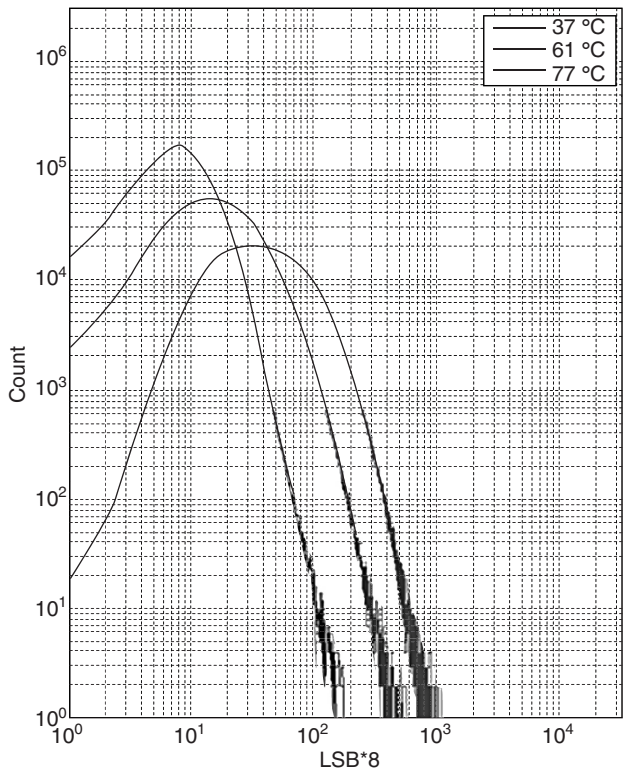
13.20 Histogram of the floating diffusion dark current at 37 °C, 61 °C and 77 °C (Jeroen Rotte, Grass Valley, Nederland BV).

lane at 1.782 Gbps. The maximum supported data rate of the 4 × 4 lanes is 28.5 Gbps.

A dark current histogram of the floating diffusion at 37 °C, 61 °C and 77 °C is shown in Fig. 13.20. The median value for the floating diffusion dark current is 114 e/sec at 37 °C, 282 e/sec at 61 °C and 720 e/sec at 77 °C, a doubling of the dark current of about 11 °C. The dark current of the photodiode doubles approximately with 6C: 2.5 e/sec at 37 °C, 38 e/sec at 61 °C and 230 e/sec at 77 °C. The dark current values could be reached through the use of a pinned photodiode (Teranishi *et al.*, 1982) and a pinned-surface underneath the transfer gate and global shutter gate.

A low dark current is important because the charge is stored at the floating diffusion and the shot noise of the dark current contributes to the read-noise too. To show the global shutter operation a frame grab was taken (Fig. 13.21) from an early test device. The scanning mode was 1920 × 1080p60 at an integration time of 1 ms. The ventilator has no rolling shutter distortion, only motion blur. The chip specification, measured at 1920 × 1080p60 after DDS (imager running at 1920 × 1080p120), is summarized in Table 13.6. The

13.21 Image after DDS at 1920 × 1080p60 in global shutter mode at 1 ms integration time (Jeroen Rotte, Grass Valley, Nederland BV).

Table 13.6 Some performance parameters of Xensium-FT

Process	CMOS 0.18 µm 1P4M with lightshield
Supply voltages	3.3 V and 1.8 V
Chip size	203 mm²
Number of light sensitive pixels and overscan pixels	1920 + 152 (H) × 1080 + 24 (V)
Pixel size	5 µm × 5 µm
Transistors per pixel	5T
Pixel fill factor	44% w/o micro-lens
Analog to digital conversion	4 ADCs
LVDS	4 * 4 lanes
Maximum data rate per LVDS lane	1.782 Gbps/lane
	28.5 Gbps all lanes
Maximum frame rate	120 frames/sec with DDS
	240 frames/sec with DDS quincunx
PRNU	<1%
Conversion gain FD	90–100 µV/e
Idark FD @ 61 °C	282 e/sec; 0.18 nA/cm²
Idark PD@ 61 °C	38 e/sec; 24 pA/cm²
Read noise global shutter mode, after DDS @27 °C	4 e at gain = 2x
Sensitivity in green (495 nm–573 nm)	2170 el.lux-sec/um²
Qmax	>15 kel
Power dissipation	1.6W @ 120 frames/sec, DDS
	1.1W @ 60 frames/sec, DDS

pixel fill factor without μ-lens is 44 and the conversion gain 95 μV/e. The read-noise after DDS is 4e at 27 °C with an analog gain of 2×. Power dissipation for the imager running at 1920 × 1080p120 is 1.1 W. The size of the active image array is 10.34 mm × 5.5 mm. Chip size is 203 mm², mounted in a customer defined ceramic package μPGA-185. With this CMOS imager the transition from CCD to CMOS was a fact.

13.9   Bit size, pixel count and other issues

In the debate about the number of bits needed to quantize the pixel signal the following observation can be made. In essence the pixel signal consists of photon generated electrons. They come in 0, 1, 2, 3, 4 electrons, up to Q_{max}. There are no half electrons! Therefore the number of bits needed to quantize the number of electrons (Fig. 13.22) in the pixel is

$$N = \frac{\log(Q_{max})}{\log(2)}$$ [13.11]

rounded to the next nearest integer. As an example, a pixel with a Q_{max} of 16 kel would need at maximum 14 bit. One ADC level then equals one electron equivalent and the quantization noise is 0.29 electron equivalent. Even the shot noise is quantized in integer numbers whereas the read-noise can have any value.

A minor but important issue is the discussion about the number of pixels. In broadcast the number of RGB-triplets are counted to define the resolution. That sets broadcast back by more than a factor of three, counting
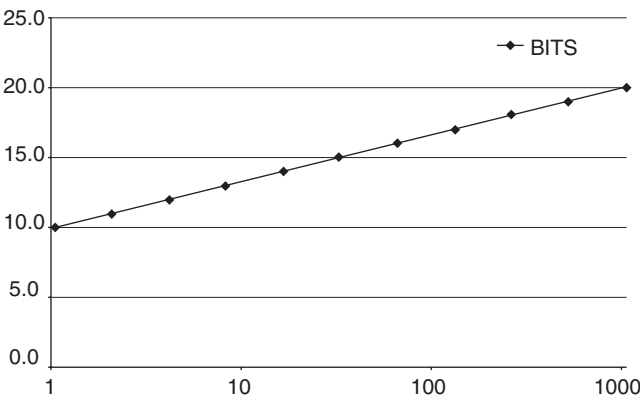
13.22 Number of bits versus Q_{max}. The units for the vertical axis are the number of bits. The horizontal axis has the number of electrons generated in the pixel in multiples of 1000.

Table 13.7 Comparison between single-imager and three-imager camera at equal resolution

Resolution comparison			
Aspect ratio: 16:9	Three-imager	Single imager	
HDTV	1280 × 720	2560 × 1440	Square pixels
HDTV	1920 × 1080	3840 × 2160 (aka 4k)	Square pixels
UHDTV 1	3840 × 2160	7680 × 4320 (aka 8k)	Square pixels
UHDTV 2	7680 × 4320	15360 × 8640	Square pixels

wise, as compared to the single imager world. The single imager world apply the well known Bayer patterned color filters or the color stripe filters. At every pixel-site one reads either R either G or, either B, they are all co-located. And the resolution is defined by adding all the pixels, counting the number of photo-sites.

An 8 Mpixel imager only has 2 kpixels of one color per line and 2 colors per line (Thorpe, 2013; Takayanagi *et al.*, 2009). For a single imager camera that is regarded as 4k scanning! But with that same token a full HDTV three-imager camera could be regarded as a 6k scanning device! This is because a three-imager full HDTV camera has 2 kpixels of one color per line and three colors per line. The fact that, in the single imager case, the color pixels are co-sited does not mean that one can resolve the amount of photo-sites. In the end it is just estimating what the red and the blue color information will be at the position where the green is and vice versa. It is the Nyquist reconstruction theorem that is calling the shots here.

The pixel shape of cinematography cameras can be rhombic on a quincunx lattice (Toyama *et al.*, 2011). This imager can be viewed as two lattices of $4k \times 2k$ each and the number of pixels is 17.67 Mpix. The approach of NHK is as it should be: in 4k one needs in a three-imager camera a total of $3 \times 4k \times 2k = 24$ Mpixels which for a single imager Bayer camera should $4k \times 4k \times 2k = 32$ Mpixel (Kitamura, 2011; Takayanagi *et al.*, 2009) and for 8k it is 3×32 Mpixel (Shimamoto *et al.*, 2012) or $4k \times 32k = 128$ Mpixel in Bayer. In Table 13.7 a resolution comparison is made between single imager and three-imager cameras.

13.10 Three-dimensional and ultra high-definition (UHD) television

13.10.1 Three-dimensional television

Films projected in the cinema can generate a good 3D user experience. Through extensive use of postproduction all the straining elements that

make you feel bad when viewing 3D can be removed. After postproduction one can offer a good 3D user experience. Live 3D is a completely different story. One does not have control over the scene and objects entering the scene at any distance and speed can cause bad 3D. Bad 3D can cause nausea and headaches, etc. Switching over from one scene to another must be done in a very controlled manner, again to prevent bad experiences. There is even discussion whether or not children should be allowed to view 3D because it might influence brain development in an adverse way. Alas, 3D has disappeared from the broadcast radar, at least for the time being.

13.10.2 UHD television (UHDTV) level 1 and level 2

UHDTV levels 1 and level 2 are the new topics. Level 1 is about 4k resolution (in RGB) and level 2 is 8k (in RGB). The common opinion is that more resolution only will not drive the market, it will not generate the wow factor. Introduction of UHDTV should be accompanied by a wider color gamut (e.g. Rec 2020), a higher frame rate (120 frames/sec) and higher dynamic range. Technically it translates into more bits per pixel and more pixels per second. Getting it off-chip and through the broadcast chain in a camera form factor that is workable for the cameraman, will be challenging. In the end it is all about the business model, i.e. the money.

The wider color gamut in general is not a technical problem. For Bayered imagers it means that the color cross talk which is inevitable between the co-sited red, green and blue pixels, must be limited. The cross talk determines the reduction of the colorspace. Rec 709 is of course not a problem, Rec 2020 could be.

Present day illuminators are shifting from tungsten type continuous spectrum illuminators to LED-type with several spectral peaks that cause metamerism and outside gamut response. A camera adjusted for one type of illumination can give a different response to another set of illuminators even though both scenes look the same (Salmon, 2012). The mapping between raw RGB from the imagers and Rec 709 colorimetry is governed by the matrix operation which is part of the video processing chain. Video values outside the color-triangle have negative values and are clipped. This can be regarded as implicit rendering. One could define the matrix against other primaries and obtain a wider color gamut. As is the case with the Rec 2020 for UHDTV.

With the advent of the CMOS-imager, whether in a single imager or three-imager camera, high frame rates, and also higher dynamic range and sufficient signal-to-noise are within reach. One has to realize, though, that 4k RGB means literally that in the single imager case it has to be 4k × 4k × 2k = 32 Mpixels (photo-sites) and 8k RGB will need four times this

amount of photo-sites, totaling to 128 Mpixels. In a three-imager camera those numbers would be 24 Mpixels and 96 Mpixels, respectively.

13.11 Conclusion

Performance-wise broadcast is a very demanding industry. Imagers in broadcast therefore have always been at the edge of what is technologically possible. In this chapter some aspects of that edge are described. The CCD brought video cameras to a high performance level. Only recently the cross over to global shuttered CMOS-imagers has begun. It will enable intrinsic better SNR in 1080p60 modes and beyond, and high dynamic range modes and other flexible readout mechanisms that are not possible with CCDs. It will be the enabler for UHDTV.

13.12 Sources of further information and advice

The Society of Motion Pictures Engineers (SMPTE), is a renowned standardization body in broadcast and cinematography (www.smpte.org). Another, European based body is the International Telecommunications Union (ITU – www.itu.int) whose broadcasters are represented by the European Broadcasters Union (www.ebu.ch).

