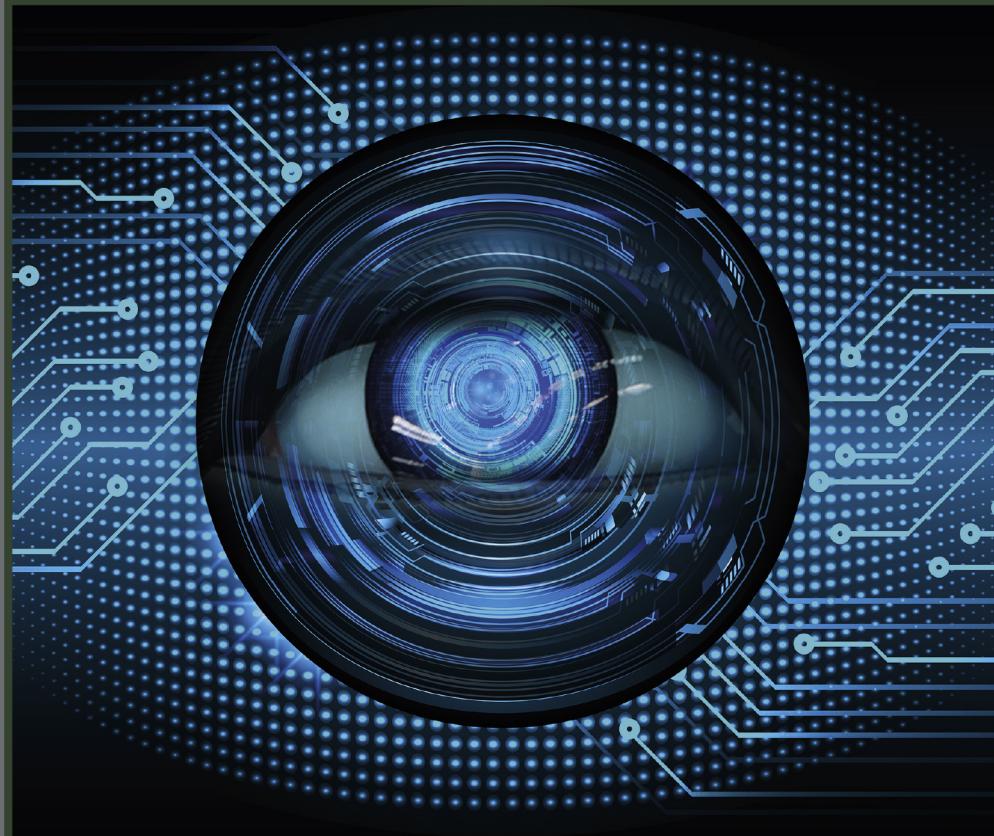


WOODHEAD PUBLISHING SERIES IN ELECTRONIC AND OPTICAL MATERIALS



# HIGH PERFORMANCE SILICON IMAGING

## FUNDAMENTALS AND APPLICATIONS OF CMOS AND CCD SENSORS

SECOND EDITION



Edited by  
**DANIEL DURINI**

# **High Performance Silicon Imaging**

This page intentionally left blank

**Woodhead Publishing Series in Electronic and  
Optical Materials**

# **High Performance Silicon Imaging**

**Fundamentals and Applications of  
CMOS and CCD Sensors**

*Second Edition*

*Edited by*

***Daniel Durini***



An imprint of Elsevier

Woodhead Publishing is an imprint of Elsevier  
The Officers' Mess Business Centre, Royston Road, Duxford, CB22 4QH, United Kingdom  
50 Hampshire Street, 5th Floor, Cambridge, MA 02139, United States  
The Boulevard, Langford Lane, Kidlington, OX5 1GB, United Kingdom

© 2020 Elsevier Ltd. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: [www.elsevier.com/permissions](http://www.elsevier.com/permissions).

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

#### Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability 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.

#### Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

#### British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-0-08-102434-8 (print)

ISBN: 978-0-08-102435-5 (online)

For information on all Woodhead publications visit our website at <https://www.elsevier.com/books-and-journals>

*Publisher:* Matthew Deans  
*Acquisition Editor:* Kayla Dos Santos  
*Editorial Project Manager:* John Leonard  
*Production Project Manager:* Sojan P. Pazhayattil  
*Cover Designer:* Mark Rogers

Typeset by SPi Global, India



Working together  
to grow libraries in  
developing countries

[www.elsevier.com](http://www.elsevier.com) • [www.bookaid.org](http://www.bookaid.org)

# **Dedication**

To Lizette.

This page intentionally left blank

# Contents

<b>Contributors</b>	xiii
<b>Preface</b>	xv

## Part One Fundamentals

<b>1 Fundamental principles of photosensing</b>	<b>3</b>
<i>D. Durini and D. Arutinov</i>	
1.1 Introduction	3
1.2 The human vision system	3
1.3 Photometry and radiometry	10
1.4 History of photosensing	12
1.5 Early developments in photodetector technology	18
References	22
Further reading	24
<b>2 Operational principles of silicon image sensors</b>	<b>25</b>
<i>D. Durini and D. Arutinov</i>	
2.1 Introduction	25
2.2 Silicon phototransduction	25
2.3 Principles of CCD and CMOS photosensing technologies	31
2.4 Metal-oxide-semiconductor-capacitor (MOS-C) structure-based photodetectors	34
2.5 <i>p-n</i> junction-based photodetectors	40
2.6 Noise considerations in pixel structures	44
2.7 High-performance pixel structures	52
2.8 Miniaturization and other development strategies followed in image sensor technologies	56
2.9 Single-photon counting	59
2.10 Hybrid and 3D detector technologies	61
2.11 Conclusion	68
References	69
Further reading	73

<b>3 Charge-coupled device (CCD) image sensors</b>	<b>75</b>
<i>M. Lesser</i>	
3.1 Introduction	75
3.2 CCD design, architecture, and operation	75
3.3 Illumination modes	82
3.4 Imaging parameters and their characterization	85
3.5 Conclusion and future trends	91
References	92
<b>4 Backside illuminated (BSI) complementary metal-oxide-semiconductor (CMOS) image sensors</b>	<b>95</b>
<i>A. Lahav, A. Fenigstein, A. Strum, and S. Rizzolo</i>	
4.1 Introduction	95
4.2 Challenges facing a scaled-down FSI sensor	97
4.3 Basics of BSI sensor process integration	101
4.4 Interface solutions to BSI sensors	111
4.5 Conclusion	114
References	115
<b>5 CMOS circuits for high-performance imaging</b>	<b>119</b>
<i>Bhaskar Choubey, Waqas Mughal, and Luiz Gouveia</i>	
5.1 High-resolution image sensors	119
5.2 Low-noise CMOS image sensors	124
5.3 High-speed image sensors	130
5.4 Low-power image sensors	140
5.5 WDR sensors	141
5.6 Other high-performance designs	152
5.7 Conclusion	153
References	154
<b>6 Smart cameras on a chip: Using complementary metal-oxide-semiconductor (CMOS) image sensors to create smart vision chips</b>	<b>161</b>
<i>D. Ginbac</i>	
6.1 Introduction	161
6.2 The concept of a smart camera on a chip	163
6.3 The development of vision chip technology	165
6.4 From special-purpose chips to smart computational chips	167
6.5 From video rate applications to high-speed image processing chips	170
6.6 Future trends	174
6.7 Conclusion	175
References	176
Further reading	181

## Part Two Applications

<b>7 CMOS image sensor technology advances for mobile devices</b>	<b>185</b>
<i>Robert J. Gove</i>	
7.1 Introduction	185
7.2 Core image/video capture technology requirements and advances in mobile applications	191
7.3 Emerging CMOS “sensor-embedded” technologies	203
7.4 Mobile image sensor architecture and product considerations	218
7.5 Future trends	224
7.6 Vision for the future of mobile silicon imaging	231
7.7 Conclusion	235
References	236
Further reading	240
<b>8 Complementary metal-oxide-semiconductor (CMOS) image sensors for automotive applications</b>	<b>241</b>
<i>C. De Locht and H. Van Den Broeck</i>	
8.1 Automotive applications	241
8.2 Vision systems	242
8.3 Sensing systems	244
8.4 Requirements for automotive image sensors	246
8.5 Future trends	253
<b>9 CMOS and CCD image sensors for space applications</b>	<b>255</b>
<i>P. Jerram and K. Stefanov</i>	
9.1 Introduction	255
9.2 Imaging modes in space applications	255
9.3 Important additional requirements for image sensors in space	258
9.4 Performance of CMOS and CCD image sensors for space	259
9.5 Space applications	273
9.6 Conclusion and longer term trends	283
References	284
<b>10 Complementary metal-oxide-semiconductor (CMOS) sensors for high-performance scientific imaging</b>	<b>289</b>
<i>R. Turchetta</i>	
10.1 Introduction	289
10.2 Detection in silicon	290
10.3 CMOS sensors for the detection of charged particles	298
10.4 CMOS sensors for X-ray detection	307
10.5 Future trends	310
10.6 Sources of further information and advice	311
References	311
Further reading	317