Test charts

The following two companies deliver test charts for cameras: DSC-Labs (dsclabs.com) and Esser (www.image-engineering.de).

Books

Charles Poynton (2003) discusses many video aspects of cameras and Albert Theuwissen (1995) the imagers.

Standardization

ITU-R BT.709. Parameter values for the HDTV standards for production and international program exchange.
  SMPTE 274M standard describes a 1920×1080 matrix of light sensitive pixels.
  SMPTE 296M describes a 1280×720 matrix of light sensitive pixels.
  ITU-R BT.1769. 'Parameter values for an expanded hierarchy of LSDI image formats for production and international programme exchange,' 2007.

SMPTE 2036-1-2009, 'Ultra high definition television – image parameter values for program production,' 2009.

SMPTE ST 2048-1: 2011 '2048 × 1080 and 4096 × 2160 digital cinematography production image formats.

ITU-R Recommendation BT.2020, 'Parameter values for ultra-high definition television systems for production and international programme exchange,' International Telecommunications Union, Geneva, 2009.

13.13 References

Asano *et al.* (2011), *A ⅔-type 2.3-Mega Pixel IT-CCD for HD 1080p60*, Hokkaido 2011, IISW.

Blackwell (1946), Contrast thresholds of the human eye, *Journal of the Optical Society of America*, 36(11), Nov. 624–643.

Blankevoort *et al.* (1994), 'A high performance, full-bandwidth HDTV camera applying the first 2.2 Milion pixel frame transfer CCD sensor,' *SMPTE Journal*, May, 319–329.

Bosiers *et al.* (2002), 'Frame transfer CCDs for digital still cameras: concept, design, and evaluation,' *IEEE Transactions on Electron Devices*, 49(3), 377–386.

Bosiers *et al.* (2006), 'Technical challenges and recent progress in CCD imagers,' *Nuclear Instruments and Methods in Physics Research*, A 565, 148–156.

Burke *et al.* (1997), 'Soft X-ray CCD imagers for AXAF,' *IEE Trans Electron Devices*, 44(10), 1633–1642.

Centen (1991), 'CCD on-chip amplifiers: noise performance versus MOS transistor dimensions,' *IEEE Trans. Electron Devices*, ED-38(5), 1206–1216.

Centen *et al.* (1995), 'Dynamic pixel management for second generation aspect ratio switching,' *Montreux International Television Symposium Record*, 8–13 June, pp. 199–211.

Centen *et al.* (2001), 'A multi-format HDTV camera head,' *SMPTE Journal*, 110(8), 510–516.

Centen *et al.* (2007), 'A ⅔-inch CMOS image sensor for HDTV applications with multiple high-DR modes and flexible scanning,' *ISSCC Dig. Tech. Papers*, Feb, 512–513.

Centen *et al.* (2009), 'A ⅔-inch low noise HDTV FT CCD-imager for 1080i180, 1080p90 and 720p120 scanning at constant image diagonal,' *International Imager Workshop (IISW2009)*, June 2009, Bergen, Norway.

Centen *et al.* (2013), 'A 4e-noise ⅔-inch global shutter 1920 × 1080P120 CMOS-imager,' *International Image Sensor Workshop (IISW2013)*, June 2013, Snow-Bird, USA.

Chen *et al.* (2012), 'A 0.7erms-temporal-readout-Noise CMOS image sensor for low-light-level imaging,' *ISSCC2012*, pp. 384–385.

Cho *et al.* (2007), 'A $\frac{1}{2.5}$ inch 8.1Mpixel CMOS image sensor for digital cameras,' *ISSCC Dig. Tech. Papers*, 508–509.

Draijer *et al.* (2005), 'Third generation large area professional DSC CCD process and design technology,' *International Image Sensor Workshop*, Karuizawa Japan, Paper 64.

Drewery (1979), 'The zone plate as a television testpattern,' *SMPTE Journal Volume 88*, November.

Fasoli and Sampietro (1996), 'Criteria for setting the width of CCD front end transistor to reach minimum pixel noise,' *IEEE Trans. Electron Devices*, ED-43(7), 1073–1076.

Harada *et al.* (1992), 'A 2/3" FIT 2M pixel FIT-CCD HDTV image sensor,' *Proceedings of the ISSCC 1992*, pp. 170–171.

Honda *et al.* (2005), 'Development of ⅔"-type 1-mega pixel progressive scan CCD for HDTV capable of high frame rate of 96fps,' *Proc. of the IEEE AIS Workshop*, Kuruizawa, 177–180.

Hynecek (1984), 'Design and performance of a low-noise charge detection amplifier for VPCCD devices,' *IEEE Trans. Electron Devices*, ED-31(12), 1713–1719.

Janesick, Klaassen and Elliott (1987), 'Charge-coupled-device charge-collection efficiency and the photon transfer technique,' *Optical Engineering*, 26(10), 972–980.

Kitamura (2011), 'A 33 Mpixel, 120 fps CMOS image sensor for UDTV application with two-stage column-parallel cyclic ADCs,' *International Image Sensor Workshop (IISW2011)*, June, Hokkaido, Japan.

Kozlowski *et al.* (2005), 'A progressive 1920 × 1080 imaging system-on-chip for HDTV cameras,' *ISSCC Dig. Tech. Papers*, Feb, 358–359.

Krymski *et al.* (2003), 'A 2 e noise 1.3 Megapixel CMOS sensor,' *Proc. of the IEEE AIS Workshop*, June.

Lang (2001), *Colour and its Reproduction, Part 1: Colorimtery*. Muster-Schmidt Verlag, Mainz, Germany.

Loose *et al.* (2001), '⅔-inch CMOS imaging sensor for high definition television,' *IEEE Workshop on Charge-coupled Devices and Advanced Image Sensors*, 7–9, June.

Martinez-Verdu *et al.* (2002), 'Calculation of the color matching functions of digital cameras from their complete spectral sensitivities,' *Journal of Imaging Science and Technology*, 46, 15–58.

Meynants *et al.* (2011), 'Backside illuminated global shutter CMOS image sensors,' *International Image Sensor Workshop (IISW2011)*, June, Hokaido, Japan.

Moeland *et al.* (1998), 'A high speed sports action camera system,' *Technical Papers International Broadcast Convention*, Amsterdam, 11–15, September, pp. 86–92.

Morimoto *et al.* (1994), 'A ⅔-inch 2 M-pixel IT-CCD image sensor with individual p-wells for separate V-CCD and H-CCD formation,' *ISSCC. Dig. of Tech. Papers*, Feb., 222–223.

Nitta *et al.* (2006), 'High-speed digital double sampling with analog CDS on column parallel ADC architecture for low-noise active pixel sensor,' *ISSCC. Dig. of Tech. Papers*, Feb., 500–501.

Peek (1999), 'A low dark current double membrane poly-Si FT-technology for 2/3 inch 6M pixel CCD imagers,' *Tech. Dig. IEDM*, Dec., 871–874.

Poynton (2003), *Digital Video and HDTV Algorithms and Interfaces*, Morgan Kaufmann, San Francisco.

Purcell *et al.* (2009), 'Image artifacts caused by pixel bias cells in CMOS imagers targeted for mobile applications,' *International Image Sensor Workshop*, Norway, Invited Paper 73.

Sakakibara *et al.* (2012), 'An 83dB-dynamic-range single-exposure global-shutter CMOS imager sensor with in-pixel dual storage,' *ISSCC Dig. Tech. Papers*, Feb, 380–381.

Salmon (2012), 'The impact of LED lighting on colour rendtition in TV and other cameras,' *Image Sensors*, 21 March, London.

Shimamoto *et al.* (2012), '120 Hz-frame-rate SUPER Hi-VISION capture and display devices,' *SMPTE Conference*, October, Hollywood.

Solhusvik *et al.* (2011), 'A 1.2 MP ⅓-inch global shutter CMOS image sensor with pixel-wise automatic gain selection,' *International Image Sensor Workshop*, Japan, Paper 53.

Spitzer *et al.* (1996), 'Design and implementation of a 3-CCD state of the art, 750-line HDTV progressive scan broadcast camera,' *Proc. NAB*, Las Vegas, pp. 171–178.

Stoldt *et al.* (1996), 'CCD imagers for broadcast applications,' *Technical Digest IEDM96*, San Francisco, 8–11 December, pp. 899–902.

Takahashi *et al.* (2007), 'A $\frac{1}{2.7}$-inch low-noise CMOS image sensor for full HD camcorders,' *ISSCC Dig. Tech. Papers*, Feb, 510–511.

Takayanagi *et al.* (2009), 'Progress in 1.25-inch digital-output CMOS image sensor developments for UDTV application,' *International Image Sensor Workshop*, Norway.

Teranishi *et al.* (1982), 'No image lag photodiode structure in the interline CCD image sensor,' *Technical Digest IEDM*, 13–15 Dec, pp. 324–327.

Theuwissen *et al.* (1991), 'A 2.2 Mpixel FT-CCD imager, according to the Eureka HDTV-standard,' *IEDM Tech. Dig*. pp. 167–170.

Theuwissen (1995), *Solid-State Imaging with Charge-Coupled Devices*. Kluwer Academic Publishers, Dordrecht.

Thorpe *et al.* (1988), 'New advances in CCD imaging,' *SMPTE Journal*, May, pp. 378–387.

Thorpe (2000a), 'Technical aspects of the new world of multiformat DTV embodying progressive, interlaced, and segmented video frame format,' *SMPTE Journal*, September, 713–719.

Thorpe (2000b), 'Acquisition in 24p and 35mm film: a technical enquiry,' *142nd SMPTE Conference*, Pasadena, October.

Thorpe (2013), 'New optical and image sensor technologies for high-resolution digital cinematography,' *SMPTE Motion Imaging Journal*, March, 33–41.

T. Toyama *et al.* (2011), 'A 17.7Mpixel 120fps CMOS image sensor with 34.8Gb/s readout.' *ISSCC Dig. Tech. Papers*, Feb, 420–421.

Van Rooy *et al.* (2002), The VIPER filmstream camera: a marriage between film and video, *Proceedings of the Broadcast Asia Technical Conference and Exhibition*, Singapore, 18–22 June, pp. 1–10.

Wang, *et al.* (2010), 'A 2.2M CMOS image sensor for high speed machine vision applications,' *Proc. SPIE*, 7536.

Wong (1996), 'Technology and device scaling considerations for CMOS imagers,' *IEEE Trans. Electron Devices*, ED-43(12), 2131–2142.

Yoshihara *et al.* (2006), 'A $\frac{1}{1.8}$-inch 6.4MPixel 60frames/s CMOS image sensor with seamless mode change,' *ISSCC Dig. Tech. Papers*, 492–493.

High-performance silicon imagers and their applications in astrophysics, medicine and other fields

S. N I K Z A D, Jet Propulsion Laboratory, California Institute of Technology, USA

Abstract: This chapter briefly describes some elements of physics and materials that constitute the essential criteria for producing high-performance silicon imagers. In particular, the chapter discusses back illumination and passivation of surface states as a key approach for enabling high performance in silicon imagers, independent of their electronic readout structure. The chapter provides a brief review of silicon imaging arrays, applications in planetary exploration and astronomy, as well as a brief mention of specific cases in the commercial and medical fields.

Key words: back illumination, surface states, spectral range, quantum efficiency, stability, delta doping, band-structure engineering.

14.1 Introduction

Since the advent of digital imaging in 1969, many inventions and much progress have been made to produce high-performance silicon imagers with ultrahigh resolution, large format, low noise, and high sensitivity. Discoveries about the birth and evolution of our universe, detection of extra-solar planets, and new windows into our own solar system have been made possible using high-performance silicon imagers in Earth orbiting telescopes such as the Hubble Space Telescope and Kepler. In addition, instruments in missions visiting or orbiting our solar system planets such as Cassini (mission to Saturn), Galileo (mission to Jupiter), MER (Mars Exploration Rover), and Mars Science Laboratory (MSL) have demonstrated the power of this imaging technology. Discoveries made possible by these missions, in turn, have created new challenges and have placed more demands on the next generation of detector and imaging arrays in order to facilitate further discoveries.

Semiconductor imagers have become ubiquitous in our daily lives. The extent of their presence is so commonplace that we no longer imagine not having the immediate gratification of instant capture and observation of images. We use digital imagers to capture fleeting precious moments; we instantly record data such as faces of people we have just met, business cards, prices of items, coupons, boarding passes, recipes; and more broadly, we use these imagers to generally aide our memory. We use our digital imagers to help park our cars, to secure our homes, and communicate in real time with loved ones. Silicon digital imagers have been used to immortalize events, report uprisings and revolutions in the making, and to snap up-close and personal images of celebrities.

As scientists, we use solid-state imagers to record data and perform real-time and *in situ* analysis of experiments. We photograph the heavens, diagnose diseases, and detect signals at levels of single photons with temporal and spatial resolution that was not believed possible even a decade ago. The revolution of digital imaging technology and the replacement of film with digital imaging in consumer applications, medical field, and scientific instrument has been a disruptive technology development which has deeply affected the way we communicate with and record the world around us. Much of this progress has been made possible because of the unique advantages of silicon as a semiconductor material.

Starting from sand, nearly perfect single crystalline silicon wafers are produced that are in turn used in fabrication of silicon microelectronics, digital imagers, and solar devices. Whereas other semiconductor materials offer higher mobility (e.g., Ge, GaAs), direct bandgap (e.g., GaAs, InAs, InSb), and heterojunctions (e.g., $GaAs/Al_{1-x}Ga_xAs$), the success of silicon in the microelectronics and digital imaging field is due to the abundance and low cost of primary material, the exceptional quality of silicon that is produced, and the reliability, stability and repeatability of silicon oxide production.

Silicon oxide has proven to be an indispensible asset in fabrication of silicon devices. As George E. Smith in his Nobel lecture in 2009 said, 'In summary, CCDs were born in the $Si-SiO_2$ revolution and, because of their unique properties, created their own revolution in widespread imaging device applications.'

This chapter is organized as follows: first, principles of operation in solid-state imaging are discussed. This discussion is followed by a short description of various electronic readout schemes commonly used today in imaging devices. The chapter focuses next on scientific imagers including photon counting detectors and the role of back illumination in the design of high-performance silicon imagers. The remainder of the chapter briefly discusses applications in astronomy and planetary science, ending with a brief note on biomedical applications and an example in commercial, machine vision.

14.2   Solid-state imaging detectors: principles of operation

Light detection in a semiconductor and the measurement of a signal can be simply described as follows (Fig. 14.1): when a photon impinges on a semiconductor, the photon is absorbed, if it is not reflected off the semiconductor surface and if the photon energy is equal or greater than the detector's semiconductor bandgap. The absorption of the photon creates an electron-hole pair. The pair is separated through the field produced by the doping profile of the detector. The electron is collected as a current or voltage depending on the architecture or readout scheme of the imager design. The measurement of current or the collection of charge is accomplished by the depletion field sweeping the photoelectrons to the region where they are converted, for example, to current as in photodiodes, or to voltage as in charge-coupled devices (CCDs) or complementary metal-oxide-semiconductor (CMOS) detectors – e.g., the floating gate diffusion contact. Devices are typically fabricated in low-doped, high quality single crystalline silicon that is epitaxially grown on a float zone silicon substrate. The silicon substrate typically has several orders of magnitude higher doping level than the epitaxially-grown layer (or epilayer). All photo-carrier generation, transport, and collection occurs in this epilayer. Therefore the quality of this epilayer, number of defects in the layer, its doping level, and its doping uniformity are crucial to the device performance. In fact, material quality plays a fundamental role in all aspects of device performance.

Metrics used to measure performance of an imager greatly depend on the application for which the imager is used. For scientific imaging, quantum efficiency (QE), signal-to-noise (S/N) ratio, the stability of the signal as a function of illumination and environmental parameters, resolution, and the
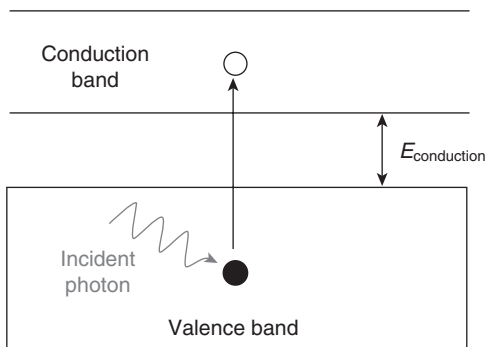
14.1 Principle of photon absorption and carrier generation in a semiconductor.

uniformity of the response are among the most important parameters of imager performance.

QE is the measure of detector's ability to 'see' or detect incident photons and can be defined as the ratio of detected photons to incident photons. In practice, what is measured or detected are electrons, so the QE is often defined as the number of electrons collected, divided by the number of incident photons at a given wavelength. As explained below, using this ratio could lead to erroneously high QE. Many phenomena affect the QE but they can be grouped into four major factors. Care must be taken to ensure accurate correlation between the measured quantity and true quantum efficiency:

$$QE = \left(\frac{\text{absorbed photons}}{\text{incident photons}}\right)\left(\frac{\text{photons generating e}-\text{h pairs}}{\text{absorbed photons}}\right)$$
$$\left(\frac{\text{photoelectrons generated}}{\text{photons generating e}-\text{h pairs}}\right)\left(\frac{\text{collected electrons}}{\text{photoelectrons generated}}\right)$$
$$\text{or} \quad QE\,(\text{measured}) = T \times P \times QY \times IQE$$

where T is the transmittance of incident photons into the silicon. T accounts for loss of photons due to reflection from the surface as well as the losses of photons due to absorption in surface oxides or applied coatings where the photon will not generate useful photoelectrons. In silicon, in most cases, T is approximated as $1\text{-}R$ where R is the surface reflectivity and it is a function of photon energy. P is the probability of an absorbed photon generating an electron-hole (e-h) pair, and in the case of most silicon detectors, this number is essentially unity.

For photons that generate electron-hole pairs, the quantum yield (QY) is the number of photoelectrons generated per photon and it can be defined as:

$$QY = E_{\text{photon}}/E_{\text{e-h}}$$

where E_{photon} is the energy of the incident photon and $E_{\text{e-h}}$ is the average energy required by the electron-hole (e-h) pair. QY is wavelength dependent as well as material dependent. A *very* rough estimate of $E_{\text{e-h}}$ in silicon is 3.65 eV where one electron-hole pair is produced for each 3.65 eV of photon energy (Wilkinson *et al.*1982). It should be emphasized that this is somewhat of an oversimplification as $E_{\text{e-h}}$ is dependent on the photon energy. Furthermore, $E_{\text{e-h}}$ has been shown in silicon to deviate from this value for lower energy (sub keV) electrons as well as charged and neutral particles (Funsten *et al.* 2004, Nikzad *et al.* 1999, Nikzad *et al.* 2006). For silicon detectors, QY is generally unity for visible photons and it is greater than unity for higher-energy photons. For ultraviolet photons, X rays and higher energy particles, it is best to directly measure the QY for each wavelength (Nikzad *et al.* 2012).
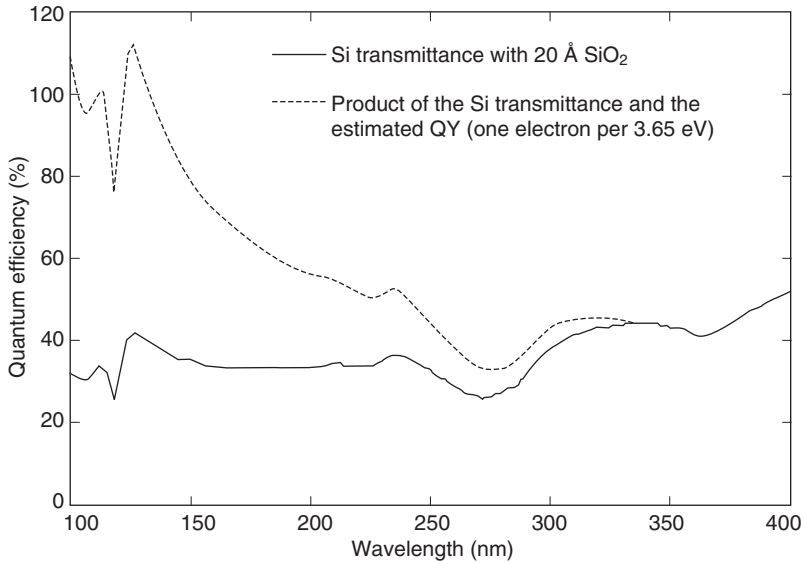
14.2 Silicon transmittance including the effect of silicon oxide is shown in solid line. The dashed line is the silicon transmittance (with oxide) including the effect of quantum yield. At higher photon energies ($\lambda < 200$ nm), more than one electron-hole pairs are produced for each photon. For regions of the spectrum where the QY is greater than unity, care must be taken for an accurate measure of QE (detected photons to the incident photons).

The final term, internal quantum efficiency (IQE), is the collection efficiency of the photoelectrons. For $P=1$ and for a QY of unity where one photoelectron is produced for every absorbed photon, and internal QE could be calculated as the ratio of the measured QE to the transmittance T. In the visible and near UV the transmittance is determined almost entirely by reflection from the bare silicon surface. The internal QE can be affected by trapping or recombination of the photoelectrons, and in a back illuminated devices, by trapping in the backside potential well of the silicon array.

Figure 14.2 shows a calculation of the above effects. The solid line is the silicon transmittance, including reflection and absorption in the oxide, and corresponds to the reflection-limited quantum efficiency (or 100% internal QE) for a QY of unity. The dashed line is the silicon transmittance (with oxide) including the effect of quantum yield. For regions of the spectrum where the QY is greater than unity, a measured ratio of greater than 100% is possible. It is important to note that true QE will not be above 100% and it is crucial to measure the QY and divide the measured ratio by QY in order to determine external QE (Nikzad *et al.* 2012).

QE of silicon detectors is not constant as a function of wavelength. Silicon detectors naturally have the highest sensitivity in the visible wavelength
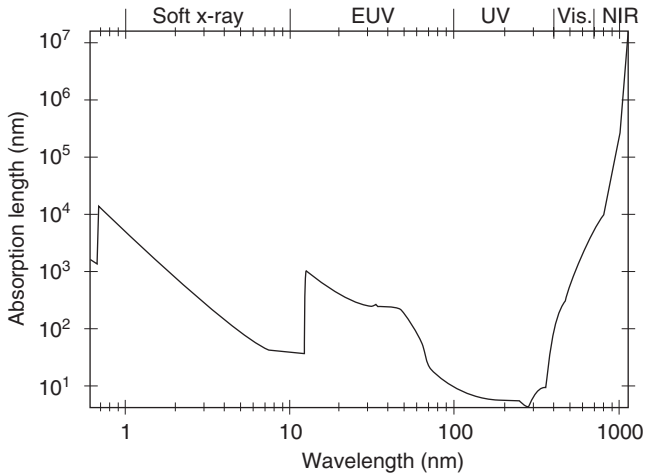
14.3 Photon absorption length in silicon as a function of photon wavelength. In the entire UV and blue section of the spectrum, the absorption is below 1 nm whereas at longer wavelengths, the absorption length is tens or hundreds of microns into silicon.

range. As we will discuss later in this chapter, silicon detectors also can be modified for high sensitivity in the ultraviolet, far ultraviolet, and even soft X rays through back illumination.