<b>11 CMOS-based optical time-of-flight 3D imaging and ranging</b>	<b>319</b>
<i>R. Lange, S. Böhmer, and B. Buxbaum</i>	
11.1 Introduction to 3D imaging and ranging	319
11.2 Concept and design considerations for ToF cameras	325
11.3 Comparison of ToF with triangulation-based approaches	338
11.4 CMOS ToF image sensors	342
11.5 Applications	360
References	373
Further reading	375
<b>12 CMOS sensors for fluorescence lifetime imaging</b>	<b>377</b>
<i>Robert K. Henderson, Bruce R. Rae, and Day-Uei Li</i>	
12.1 Introduction	377
12.2 Fluorescence lifetime imaging	378
12.3 CMOS detectors and pixels	386
12.4 FLIM system-on-chip	395
12.5 Outlook	405
References	406
<b>13 Complementary metal-oxide-semiconductor (CMOS) X-ray sensors</b>	<b>413</b>
<i>A. Strum, A. Fenigstein, and S. Rizzolo</i>	
13.1 Introduction	413
13.2 Intraoral and extraoral dental X-ray imaging	414
13.3 Medical radiography, fluoroscopy, and mammography	415
13.4 CIS-based FPD technology	416
13.5 Pixel design considerations for CMOS-based FPDs	417
13.6 Key parameters for X-ray sensors	425
13.7 X-ray sensors: Types and requirements	428
13.8 Direct X-ray sensors	429
13.9 Conclusion and future trends	435
References	435
<b>14 Complementary metal-oxide-semiconductor (CMOS) and charge coupled device (CCD) image sensors in high-definition TV imaging</b>	<b>437</b>
<i>P. Centen</i>	
14.1 Introduction	437
14.2 Broadcast camera performance	438
14.3 Modulation transfer function, aliasing and resolution	439
14.4 Aliasing and OLP filtering	443
14.5 Opto-electrical matching and other parameters	447
14.6 Standards for describing the performance of broadcast cameras	451

---

14.7	CCD and CMOS image sensors used in broadcast cameras	452
14.8	Signal-to-noise ratio	456
14.9	Bit size, pixel count, and other issues	464
14.10	Three-dimensional and ultrahigh-definition (UHD) television	467
14.11	Conclusion	468
14.12	Sources of further information and advice	468
	References	469
<b>15</b>	<b>High-performance silicon imagers, back illumination using delta and superlattice doping, and their applications in astrophysics, medicine, and other fields</b>	<b>473</b>
	<i>S. Nikzad and M.E. Hoenk</i>	
15.1	Introduction	473
15.2	Solid-state imaging detectors: Principles of operation	474
15.3	Scientific imaging detectors	478
15.4	Readout structures	479
15.5	Photon counting detectors	480
15.6	High-performance imaging through back illumination	485
15.7	Planetary and astronomy applications	492
15.8	Commercial applications of high-performance imaging detectors	495
15.9	Brief note on biological and medical applications	495
	References	496
	Further reading	500
<b>Index</b>		<b>503</b>

This page intentionally left blank

# Contributors

**D. Arutinov** Central Institute of Engineering, Electronics and Analytics  
ZEA-2—Electronic Systems, Forschungszentrum Jülich GmbH, Jülich, Germany