At the other end of the spectrum, the spectral response of silicon imagers typically decreases rapidly beyond 700 nm range and drops to a few percent around 1000 nm. This is again a matter of absorption of photons and the inefficient absorption of these longer wavelength photons in the indirect bandgap. Figure 14.3 shows absorption length of photons in silicon as a function of wavelength. To allow efficient detection of $\lambda \sim 1$ µm light, a thickness of at least 100 micron is required. Holland and team, have developed fully depleted CCDs in several micron-thick ultrahigh purity silicon in a p-channel design (Holland *et al.* 2003). Note that silicon's indirect bandgap, and the consequent long absorption length of visible and near infrared photons in this material, requires that high quality, thicker layers of silicon be used for detector fabrication. The fact that imaging arrays can be produced in this indirect bandgap material, is due (as mentioned earlier in the introduction) to our ability to produce several micron-thick to several hundreds of micron-thick high quality layers of silicon.

When back illuminated, these fully-depleted CCDs have a broad spectral response. Especially when combined with band-structure engineering through molecular beam epitaxy (MBE), i.e., delta doping, these devices have exhibited high quantum efficiency over a wide spectral range (near UV to near IR) (Blacksberg *et al.* 2005, 2008). An added advantage of the design by Holland and team is the higher radiation tolerance due to the

p-channel design and the lower probability of vacancies in this design. The higher radiation tolerance of these detectors has been well documented in the literature (Bebek 2002).

14.3   Scientific imaging detectors

The requirements placed on detectors are defined and reduced from the scientific objective by the end user (observer). These requirements often assign perfect performance parameters for the detector such as zero noise, 100% photon detection probability, 100% fill factor, 100% quantum efficiency, and complete tolerance and robustness to the radiation. Increased spectral range sensitivity over a broad range is desired in some applications while in other applications high sensitivity is required in a spectral band along with no out-of-band sensitivity.

Although a perfect combination of the attributes described above does not exist in any single detector, certain readout designs and architectures are more suitable for a given application. Here we focus on a common feature that could improve many of the detector performance parameters for scientific and high-performance applications.

To achieve the highest performance in silicon imaging arrays, these detectors need to be back illuminated. Typically images are illuminated through the VLSI –fabricated side of the device. This VLSI-fabricated pixel side is conventionally called the front surface and when illuminated at this surface the device is referred to as front-illuminated. Layers of material in the pixel fabrication interfere with efficient absorption of the photons in the sensing area. For this reason, by flipping the chip and using the 'back surface' as the illumination surface, photons only interact with silicon. In this 'back illuminated' configuration, the photoelectrons can be directed into corresponding pixels in the front surface under the fields that extend from the front surface.

Back illumination allows expansion of spectral range sensitivity beyond visible and into ultraviolet, it allows 100% fill factor, and it enables high QE. This statement was made over a decade and half ago in a NASA Detector workshop (from X-ray to X Band, STSCI, Baltimore, MD, 2000) indicating the application of back illumination to all silicon imagers and not just CCDs, at a time when back illumination seemed mostly the concern of scientific silicon imaging detectors (Nikzad *et al.* 2000). At the time, back illuminated CCDs had been used in several scientific instruments and they were being planned for the Kepler mission to detect extra-solar planets. Over a decade later, it has become clear that for other silicon imagers, e.g., CMOS imagers, back illumination is critical for achieving high performance. This is true for non-scientific applications as well. Demands for higher resolution in consumer cameras, dictating larger numbers of pixels while keeping the

camera small, requires small pixels. On the other hand, increasing demand for on-chip processing means that in a CMOS imager more function and therefore more electronics circuitry is packed into the pixel. This leaves only a small area for photon sensing. By using back illumination, the part of the pixel's area that is devoted to electronics is restored to photon sensing, allowing near 100% fill factor to be achieved in a CMOS pixel. The consumer market push for high resolution and the scientific imaging field's requirement for high QE and broadband response has spurred on the fast development of back illuminated, high-performance imagers. Before discussing back illumination as a common process for all scientific imagers, we briefly outline the most common readout designs in silicon imaging.

14.4   Readout structures

As described above, absorption of photons in a semiconductor leads to production of electron-hole pairs that are then separated, collected, and read out through various schemes. The most commonly used readout designs are CCDs and CMOS imaging arrays. Excellent review papers and books describing fundamentals of design and operation of CCDs and CMOS arrays can be found (for example, see Janesick and Putman 2003, Theuwissen 1995). Here we will mention only the elementary anatomy of these readout schemes. In a CCD, through gate voltages and clocking, potential wells are formed and collapsed that allow the transfer of charge serially through the array and finally through a serial register and a readout amplifier. In a CMOS array, each pixel includes a readout amplifier allowing a parallel readout, in contrast to the CCD's serial readout. Figure 14.4 schematically illustrates these readout approaches.

In addition to these monolithic silicon imager designs, hybrid detectors offer another option for silicon imaging structures. PIN diode arrays can be hybridized with CMOS multiplexers using indium bump bonding or other forms of hybridization. These detectors have been developed by various groups and are commercially available (Nikzad *et al.* 2000, Bai *et al.* 2000). The hybrid structure allows for the independent optimization of detector and readout. Hybridization to detector fabricated in other semiconductors clearly allows expansion to the detection of a broader spectral range. Hybridization, however, also adds to the complexity and cost of fabrication.

14.5   Photon counting detectors

Scientific applications often drive the demands on detector performance. For example, in ultraviolet astrophysics, the objects of study often have such faint signal that photon counting is desirable. It is therefore appropriate to briefly discuss single photon detection. In order to count photons in the

CCD pixels

Serial resister

Readout amplifier

Row decoder

Pixel array

Analog output

Column processing circuitry

Column decoder
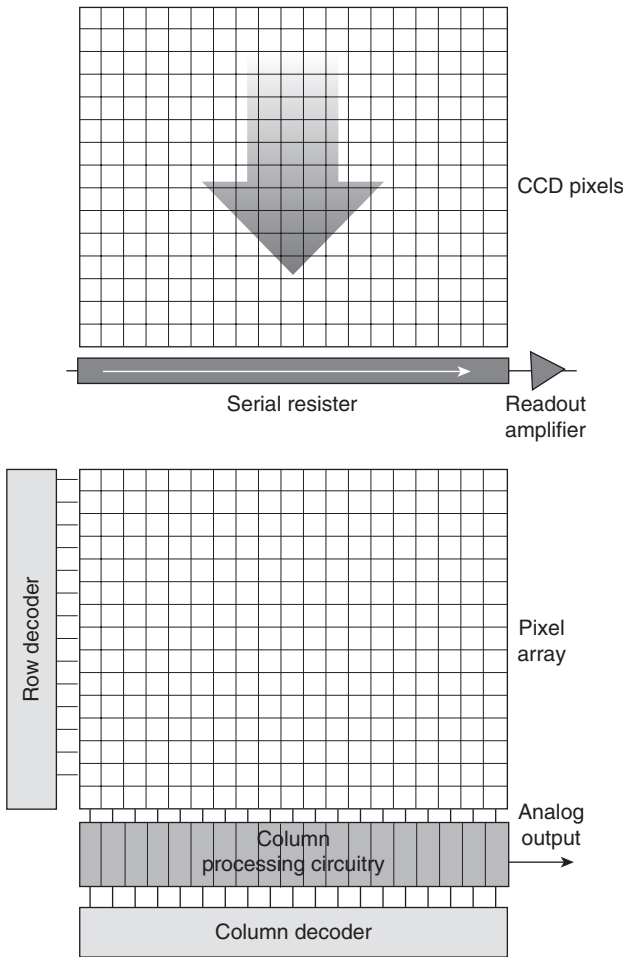
14.4 The conceptual structure of CCD (top) with serial readout and single output amplifier and CMOS imagers with parallel readout scheme (below).

presence of several to many electron background noise, gain is necessary to increase the number of electrons per photon. Detectors that possess gain are divided into two general categories: image tube-based technologies and all-solid-state imagers with avalanche gain.

Photomultiplier tubes, microchannel plates (MCPs), and electron-bombarded arrays including electron bombarded CCDs and electron bombarded CMOS are in the image-tube-based class of devices. In this general design, photons are absorbed in a photocathode; photoelectrons produced in the photocathode are then accelerated into vacuum under an electric field of typically several thousand volts. These electrons undergo

amplification by multiplication and are then detected by different schemes that are described below. In almost all of these various schemes, there is still a role for digital imagers as the final signal is often measured either by direction-detection of electrons using a silicon detector array, or by conversion of the electrons to photons by impinging on a phosphor screen, with a silicon detector arrays used to detect these converted photons.

14.5.1 Electron-bombarded (EB) arrays

As described above, photons that possess enough energy are absorbed in the photocathode material, producing electron-hole pairs. The photoelectrons escape the surface of the photocathode and are accelerated into the vacuum under the accelerating voltage V into the electron detector. In the case of the electron-bombarded (EB) CCDs or EBCMOS, the electron detector is a back illuminated CCD or CMOS array. The applied acceleration voltage is partially dependent on the low energy electron detection threshold of the back illuminated array. In recent years, back illuminated devices have been developed that have shown sensitivity to electrons with energies in the < 1000 eV range which could enable very compact EB arrays. In recent years, more compact designs of EB arrays are being developed that take advantage of better performance of back illuminated silicon detectors. Some of these designs have enabled lower operating voltages (Funsten *et al.* 2004, Nikzad *et al.* 1997].

Because the CCD or CMOS detectors are also sensitive to visible photons that pass through the photocathode without being absorbed, the electron detector is often placed off axis and photoelectrons are deflected by a magnetic field to strike the detector (Fig. 14.5(a)). Line of sight photocathode and electron detectors with compact design and without the use of a magnet have been devised. Rejecting out of band photons can be obtained using thresholding for the signal produced by these photons. An example of an electron-bombarded array in proximity focus or line of site configuration is shown in Fig. 14.5(b) and (c).

A few words are warranted here regarding photocathodes as they determine the efficiency of the entire detector as well as the reliability and manufacturability of the device. Historically, photocathode activation or the ability to allow escape of electrons from the surface of the material is achieved by cesiation, the deposition of an ultrathin layer of cesium on the surface of the photocathode material. Cesium's low work function creates a surface dipole that produces negative electron affinity (NEA) by lowering the vacuum level below the edge of the conduction band. Because Cs is highly reactive, the device must be protected from air exposure throughout the fabrication and operation. Using Cs therefore adds complexity and reliability issues to this class of devices. For stability and reliability purposes,
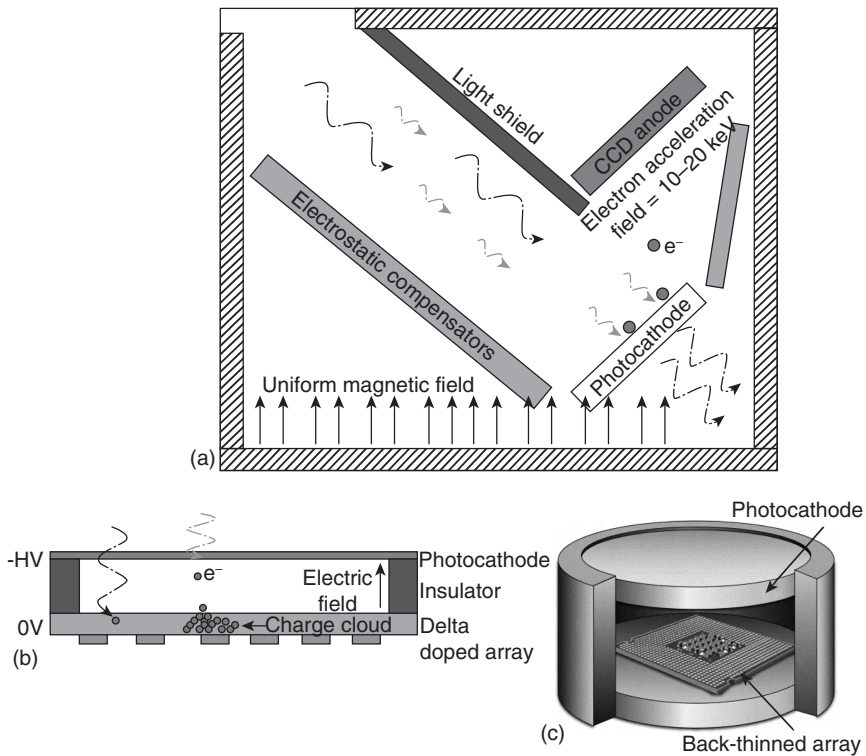
14.5 (a) Schematic of the operation of a conventional electron bombarded array. In this configuration, the photoelectrons generated in the photocathode from the in-band photons are deflected and focused on the CCD anode using a magnetic field (b and c). Cross-sectional schematic (b) illustration and cutaway illustration (c) of an advanced electron bombarded array with lower operating voltage, compact packaging, and no magnet requirement.

it is desirable to achieve NEA without use of the reactive material. Work is underway to produce NEA in gallium nitride and alloys (such as GaN and $Al_{1-x}Ga_xN$)-based photocathodes through surface band-structure engineering for UV applications (Tripathi *et al.* 2010a, b).