**S. Böhmer** PMD Technologies AG, Siegen, Germany

**B. Buxbaum** PMD Technologies AG, Siegen, Germany

**P. Centen** Grass Valley, Nederland BV, Breda, The Netherlands

**Bhaskar Choubey** Chair of Analogue Circuits and Image Sensors, Siegen University,  
Siegen, Germany

**H. Van Den Broeck** Melexis Technologies NV, Tessenderlo, Belgium

**D. Durini** National Institute of Astrophysics, Optics and Electronics (INAOE),  
Tonantzintla, Puebla, Mexico

**A. Fenigstein** TowerJazz, Newport Beach, CA, United States

**D. Ginhac** Université de Bourgogne, Dijon, France

**Luiz Gouveia** IMSE-CNM - Seville Institute of Microelectronics, Sevilla, Spain

**Robert J. Gove** Synaptics Inc., San Jose, CA, United States

**M.E. Hoenk** Jet Propulsion Laboratory, California Institute of Technology,  
Pasadena, CA, United States

**P. Jerram** Teledyne e2v, Chelmsford, United Kingdom

**A. Lahav** Tower Semiconductor Ltd, Migdal HaEmek, Israel

**R. Lange** University of Applied Sciences Bonn-Rhein-Sieg, Sankt Augustin,  
Germany

**M. Lesser** Steward Observatory, UA Imaging Technology Laboratory, University of  
Arizona, Tucson, AZ, United States

**Day-Uei Li** University of Strathclyde, Glasgow, United Kingdom

**C. De Locht** Melexis Technologies NV, Tessenderlo, Belgium

**Waqas Mughal** School of Engineering, University of Glasgow, Glasgow, United Kingdom

**S. Nikzad** Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, United States

**Bruce R. Rae** ST Microelectronics, Edinburgh, United Kingdom

**S. Rizzolo** Institut Supérieur de l’Aéronautique et de l’Espace, Toulouse, France

**Robert K. Henderson** University of Edinburgh, Edinburgh, United Kingdom

**K. Stefanov** The Open University, Milton Keynes, United Kingdom

**A. Strum** TowerJazz, Newport Beach, CA, United States

**R. Turchetta** IMASENIC, Barcelona, Spain

# Preface

The first edition of this book was aimed at presenting a unifying discussion on different aspects of manufacturing and applications of high-performance silicon image sensors in the context of current state-of-the-art developments carried out by the industry on the one side, and the academic discussions regarding the fundamental principles of these technologies and the manufacturing challenges they face on the other. One additional aspect we introduced was a historical context of the milestones these technologies have been reaching in time and an explanation on how these milestones were attained and who stood behind the developments that took this industry to the point where it stands today. Following the same approach, this second edition tries further on to maintain a balance between theory and the technological developments and applications pursued by this constantly changing industrial sector, including the developments carried out since 2014—the year the first edition of this book appeared. In this second edition we added one additional chapter accompanied by a thorough update of all the chapters that formed part of the first edition.

I cannot thank enough Robert Lange, formerly from PMD Technologies AG from Siegen in Germany and currently a professor at the University of Applied Sciences Bonn-Rhein-Sieg, and Stephan Böhmer from PMD Technologies AG for contributing the completely new chapter ([Chapter 11](#)) on CMOS-based optical time-of-flight 3D imaging and ranging, a very important topic only superficially addressed in the first edition of this book. My profound thanks goes also to Paul Jerram from Teledyne e2v in the United Kingdom, and Konstantin Stefanov from the Open University, UK, for completely rewriting [Chapter 9](#) that appeared in the first edition originally written by Jan Bogaerts from CMOSIS (acquired by AMS AG in 2015) in Belgium. The new updated [Chapter 9](#) presents very useful information concerning reliability tests, an extensive explanation of different space applications for image and radiation sensors, a very useful description of multispectral imagers used nowadays, and an interesting review on different aspects of Sun imaging. My gratitude goes also to Serena Rizzolo from ISAE SUPAERO—Institut Supérieur de l’Aéronautique et de l’Espace in France, who kindly accepted to review the original [Chapter 4](#) dealing with technology issues and applications of backside illuminated (BSI) CMOS image sensors, as well as [Chapter 12](#) from the first edition that became [Chapter 13](#) in this second one, addressing CMOS X-ray sensors. Both original chapters were very well written by Assaf Lahav, Amos Fenigstein, and Avi Strum from TowerJazz, Israel.

I am in great debt with all the authors of the different chapters included in this second edition, who revised and updated their original texts in order to present all the new developments carried out over the past 5 years. I really hope this book can help all the

undergraduate and graduate students pursuing their studies in electrical-electronic engineering or physics and all the professionals engaged in research, design, and manufacturing of silicon image sensors to have a better understanding of the different issues involved in this beautiful area of research and development that still has so much to offer. I hope we can meet soon and exchange ideas on these endless topics.

**Daniel Durini**

# **Part One**

# **Fundamentals**

This page intentionally left blank

# Fundamental principles of photosensing

1

D. Durini<sup>a</sup>, D. Arutinov<sup>b</sup>

<sup>a</sup>National Institute of Astrophysics, Optics and Electronics (INAOE), Tonantzintla, Puebla, Mexico, <sup>b</sup>Central Institute of Engineering, Electronics and Analytics ZEA-2—Electronic Systems, Forschungszentrum Jülich GmbH, Jülich, Germany

*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 colored 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 human brain and the human vision system do the rest of the job. The smaller these picture elements are and the more different intermediate values between complete darkness and complete illumination or different individual colors they might possess, the better is the resolution and the quality of the resulting image in our brain. The concept of mosaic painting has 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 third millennium BC.

## 1.2 The human vision system

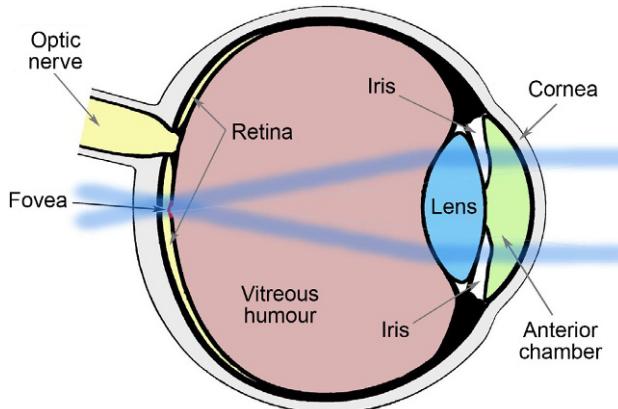
The concept of mosaic painting proved to be very successful mainly because it was developed based on the empiric knowledge of the functionality of our own human vision system and the ability of our brain to interpret the incoming information in a logical manner. Human vision is based on an optical system 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 center 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 trichromatic system to detect and separate colors, with photonic energy being measured using three different types of band-pass



**Fig. 1.1** Detail showing the Greek God Dionysus, part of the central image of the Dionysus Mosaic from around AD 220–230 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  $10.6 \text{ m} \times 7 \text{ m}$  and was made out of approximately 1.5 million stone tiles

Photograph taken by Daniel Durini in 2012.



**Fig. 1.2** A schematic diagram of the human eye.

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.

The light entering the human eye first interacts with the cornea (see Fig. 1.2), where the air-cornea interface transmittance happens to be of some 98% (Kaschke et al., 2014). This remarkable transparency of the human cornea is mainly caused by the stacked *lamellae* building the cornea tissue. They are constituted by collagen fibrils running parallel to each other and having regular spacing. These collagen fibrils happen to be very poor scatterers as their diameters (25–35 nm) are much smaller than the wavelengths forming part of the visible part of the spectra (380–780 nm) and their spatial distribution reduces additional scattering due to induced destructive interference (Kaschke et al., 2014). So, after the incident light rays have been minimally refracted by the cornea, they travel through the anterior eye chamber and cross the iris. The iris, with its variable inner aperture diameter, limits the so-called visual field of view (FOV) for incident light rays to 105 degrees at the outer parts of the human eyes, and to some 60 degrees on the nasal side (Kaschke et al., 2014) per eye. In ophthalmology and optometry, the term “pupil” is often referred to as the hole of the iris aperture, although technically the iris is actually the aperture stop. It is the image of this aperture that will finally be formed by the cornea.

After passing the iris, the light rays travel through the posterior chamber and enter the eye’s lens, adjustable to the distance of objects being fixated. The eye’s lens consists of multiple shells stacked layer by layer, each with a different refractive index starting with 1.42 found in the core and the varying downwards (Kaschke et al., 2014). Passing the eye’s lens, the light rays pass through the vitreous humor and are eventually impinging the retina.

According to Østerberg (1935, as cited by Rippes 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 (1866, as cited by Rippes and Weale in Davson, 1976), they are associated with scotopic (nocturnal) and photopic (diurnal) visions, respectively, which formed

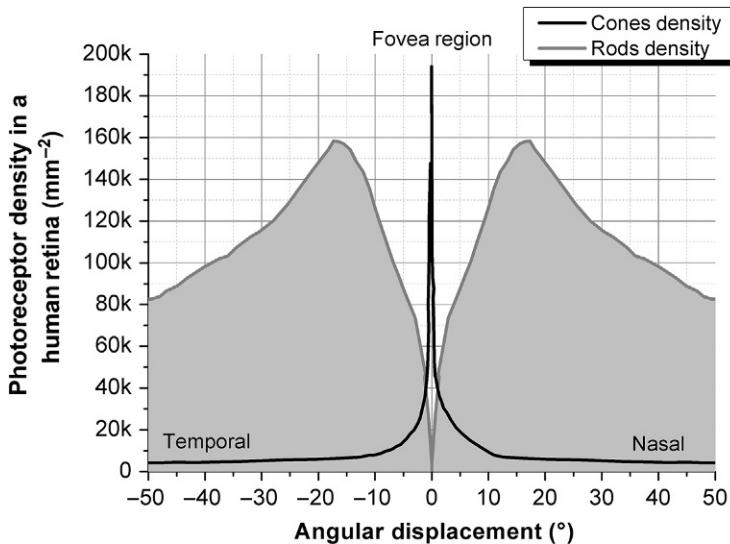
**Table 1.1** The physiological basis of the duality theory in a human being (Ripps and Weale in [Davson, 1976](#))

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
Color vision	Absent	Present

the basis for the so-called “duality 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 center (called “fovea centralis”) is of around 195,000 per  $\text{mm}^2$ , as can be observed in [Fig. 1.3](#), abruptly falling to about 9500 per  $\text{mm}^2$  in the parafovea region ( $\sim 2 \text{ mm}$  from the foveal center), and changes little from there to the retinal border—the “ora serrata.” The cones in the region of the foveal center, with an average spacing of approximately  $1.9 \mu\text{m}$  between the neighboring cells, are responsible for delivering a highly resolved image to the brain. Rods, on the other hand, are first encountered  $130 \mu\text{m}$  from the foveal center, their numbers reaching a maximum some  $5\text{--}6 \text{ mm}$  from the fovea center, 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 degrees in contrast to the 105 degrees or more ([Kaschke et al., 2014](#)) of an average human field of view. It should be noted that on the nasal side, this angle is reduced to some 60 degrees by the nose itself. 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  $260 \mu\text{m}$ , 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 8000 cones, that is, 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](#)).