14.5.2 Solid state detectors with gain

In a semiconductor, gain is achieved by inducing avalanche multiplication. Various device designs achieve this using different formats and different readout schemes. Three of these designs are briefly described below.

To increase the S/N ratio in a semiconductor imager, efforts are made to both decrease the noise and amplify the signal. A gain of several tens of

thousands in the signal can be achieved by inducing avalanche multiplication in the semiconductor. Avalanche photodiode arrays (APDs) are variations of PIN diode arrays, often with absorption and multiplication separated to increase signal while maintaining lower noise, that are operated in the avalanche Geiger or linear mode. The first avalanche photodiode arrays were fabricated in silicon; but because the major motivation for requiring the high S/N ratio also required sensitivity in one micron and larger wavelengths, the material was abandoned in favor of other semiconductors such as InGaAs and mercury cadmium telluride. Since that time, silicon APD popularity has been renewed partially due to the advancement in the imaging technology with gain and partially due to the need for high S/N ratio imagers in the visible and ultraviolet spectral range. Here we briefly review the broad classifications of silicon imagers that possess gain along with their specific advantages and applications.

CCDs as integrating detectors with gain: electron multiplied CCDs (EMCCDs)

In the last 15 years an innovation in CCDs has made it possible to achieve a high S/N ratio in the popular CCD format. Benefiting from the uniformity and linearity CCDs as well as their low noise readout amplifier, gain is achieved in this scheme by adding an additional register to the final serial register of the device prior to the readout amplifier. Called the multiplication register, high voltage in the range of 20–40 volts is applied for the each stage of the electron transfer in this register. This increase in transfer voltage produces a small avalanche gain in each stage of the transfer, causing a final gain of 1000 or more while maintaining low noise at the readout amplifier. This approach allows a significant increase in the S/N ratio thus allowing the detection of faint signals. Figure 14.6 schematically shows the EMCCD structure (Jerram *et al.* 2001, Hynecek 2001).

Avalanche photodiodes

Avalanche photodiode or array operations can be divided into several stages depending on the magnitude of the reverse bias voltage. For low applied voltages, a small photoresponse will be detected. As the reverse bias voltage is increased, an output current is detected which is proportional to the incident optical power with a gain of unity. At still higher voltages, electron avalanche produces multiplication. In the avalanche mode, there are two regions of operation: in the normal or linear region, the output current is proportional to the incident optical power. At voltages near avalanche breakdown, incident photons can move the APD into breakdown, producing a flood of electrons with very high effective gain. Geiger-mode
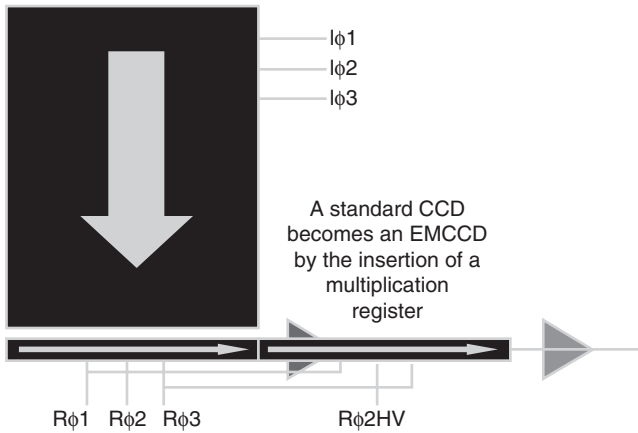
14.6 Electron multiplied CCD (EMCCD) concept shown here in the e2v Inc's L3CCD. The multiplication or gain register added as an additional serial register, creates avalanche in each stage of the transfer producing gain at the end prior to readout through the low noise output amplifier. Courtesy of e2v Inc.

measurements employ quenching electronics that turn photon detection into a series of digital pulses. Pulse counting allows single-photon detection without the usual noise contributions of analog measurements. In the linear gain region as well as the sub-Geiger region, single photon counting is either not possible due to external electronic thermal noise or not recommended to high gain fluctuations with small temperature (Farr 2009).

Single photon avalanche diodes (SPADs)

During the last decade, due to the progression of the ability to fabricate smaller and smaller features, it has become possible to include gain in the pixel of a CMOS imager. The single photon avalanche diode (or SPAD) has progressed as shown in Fig. 14.7 (Charbon 2013).

As described above, the reverse-biased pn junction operating in the Geiger mode produces a pulse in order to detect single photons. A SPAD can be implemented in CMOS processes in which a CMOS-compatible digital signal is generated when a photon is detected. Figure 14.8 shows a schematic of a SPAD structure. Because standard CMOS processes are used in SPAD fabrication, this device can be produced at relatively low cost and operated at relatively low power (Niclass *et al.* 2008, Charbon 2011, Mandai and Charbon 2012). With high timing resolutions of 100 ps per pixel achieved on relatively large arrays, SPADs are especially useful for time-resolved measurements (Charbon 2014).
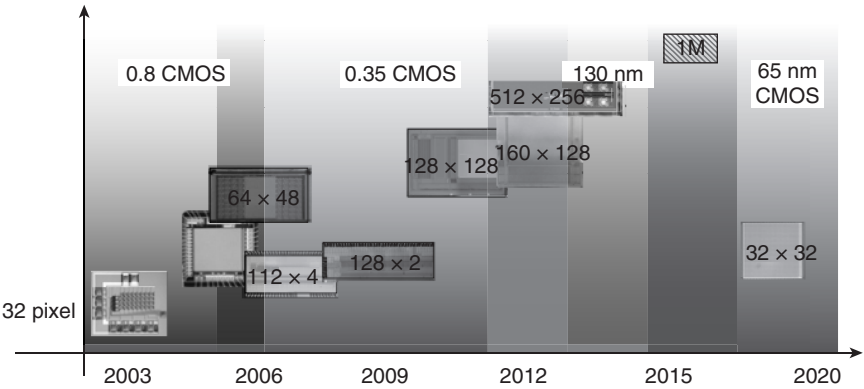
14.7 Evolution of SPADs with the subtitle of SPADs run faster than Moore's law. Courtesy of Professor Edoardo Charbon.
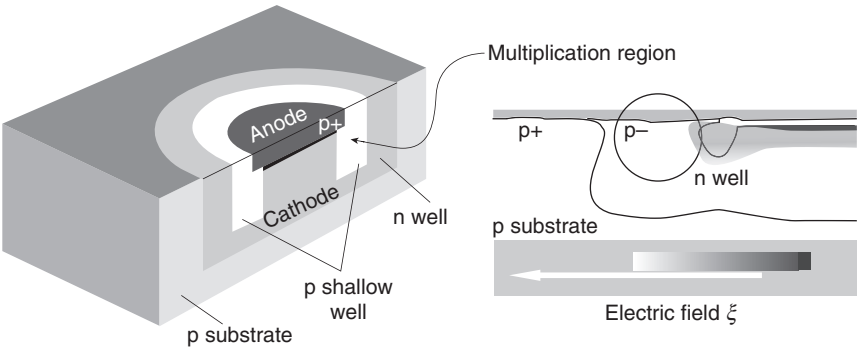
14.8 Schematic of a SPAD. Courtesy of Professor Edoardo Charbon.

As with other silicon detectors, SPADs can be operated front or back illuminated. Most SPADs are operated in front illumination configuration; however, these devices would particularly benefit from back illumination due to the high fraction of pixels devoted to circuitry. As with all other silicon detectors, when operated in the back illuminated mode, a surface processing method such as delta doping (described below) will be necessary for passivation as well as in enabling detection of high energy photons. This could significantly improve QE while potentially improving the noise performance. Dark counts may be further improved by reducing tunneling

by way of appropriately doping the multiplication regions and improving the guard rings used in SPADs for premature edge breakdown mitigation (Charbon2011).

14.6   High-performance imaging through back illumination

In the context of our discussion here, high-performance metrics are defined by wide spectral range in one detector (e.g., UV, visible, NIR in silicon), high sensitivity (or high quantum efficiency), high fill factor, stability of response, uniformity of response, and surface generated dark current. These are among the major detector metrics that can be affected by back illumination regardless of the detector readout structure (e.g., CCD, CMOS, or APD). Using these criteria, we can see how back illumination will help the performance of the detector or imaging array. By illuminating from the back surface and avoiding the VLSI fabrication surface, one avoids the absorbing layers of metals and poly silicon that are the necessary for the pixel structure. This allows more efficient photon absorption and subsequently higher quantum efficiency and potentially allows sensitivity to shorter wavelength photons that are absorbed in the first few nanometers in silicon. Furthermore, as stated before, because in CMOS imaging arrays much of the pixel area is devoted to electronic signal processing and readout, the fraction of the pixel area devoted for sensing area or the fill factor is reduced. By illuminating from the back surface, the real estate lost to electronics will be restored to photon sensing therefore allowing for 100% fill factor and higher photon detection efficiency.

14.6.1 Processing steps of back illuminated devices

For back illumination operation, all the excess silicon, i.e. the substrate, will be removed down to the epitaxial layer where all the device operation occurs. Thinning or substrate removal provides a fresh silicon surface with many dangling bonds that will immediately get saturated by forming a native silicon oxide. This new surface and its native oxide have an interesting electronic structure that is determined by the quality of the oxide as well as the silicon crystal orientation. In p-type materials, the silicon-silicon oxide interface has a high density of states, which attracts positive charge. This positive charge affects the nearby field and causes the edge of the conduction band to bend downwards forming a backside potential well, which can trap photoelectrons. An outline of various techniques for removing this trapping potential, and more details of band-structure engineering using MBE, for an ideal solution to this problem, are discussed in the next section.

14.6.2 Band-structure engineering through atomic level control: MBE for delta doping and superlattice doping

It is a well-documented phenomenon that the interface states of silicon and silicon oxide trap charge which unfavorably affects the near surface band structure of the imaging array (Bardeen 1947, Grunthaner *et al.* 1988, Hoenk *et al.* 1992, Nikzad *et al.* 1994a,b, 2012). This unfavorable band bending is independent of the frontside readout circuitry as schematically illustrated in Fig. 14.9. Several solutions have been devised since the invention of back illuminated CCDs, such as the UV flood, the Schottky barrier 'flash gate' using platinum or iridium, chemisorption charging, ion implantation, and MBE-processed band-structure engineering (Fig. 14.10). These back surface passivation solutions fit in three general categories:

1. Chemical methods of attracting charge with opposite polarity to the interface trapped charge such as UV flood and chemisorption (Williams *et al.* 1997, Lesser and Iyer 1998).
2. Physical methods of placing charge near the silicon-silicon oxide interface which includes ion implantation and post anneal (Janesick *et al.* 1989) and epitaxial-based band-structure engineering (Hoenk *et al.* 1992).
3. Field-effect techniques that induce variation in the near-surface fields by applying a bias to a deposited thin metal electrode (Janesick 2001).
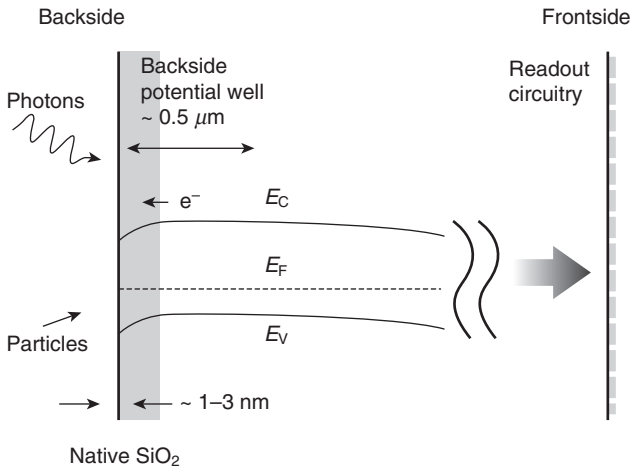
14.9 Schematic illustration of cross section of silicon imager. The detail of the frontside readout circuitry is not depicted to illustrate the point of the independence of the back surface band bending and charge trapping issue from the method of the readout circuitry.
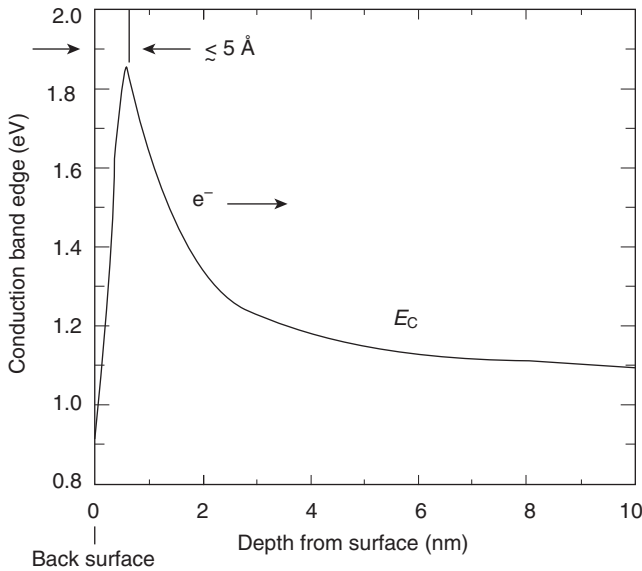
14.10 Schematic of the edge of the conduction band of a silicon imager after modification of back surface bandstructure using MBE process of delta doping. Note that the trapping well is now reduced to less than 5 Å while a field is also produced to guide the photoelectrons toward the readout circuitry in the frontside of the imager.

Recently the surface doping of pure boron has shown some success in producing arrays with high sensitivity in the ultraviolet (Nanver *et al.* 2012).

In all the cases, different techniques are used to provide charge near the surface in order to favorably affect the near surface field. In the first category, charge is chemically induced by using electronegative material on the surface. In the second category, charge is physically placed and incorporated in the silicon crystal lattice near the silicon-silicon oxide interface. In ion implant and anneal, this is achieved by implanting a dopant such as boron into the lattice, followed by an annealing process either through high temperature anneal of the entire device or by surface heating using a laser source. The annealing step is required to incorporate the dopant and to partially repair the damage induced by ion implantation. Delta doping is also in the second category, where a high sheet density of dopant with a sharply peaked distribution is incorporated in the lattice near the silicon-silicon oxide interface. Unlike ion implantation, MBE growth does not require an anneal step. Finally, the third category applies a bias on a metal contact to directly manipulate the field using the applied voltage.

Each of these techniques provides improvement with some degree of success. Physical techniques have the advantage of directly incorporating

charge into the crystal lattice to counteract the effect of the charge trapped in the interface states of silicon and its native oxide.

14.6.3 Delta doping technology as an ideal solution for back illuminated devices

Delta doping is a technology invented at the Jet Propulsion Laboratory (JPL) for producing high-performance back illuminated silicon imagers (Hoenk *et al.* 1992). Delta-doped devices with nearly 100% internal QE (Fig. 14.11(a)) have been demonstrated on various platforms, designs, and formats. These include conventional CCDs (Hoenk *et al.* 1992, Nikzad *et al.* 1994a,b), p-channel CCDs (Blacksberg *et al.* 2005), large-format arrays (Blacksberg *et al.* 2008), CMOS arrays (Hoenk *et al.* 2009, 2013), and electron-multiplied CCDs (Nikzad *et al.* 2012). Delta doping falls in the category of physical techniques for the surface passivation of back illuminated devices where charge in the form of dopant is incorporated into the crystal lattice of the back illuminated imager. Invented in 1992 at JPL (Hoenk *et al.* 1992), the technique development was motivated by the instability and low QE observed in the back illuminated UV flood CCD of the Hubble Space Telescope's wide field planetary camera's (WF/PC). The invention of delta doping was made possible by the meticulous work by Paula Grunthaner and Frank Grunthaner to understand the silicon-silicon oxide interface state and their subsequent invention of low-temperature silicon homoepitaxy (Grunthaner *et al.* 1988).

In the delta doping process, the temperature never rises above 450° C and therefore, fully fabricated devices can be processed. An ultra-thin, high quality single crystal silicon layer is grown on the thinned and atomically clean surface of silicon imager using MBE. In this layer, a very high density of boron is incorporated, nominally in a single atomic sheet overgrown with protective undoped silicon. This process is scalable to large wafer size and high throughput (Nikzad *et al.* 2013).

14.11 (a) Typical 100% internal quantum efficiency of delta doped arrays, shown here for CCDs from extreme ultraviolet to the visible part of the spectrum. The CCDs are thinned to a ~ 10 micron and the response beyond 700 nm rapidly falls to a few percent by 1000 nm part of the spectrum, (b) quantum efficiency of fully depleted delta doped p-channel high resistivity CCDs with and without antireflection (AR) coatings. The thickness of these devices allows efficient detection of higher efficiency near 1000 nm. (c) Quantum efficiency in the far ultraviolet and near ultraviolet achieved with atomic layer deposition for AR coatings and molecular beam epitaxy for bandstructure engineering of silicon – silicon oxide interface [after Nikzad, 2012].

(a)

(b)

(c)
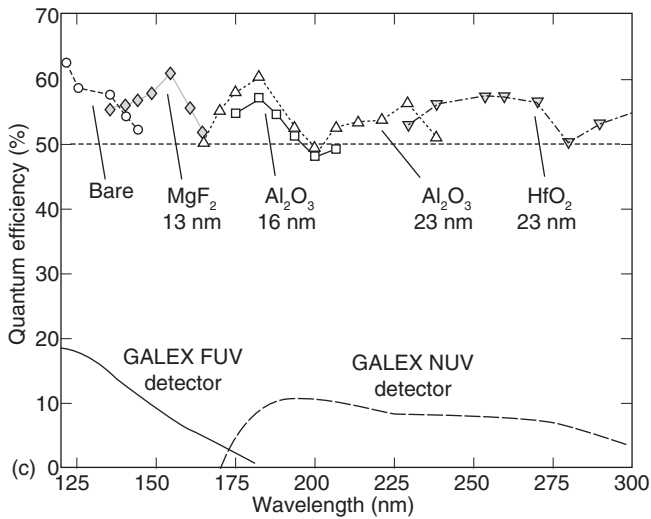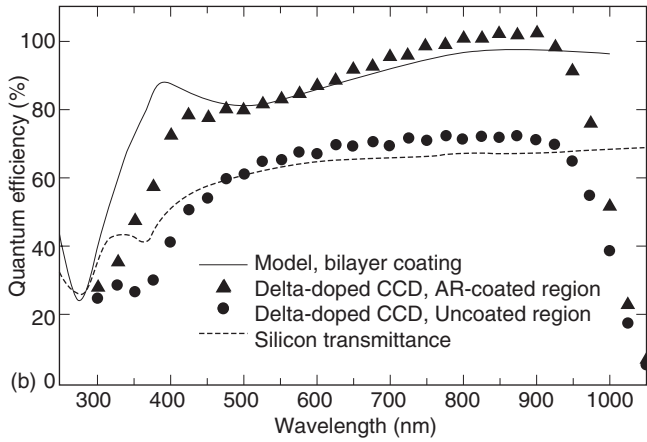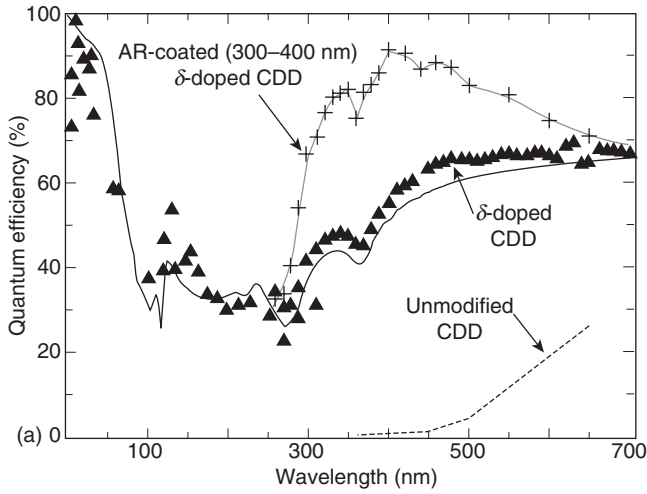
14.6.4 Reducing photon losses with advanced antireflection coatings

In the previous section, we dealt with collection of all the photoelectrons that are produced by the absorbed photons in the detector. It is of course possible to reduce reflection by depositing layers of antireflection (AR) coatings. Various techniques such as sputtering, thermal evaporation, and electron-beam evaporation, and more recently atomic layer deposition (ALD) have been employed to deposit suitable high index materials or a combination of high and low-index material on the illumination surface of back illuminated silicon detectors (Greer *et al.* 2013). The use of ALD in the past few years has opened exciting new possibilities for creating ultrathin, single layer or multilayer of sharply defined, high quality stack of films as antireflection coatings (Hamden *et al.* 2011, Nikzad *et al.* 2012). These films, in combination with delta doping, have shown record high QE (> 50%) in far and near ultraviolet (Fig. 14.11(c)).

14.6.5 Delta doped arrays and the extension of application into UV, far UV, extreme UV, soft X-rays and low energy particles

We have so far focused on applications of back illuminated imagers in the ultraviolet to near infrared spectral range. It should be noted that the band-structure engineering that removes the traps for photoelectrons, allowing higher energy photons in ultraviolet and extreme ultraviolet to be detected, also removes the so-called dead layer of particle detectors. Work has been carried out by several groups to produce solid state detectors that have a lower energy detection threshold for particles and therefore do not require the acceleration of charged (and neutral) particles prior to their detection (Funsten *et al.* 2004, Nikzad *et al.* 1998, 1999, 2006]. An advantage of these developments is that smaller, less massive plasma instruments can be developed. Applications are as diverse as miniaturized electron microscopes, compact plasma instruments for studying the solar wind, more compact space weather instruments, compact and less massive *in situ* planetary instruments such as miniature mass spectrometers, and advanced electron bombarded arrays with lower operating voltages.

14.7   Planetary and astronomy applications

Ever since Galileo used his innovative telescope design and discovered the moons of Jupiter, major leaps in discovery in space have followed technological development. NASA's Voyager spacecraft produced up close images of solar system planets using vidicon imaging technology. Since

Voyager, numerous missions for remote sensing as well as *in-situ* probing of the solar system planets have used the many advances of silicon imaging technologies. Some of these missions include Cassini (to Saturn), Galileo (to Jupiter), Kepler (exoplanet finder), MESSENGER (Mercury), and most recently the Mars Science Laboratory (MSL).

In 1976, only seven years after the invention of CCDs, the first high resolution planetary image was obtained in a ground-based observatory (Janesick 2001). Within a few years after the invention of CCDs, the Galileo spacecraft and the Hubble Space Telescope (HST) used CCDs for the first time in space. At 800×800 pixels, the CCD used in the wide field/planetary camera (WF/PC2) of HST seems small compared to even consumer cameras that are ubiquitous today. Using four CCDs at this format in the unique design of three wide field cameras and one high resolution (planetary) camera, HST's WF/PC2 sent back, in the characteristic 'staircase' format, incredible images of our universe that changed our perception of it, for example, by imaging an area of the sky previously believed dark and revealing a seemingly infinite field of galaxies. These HST deep field images have changed the way we perceive our universe.