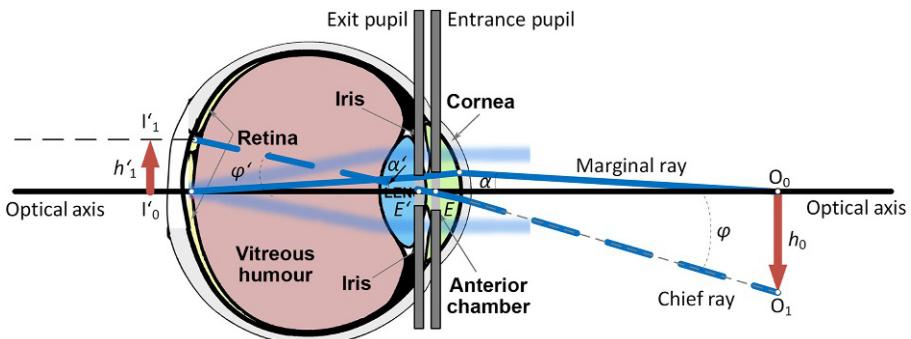


**Fig. 1.3** Average density of rods and cones in a human retina, plotted as a function of the angular separation from the fovea region.

After data published by Rippes, H., Weale, R.A., 1976a. The visual photoreceptors. In: Davson, H. (Ed.), *The Eye: Visual Function in Man*, vol. 2A, second ed, Academic Press Inc., New York, pp. 101–129 in Davson, H., 1976. *The Eye: Visual Function in Man*, vol. 2A, second ed., p. 616, Academic Press, New York.

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 Rippes and Weale in [Davson, 1976](#)).

As shown in [Fig. 1.4](#), the image formed on the retina is “inverted” if compared to the orientation of the actual image being looked at, which is analogous to the imaging of a single lens. Nevertheless, another “inversion” process happens in the human brain ([Kaschke et al., 2014](#)), which then results in the correct visual perception of the image we try to perceive. If we follow the paths of a marginal ray and the chief ray, respectively, to understand how an image is being formed by the human vision system, and we assume that the iris is centered on an optical axis, we can conclude that, as shown in [Fig. 1.4](#), the chief ray emanates from the outermost off-axis point  $O_1$  of the focused object, is refracted by the cornea, it crosses the center of the iris, and is again refracted by the eye lens to finally hit the retina at point  $I'_1$  ([Kaschke et al., 2014](#)). The marginal ray, on the other hand, emanates from the object point  $O_0$  and, after crossing the cornea, the iris, and the eye lens, finally impinges the inner edge of the iris. When we extend the object- and image side parts of the marginal ray to the pupil planes, we



**Fig. 1.4** Geometry of the light rays penetrating the human eye creating our capability to produce images on the retina. The analysis is based on the chief ray and marginal ray optical paths, respectively, starting at the object observed and ending at the eye's retina.

After Kaschke, M., Donnerhacke, K.H., Rill, M.S., 2014. Optical Devices in Ophthalmology and Optometry: Technology, Design Principles and Clinical Applications, Wiley-VCH Verlag GmbH & Co. KGaA, Germany, p. 638, ISBN: 978-3-527-41068-2.

obtain the inner diameters of the entrance and exit pupils, respectively (Kaschke et al., 2014). To determine  $|h'_1|$ , the image size on the retina, the chief ray can be used, resulting in the relation expressed by (Kaschke et al., 2014)

$$|h'_1| = \phi' \overline{E'I'_0} \quad (1.1)$$

In Eq. (1.1),  $\overline{E'I'_0}$  stands for the distance between the center of the exit pupil  $E'$  and the on-axis image point  $I'_0$ , and  $\phi'$  is the angle formed by the chief ray and the optical axis inside the vitreous humor, once the chief ray has crossed the exit pupil (see Fig. 1.4). The iris of each human eye is decentered nasally by about 0.5 mm relative to the optical axis formed by the cornea and the eye lens (Kaschke et al., 2014). So, in contrast to centered optical systems, the angles  $\phi$  and  $\phi'$  are not equal, and it can be shown that the relation  $\phi'/\phi = m$  is constant (Kaschke et al., 2014).

In healthy eyes, the refractive power of the eye lens contributes only  $\leq 30\%$  to the total eye refraction, and within a certain limit, in a process called *accommodation*, it is able to change the refractive power so that nearby, as well as distant objects can be sharply imaged on the retina. In this process, if the eye focuses on nearby objects, the ciliary muscle, surrounding the eye lens, is contracted and the zonular fibers are relaxed giving the lens a more spherical shape (for an accommodated eye). The eye's total power increases to enable near vision if the lens becomes more strongly curved. For far vision, where the eye has to focus on objects located far away, the deformable, elastic lens is brought to a more elliptical shape by relaxing the ciliary muscle (for a relaxed eye) (Kaschke et al., 2014).

To enable humans to get a three-dimensional impression of the environment, both eyes, arranged in a common plane and separated by an interpupillary distance of some 50–75 mm, enable for binocular vision or *stereopsis*. This feature allows our brain to combine the slightly different two-dimensional and horizontally shifted retinal images

from both eyes to “generate” a three-dimensional image and distinguish between any objects located in different planes along the viewing direction (Kaschke et al., 2014).