Cassini used a CCD with a 1024×1024 format that was a variation on the design of the WF/PC2 CCD on HST. Both of these devices were operated as front illuminated; however, because of the need for ultraviolet sensitivity, the HST device was coated by a phosphor that would absorb the UV and re-emit in the green (~ 500 nm) wavelengths where CCD sensitivity was high. The efficiency of the WF/PC2 CCD was only ~10–12% and while the camera produced beautiful images, next-generation devices were required to have higher efficiency. The subsequent HST cameras have employed back illuminated CCDs in the advanced camera and the wide field camera (WFC).

Beyond HST, ultraviolet and optical telescopes for future missions are planned with larger apertures and larger focal plane arrays (FPAs) populated with large-area silicon imaging arrays. Figure 14.12 shows a few examples of these FPAs already in space, under construction, or in planning. This new demand for large focal plane arrays changes the paradigm of production and selection of science grade devices in order to allow reasonable mission cost (Scowen *et al.* 2010, Blandford 2010). With high-throughput batch processing of large wafers, it is possible, for example, to delta dope a 'lot run' of wafers in a matter of days to produce high-performance, scientific imagers (Nikzad *et al.* 2013).

Many planetary studies require UV instruments in order to understand the history of our solar system. In all cases, NASA's requirement for small and low-cost missions places exacting requirements on power, volume, and mass budgets in mission payloads. Traditionally, MCPs have been used for these applications. While these work well, they have limits, and for decades
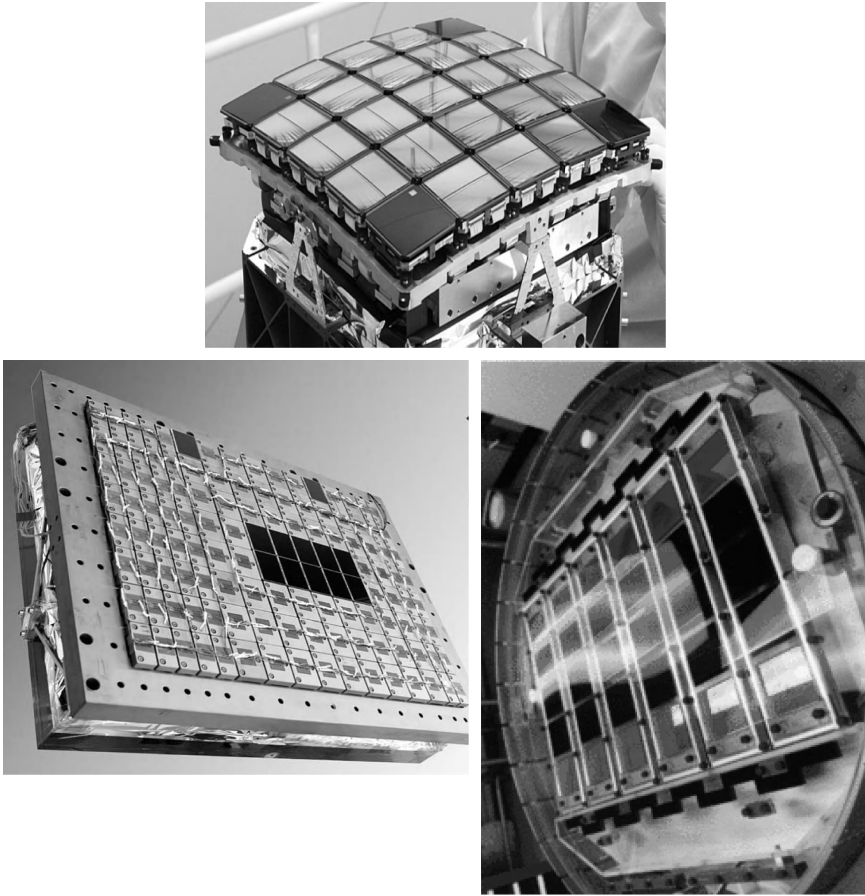
14.12 Examples of large focal plane arrays used in large ground or space based telescope. Clockwise from top the focal plane array for NASA mission Kepler (42 CCDs with 1024 × 2048 pixels), Kepler FPA (Courtesy of NASA); Sloan Digital Sky Survey or SDSS (30 CCDs of 2048 × 2048 pixel format) (Courtesy of SDSS official website); and Giai with 106 CCDs of 4500 × 1966 pixels (Courtesy ESA).

scientists have attempted the development of newer technology solar-blind UV detectors that would operate without high voltage, have low noise background, large pixel arrays with a uniform response across the array, and not be limited by count rates as MCP detectors are locally and globally. Advances in silicon imaging to achieve high efficiency and photon counting such as delta doped EMCCDs as seen in Fig. 14.11(c) or superlattice doping of SPADs or other silicon imagers will enable miniaturization, simplification, and reduced power consumption while improving the performance of UV/ VUV instruments.

14.8   Commercial applications of high-performance imaging detectors

Small digital cameras for still and video capture or live streaming have many familiar and obvious applications. As consumer products, these imagers are not held to the same exacting requirements as scientific imagers. Here we limit our discussion to applications where high performance in terms of sensitivity, noise, and stability are demanded. Specifically, we will discuss one case of machine vision, i.e., semiconductor inspection.

In the introduction, we enumerated a few of the most commonly used machine vision including cameras in consumer automobiles, surveillance cameras, and security cameras. State-of-the-art wafer and reticle inspection systems use deep ultraviolet (DUV) lasers at 263 nm and 193 nm for optical detection of defects down to the range of 20 nm and below. DUV radiation creates high densities of trapped charge in oxides and at interfaces in silicon detectors. Radiation-hardened oxides help to limit the damage, and doping the surface helps to mitigate its effects, but the fundamental problem remains. State-of-the-art silicon detectors have repeatedly failed under irradiation by pulsed DUV lasers.

Imaging systems require resolutions of approximately 8 megapixels and frame rates of up to 1000 frames per second. High-brightness sources and efficient detectors are required for maximum throughput. While back illuminated CMOS imaging arrays meet the requirements for resolution, frame rate, and noise, stability and durability under DUV illumination present major technological challenges. DUV-stable silicon detectors that would meet these requirements have been sought for years.

Recently, a different approach has been developed by using MBE to embed multiple layers of dopant atoms within a few nanometers of the silicon surface. The quantum properties of the superlattice lead to the required DUV stability and a new camera has been developed with this technology. Figure 14.13 shows the camera. The stability response to 193 nm and 260 nm photons has been measured at 500× saturation for over 24 hours showing unprecedented durability (Nikzad *et al.* 2012, Hoenk *et al.* 2013).

14.9   Brief note on biological and medical applications

Applications of silicon imaging in medical fields have been mostly for radiation detection. Excellent reviews can be found on this subject (see for example the introduction in Llacer 2005). Leveraging from the advances made in silicon imaging for astronomy, highly efficient imagers have been implemented in medical imaging, making it possible to reduce exposure of patients to radiation, reduce diagnosis turnaround time, and produce more compact medical radiation devices.

14.13 A super-lattice doped CMOS camera optimized in deep UV (260 nm) for wafer and reticle inspection applications.

More recently optical and UV imaging is being explored for intraoperative tumor delineation or post-operative diagnosis. The use of semiconductor imaging could allow minimally invasive approaches for investigation of disease. Challenges in detection and imaging in medical fields are numerous. These challenges are specific to a given application but two of the most prominent challenges for *in-vivo* imaging of the body and internal organs are imager survivability in the imaging environment and the need for doing no harm while imaging. Endoscopic applications such as the pill camera is a prime example of power of compact yet powerful imaging. Compact, high resolution, high sensitivity imagers equipped with sophisticated image processing could be used as a powerful tool during diagnosis or surgical stages. Combined with other semiconductor imaging, silicon imaging can offer a powerful multimodal imaging tool for medical applications.

14.10   References and further reading

Baggett, S. M., MacKenty, J. W., Kimble, R. A., *et al.* (2010) 'WFC3 detectors: on-orbit performance', *Proc. of SPIE*, 7731, 1.

Bai, Y., Montroy, J., Blackwell, J., Farris, M., Kozlowski L. and Vural K. (2000) 'Development of hybrid CMOS visible focal plane arrays at Rockwell', *Proc. SPIE 4028, Infrared Detectors and Focal Plane Arrays VI*, 174, July.

Bardeen, J. (1947) 'Surface states and rectification at a metal semi-conductor contact', *Phys. Rev.*, 71, 717.

Bebek, C. (2002) 'Proton radiation damage in p-channel CCDs fabricated on high-resistivity silicon', *IEEE Trans. on Nuclear Science*, 49, 1221.

Blacksberg, J., Hoenk, M. E., Elliott, T. S., Holland, S. E. and Nikzad, S. (2005) 'Enhanced quantum efficiency of high purity silicon imaging detectors by ultra-low temperature surface modification using Sb-doping', *Applied Physics Letters*, 85, 254101.

Blacksberg, J., Nikzad, S., Hoenk, M. E., Holland, S.E. and Kolbe, W. (2008) 'Near-100% quantum efficiency of delta doped large-format UV-NIR silicon imagers', *IEEE Trans. On Electron Devices*, 55, 3402.

Blandford, R. (2010) 'New worlds, new horizons in astronomy and astrophysics', National Academy of Sciences, Washington.

Charbon, E. (2011) 'Evolving CMOS technology for high-performances single-photon detection', *Optics & Photonics News*, 22(5), 14–15.

Charbon, E. (2013) Private communication.

Charbon, E. (2014) 'Single-photon imaging in CMOS processes', to appear, *Trans. Royal Society*.

Collins, N. R., Boehm, N., Delo, G., *et al.* (2009) 'Wide field camera 3 CCD quantum efficiency hysteresis: characterization and mitigation', *Proc. of SPIE*, 7439, 1.

Deelman, P., Nikzad, S. and Blouke, M. (2000) 'Delta doped CCDs with integrated UV coatings', *Proc. of SPIE*, 3965, 462.

Farr, W.H. (2009) 'Negative avalanche feedback detectors for photon-counting optical communications', *Proc. of SPIE*, 7199, 71990Q-1.

Funsten, H. O., Ritzau, S. M. and Harper, R. W. (2004) 'Fundamental limits to detection of low-energy ions using silicon solid-state detectors', *Applied Physics Letters*, 84, 3552.

Greer, F., Hamden, E., Jacquot, B. C., Hoenk, M. E., Jones, T., Dickie, M., Monacos, S. P. and Nikzad, S. (2013) 'Atomically precise surface engineering of silicon CCDs for enhanced UV quantum efficiency', *Journal of Vacuum Science and Technology A*, 01A103-1.

Grunthaner, P. J., Grunthaner, F. J., Fathauer, R. W., Lin, T. L., Schowengerdt, F. D., Pate, B., Mazur, J. H., Bean, J. C. and Schowalter. L.J. (1988) *Silicon Molecular Beam Epitaxy*, The Electrochemical Society, New Jersey, p. 375.

Grunthaner, P.J., Grunthaner, F.J., Fathauer, R.W., Lin, T.-L., Hecht, M.H., Bell, L.D., Kaiser, W.J., Schowendgerdt, F.D. and Mazur, J.H. (1989) 'Hydrogen-Terminated Silicon Substrates for Low-Temperature Molecular Beam Epitaxy', *Thin Solid Films*, 183, 197.

Hamden, E., Greer, F., Hoenk, M. E., Blacksberg, J., Jones, T. J., Dickie, M., Jacquot, B., Monacos, S., Martin, C., Morrissey, P., Schiminovich, D. and Nikzad, S. (2011) 'Antireflection coatings designs for use in UV imagers', *Applied Optics*, 50, 4180.