Human evolution determined the visual abilities we have today, which combined with the absorption, reflection, 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 400 nm (which we perceive as near ultraviolet and blue radiation) and 700 nm (which we perceive as red in color). 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 ( $\lambda$ ) band lying between  $\lambda$  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 preretinal 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, that is, 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, we would come to the conclusion that it could be delivered only by the astonishing amount of more than  $340 \times 10^6$  pixels. Here, the average angle for the human field of view in all directions of some 120 degrees, and 0.39° of arc being the acuity of a single photoreceptor are both respectively taken into account. 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 Ripp 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 min, a period that might be cut to no more than 10 min 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

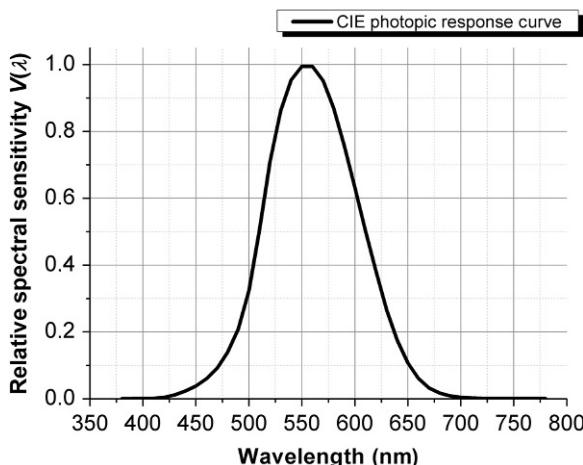
The universal stimulus specified for the study of human visual perception was white light. To define it, in 1900 Max Planck described the electromagnetic radiation of the so-called “black body” (Planck, 1900), that is, a radiation originated in a cavity in thermodynamic equilibrium with rigid opaque walls that might not be perfectly reflective at any wavelength, and that follows what is now called Planck’s law as expressed by Eq. (1.2). The radiation defined in Eq. (1.2) is assumed to be isotropic, homogeneous, unpolarized, and incoherent.

$$L(T) = \frac{2hc^2}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \quad (1.2)$$

In Eq. (1.2)  $L(T)$  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 meter or  $\text{W sr}^{-1} \text{ m}^{-2}$ ),  $T$  is the absolute temperature (in K),  $\lambda$  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  $101,325 \text{ N m}^{-2}$  (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  $540 \times 10^{12} \text{ Hz}$  and that has a radiant intensity in that direction of  $1/683 \text{ W sr}^{-1}$  (Kodak, 2008). One square centimeter of such a radiator defines the radiometric unit of candela (cd) which is equal to the luminous intensity radiated by it perpendicularly to its orifice. A hypothetical point-source of 1 cd emits a luminous flux of 1 lm (lm) per steradian. The flux density due to such a source at a distance of one meter is  $1 \text{ lm/m}^2$  or 1 lx.

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. 1.5. 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 color 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.



**Fig. 1.5** Photopic relative luminous sensitivity, as used by the CIE, based on data provided in Ripps and Weale (1976b) in Davson (1976).

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  $\Phi$  (measured in joules per second or watt), related to the photometric unit of lumen—the measure of luminous flux  $\Phi_v$ . In this sense, Eq. (1.3) (Kodak, 2008) defines a radiant flux  $\Phi$  emitted by a source, proportional to the area enclosed by its spectral (wavelength dependent) distribution curve  $\Phi(\lambda)$  which must be known.

$$\Phi = \int_0^{\infty} \Phi(\lambda) d\lambda \quad (1.3)$$

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  $\text{W/m}^2$ , related to the photometric unit of lux—a measure of illuminance  $E_v$
- the radiant intensity  $I_R$ , measured in  $\text{W sr}^{-1}$ , 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  $\text{W m}^{-2} \text{sr}^{-1}$  and related to the photometric entity of luminance  $L_v$ , measured in  $\text{cd/m}^2$ , among others.

To determine the capacity of the radiant flux  $\Phi$  to create the sensation of light, the  $\Phi(\lambda)$  curve from Eq. (1.3) 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. 1.5, and the maximum luminous efficacy factor  $K_m$  as expressed in the following equation (Kodak, 2008):

$$\Phi_v = K_m \cdot \int_{\lambda_{\min}}^{\lambda_{\max}} \Phi(\lambda) \cdot V(\lambda) d\lambda \quad (1.4)$$

For photopic vision, the maximum luminous efficacy of radiant flux as derived from its definition mentioned above is  $K_m = 683 \text{ lm/W}$ . In Eq. (1.4)  $\lambda_{\min}$  and  $\lambda_{\max}$  can be set to define the wavelength bandwidth in which the product of  $\Phi(\lambda)$  and  $V(\lambda)$  is nonzero. 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 photo-sensing technology performances in unforeseen manners.