Hester, J. J., Scowen, P. A., Sankrit, R., Lauer, T. R., Ajhar, E. A., Baum, W. A., Code, A., Currie, D. G., Danielson, G. E., Ewald, S. P., Faber, S. M., Grillmair, C. J.; Groth, E. J., Holtzman, J. A., Hunter, D. A., Kristian, J., Light, R. M., Lynds, C. R., Monet, D. G., O'Neil, E. J., Jr., Shaya, E. J., Seidelmann, K. P. and Westphal, J. A. (1996) 'Hubble Space Telescope WFPC2 imaging of M16: photoevaporation and emerging young stellar objects', *Astronomical Journal*, 111, 2349.

Hoenk, M. E. (2010) 'Delta doping technology for UV photon counting detector arrays', Invited Presentation, KISS Single Photon Counting Detectors Workshop, California Institute of Technology, Pasadena, California, 28 January, http://www.kiss.caltech.edu/workshops/photon2009/presentations/hoenk.pdf

Hoenk, M. E., Grunthaner, P. J., Grunthaner, F. J., Fattahi, M., Tseng, H.-F. and Terhune, R. W. (1992) 'Growth of a delta-doped silicon layer by molecular-beam epitaxy on a charge-coupled device for reflection-limited ultraviolet quantum efficiency', *Applied Physics Letters*, 61, 1084.

Hoenk, M. E., Jones, T. J., Dickie, M. R., Greer, F., Cunningham, T. J., Blazejewski, E. R. and Nikzad, S. (2009) 'Delta-doped back-illuminated CMOS imaging arrays: progress and prospects', *Proc. of the SPIE*, 74190, 74115.

Hoenk, M.E., Jacquot, B. C., Monacos, S. P. and Nikzad, S. (2010) 'The photometric stability of delta-doped detectors', *Proc. of SPIE*, 7742.

Hoenk, M.E., Carver, A., Jones, T., Dickie, M., Cheng, P., Greer, F. and Nikzad, S. (2013) 'DUV stability of superlattice doped CMOS detector arrays', 2013 International Image Sensor Workshop, 12–16 June, Snowbird, Utah.

Holland, S. E., Groom, D. E., Palaio, N. P., Stover, R. J. and Wei, M. (2003) 'Fully Depleted, Back-Illuminated Charge-Coupled Devices Fabricated on High-Resistivity Silicon', *IEEE Transaction on Electron Devices*, 50, 225.

Hynecek, J. (2001) 'Impactron – a new solid state image intensifier', *IEEE Transactions on Electron Devices*, 48(10), 2238, 2241.

Jacquot, B. C., Hoenk, M. E., Jones, T., Cunningham, T. J. and Nikzad, S. (2012) 'Direct detection of 1000–5000 eV electrons with delta doped silicon CMOS and electron-multiplying CCD imagers', *IEEE Transactions on Electron Devices*, 59, 1988.

Janesick, J. R. (2001) *Scientific Charge Coupled Devices*, SPIE Press Bellingham, Washington.

Janesick, J. and Putman, G. (2003) 'Developments and applications of high-performance CCD and CMOS imaging arrays', *Annual Review of Nuclear and Particle Science*, 53 (December), 263–300.

Janesick, J., Elliott, T. and Pool, F. (1989) 'Radiation damage in scientific charge-coupled devices', *IEEE Transactions on Nuclear Science*, 36(1), 572, 578.

Jenkins, J. M., Barouki, W. J., Dunham, E. W. and McDonald, J. S. (1997) 'High precision photometry with back-illuminated CCDs', in *Planets Beyond our Solar System and Next Generation of Space Missions*, (ed.) Sodarblum, D., ASP Conf. Ser. 119, 277–280.

Jerram, P., Pool, P., Bell, R., Burt, D., Bowring, S. and Spencer, S. (2001) 'LLLCCD—low light level imaging without the need for an intensifier', *Proc. SPIE* 4306, 178.

Jiang, Z. M., Xiu, L. S., Jiang, X. M., Zheng, W. L., Lu, X. K., Zhu, H. J., Zhang, X. J. and Wang, X. (1997) Chinese *Phys. Lett.*, 14, 686.

Kasdin, J. (2009a) 'Exoplanets and disks: their formation and diversity', *AIP Conference Proceedings*, 1158, 359.

Kasdin, J. (2009b) 'THEIA: telescope for habitable exoplanets and interstellar/intergalactic astronomy', http://www.astro.princeton.edu/~dns/Theia/nas_theia_v14.pdf

Knezevic, T., Suligoj, T., Sakic, A. and Nanver, L.K. (2012) 'Modeling of electrical characteristics of ultra-shallow pure amorphous boron p+n junctions', *Proc. of the 35th International Convention MIPRO 2012*, Opatija, Croatia, 21–25 May, pp. 36–41.

Lepri, S., Nikzad, S., Jones, T., Blacksberg, J. and Zurbuchen, T. H. (2006) 'Response of a delta-doped charge-coupled device to low energy protons and nitrogen ions', *Rev. Sci. Instrumation*, 77, 053301.

Lesser, M. and Iyer, V. (1998) 'Enhancing back illuminated performance of astronomical CCDs', *Proc. of the SPIE*, 3355, 446.

Llacer, G. L. (2005) 'Silicon detectors for medical imaging development of a first Compton probe prototype for improved prostate imaging', *PhD Thesis, Departement de Fisica Atomica Molecular i Nuclear*. (See Introduction and references therein).

Mackay, C. D., Tubbs, R. N., Bell, R., Burt, D., Jerram, P. and Moody, I. (2008) 'Sub-electron read noise at MHz pixel rates', *Proc. SPIE*, 4306, 289.

Mandai, S. and Charbon, E. (2012) 'A wide spectral range single-photon avalanche diode fabricated in an advanced 180nm CMOS technology', *Optics Express*, 20, 5849.

Nanver, L. K., Sammak, A., Sakic, A., Mohammadi, V., Derakhshandeh, J., Mok, K. R. C., Qi, L., Golshani, N., Scholtes, T. M. L. and de Boer, W. B. (2012) 'Applications of PureB and PureGaB ultrashallow junction technologies', 2012 IEEE 11th International Conference on Solid-State and Integrated Circuit Technology (ICSICT), October, pp. 1, 4.

Niclass, C., Favi, C., Kluter, T., Gersbach, M. and Charbon, E. (2008) 'A 128 × 128 single-photon image sensor with column-level 10-bit time-to-digital converter array', *IEEE Journal of Solid-State Circuits*, 43, 2977.

Nikzad, S., Hoenk, M.E., Grunthaner, P.J., Terhune, R.W., Grunthaner, F.J., Winzenread, R., Fattahi, M. and Tseng, H-F. (1994a) 'Delta-doped CCDs: high QE with long-term stability at UV and visible wavelengths', *Proc. of SPIE*, 2198, 907.

Nikzad, S., Hoenk, M. E., Grunthaner, P. J., Terhune, R. W., Winzenread, R., Fattahi, M., Tseng, H-F. and Grunthaner, F. J. (1994b) 'Delta-doped CCDs as stable, high sensitivity, high resolution UV imaging arrays', *Proc. of SPIE*, 2217, 355.

Nikzad, S., Yu, Q., Smith, A. L., Jones, T. J., Tombrello, T. A. and Elliott, S. T. (1997) 'Direct detection and imaging of low-energy electrons with delta-doped charge-coupled devices', *Applied Physics Letters*, 73, 3417.

Nikzad, S., Yu, Q., Smith, A. L., Elliott, S. T., Jones, T. J. and Tombrello, T. A. (1998) 'Direct detection and imaging of low-energy electrons with delta-doped charge-coupled devices', *Applied Physics Letters*, 241.

Nikzad, S., Croley, D., Elliott, S., Cunningham, T. J., Proniewicz, W. K., Murphy, G. B. and Jones, T. J. (1999) 'Single-event keV proton detection using a delta-doped charge-coupled device', *Applied Physics Letters*, 75(17), 2686–2688.

Nikzad, S., Jones, T., Elliot, S. T., Cunningham, T., Deelman, P., Walker, A. and Oluseyi. H. (2000) 'Ultrastable and uniform EUV and UV detectors', *Proc. SPIE 4139, Instrumentation for UV/EUV Astronomy and Solar Missions*, 250.

Nikzad, S., Cunningham, T. J., Hoenk, M. E., Ruiz, R. P., Soules, D. M. and Holland, S. E. (2006) 'Direct detection of 0.1–20 keV electrons with delta-doped, fully depleted, high purity silicon p-i-n diode arrays', *Applied Physics Letters*, 89, 182114.

Nikzad, S., Hoenk, M., Greer, F., Jacquot, B., Monacos, S., Jones, T., Blacksberg, J., Hamden, E., Schiminovich, D., Martin, C. and Morrissey, P. (2012) 'Delta-doped electron-multiplied CCD with absolute quantum efficiency over 50% in the near

to far ultraviolet range for single photon counting applications', *Applied Optics*, 51, 356–369.

Nikzad, S., Hoenk, M. E., Carver, A. G., Jones, T. J., Greer, F., Hamden, E. and Goodsall, T. (2013) 'High throughput, high yield fabrication of high quantum efficiency back-illuminated photon counting, far UV, UV, and visible detector arrays', *Proc. of the 2013 International Image Sensor Workshop*, 1, 37.

Postman, M. (2009) 'Advanced technology large aperture space telescope (ATLAST): a technology roadmap for the next decade', NASA Strategic Mission Concept Study.

Scowen, P., Nikzad, S., Hoenk, M. E., *et al.* (2010) 'Large focal plane arrays for future missions', *Astro2010: The Astronomy and Astrophysics Decadal Survey, Technology Development Papers*.

Smith, D.R., Ingley, R. and Holland, A. (1987) 'Ultraviolet and extreme ultraviolet response of charge-coupled devices','' *Optical engineering*, 26, 875.

Smith, D. R., Ingley, R. and Holland, A. (2006) 'Proton irradiation of EMCCDs', *IEEE Transactions on Elect. Devices*, 53(2), 205.

Stahl, H. P., Stephens, K., Henrichs, T., Smart, C. and Prince, F. A. (2010) 'Single-variable parametric cost models for space telescopes', *Opt. Eng.*, 49, 073006.

Theuwissen, A.J.P. (1995) 'Solid-state imaging with charge-coupled devices', Kluwer Academic Publishers, Dordrecht, Boston, London.

Tripathi, N., Bell, L. D., Nikzad, S. and Shahedipour-Sandvik, F. (2010a) 'Effect of n+GaN cap thickness on Cs-free GaN photocathode characteristics', *Applied Physics Letters*, 97, 052107.

Tripathi, N., Bell, L. D., Nikzad, S., Tungare, M., Suvarna, P. H. and Shahedipour Sandvik, F. (2010b) 'Novel Cs-free GaN photocathodes', *Journal of Electronic Materials*, 40, 382.

Westhoff, R. C., Rose, M. K., Gregory, J. A., Berthiaume, G. D., Seely, J. F., Woods, T. N. and Ucker, U. (2007) 'Radiation-hard, charge-coupled devices for the extreme ultraviolet variability experiment', *Proc. of SPIE*, 6686, 4.

Westhoff, R. C., Burke, B. E., Clark, H. R., Loomis, A. H., Young, D. J., Gregory, J. A. and Reich, R. K. (2009) 'Low-dark-current, back-illuminated charge-coupled-devices', *Proc. SPIE*, 7249, 0J.

Williams, G., Janesick, J., Dosluoglu, T. and Reinheimer, A. (1997) 'Advances in Pixelvision, Inc. Back-Illuminated CCD Imaging Technologies for Low Light Level Imaging Applications', DTIC Online.

Wilkinson, F. J., Farmer, A. J. D. and Geist, J. (1982) 'The near ultraviolet quantum yield of silicon', *J. Appl. Phys.*, 54, 1172.