Radiation detection far beyond the visible part of the spectra in both direction of X-rays and high-energy  $\gamma$  particles (photons) on the one side and millimeter wavelengths on the other, single-photon counting ability for extremely low radiation, sub-micrometer 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 the 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—quanta—which are proportional to the frequency of the radiation as expressed in Eq. (1.5), where  $E_{ph}$  is the energy of the emitted radiation,  $\nu$  its frequency, and  $h$  the proportionality constant named after Max Planck. The frequency of radiation could also be expressed in terms of its wavelength  $\lambda$ , as also shown in the following equation:

$$E_{ph} = h\nu = \frac{hc}{\lambda} \quad (1.5)$$

Based on Eq. (1.5), 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’s 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 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 to 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 a 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](#)). 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 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 toward 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 coproduct 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 forever.

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 it is the case 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 behavior) 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 a so-called long-wavelength cut off 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 of the light intensity.
- The photoelectric effect is a noninertial 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 is 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 photon emission.

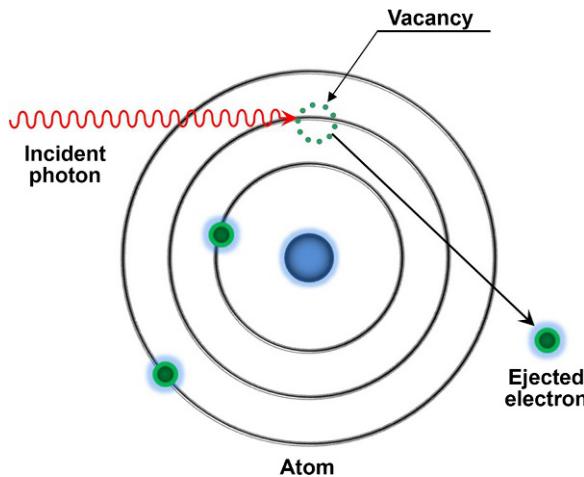
The mathematical description of the so defined external photoelectric effect is expressed by Eq. [\(1.6\)](#), where  $T$  is the kinetic energy of the ejected electron,  $h$  the Planck constant,  $\nu$  is the frequency of the incoming light, and  $W$ , the so-called work function of a certain material, is the minimum energy required to liberate a single electron to a position outside the solid surface.

$$T = h\nu - W \quad (1.6)$$

The work function  $W$  determines the threshold frequency of the external photoelectric effect as shown in Eq. [\(1.7\)](#), where  $\nu_{\min}$  is the minimal threshold frequency at which the photoelectric effect can still occur. This discovery earned Albert Einstein the Nobel Prize in Physics in 1921.

$$W = h\nu_{\min} \quad (1.7)$$

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 with it. One can clearly see that for the



**Fig. 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

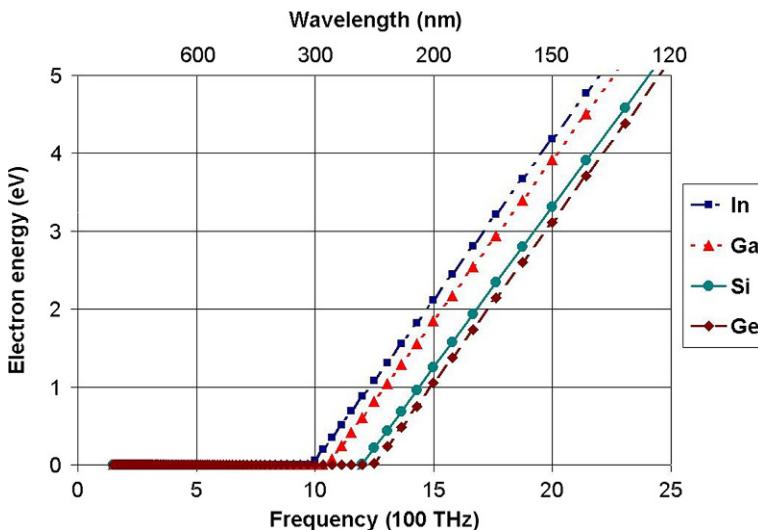
Material	Work function (eV)
Indium (In)	4.08
Gallium (Ga)	4.35
Silicon (Si)	4.95
Germanium (Ge)	5.15

Based on data from Hödlz, J., Schulte, F.K. 1979. Work function of metals. In Solid Surface Physics, vol. 85, Springer Tracts in Modern Physics, Springer-Verlag, Berlin Heidelberg, New York, pp. 1–50.

materials chosen no photoelectric effect can take place for frequencies of impinging radiation below 1000 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 V. This unit is widely used in particle physics due to its small value compared to the unit of the joule defined in SI, where  $1 \text{ eV} \approx 1.6 \times 10^{-19} \text{ J}$ .

According to the already introduced Rutherford-Bohr model, the individual electrons bound to the nucleus of a certain atom, for example, 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, for example, by absorbing incident radiation on them—which gives birth to the internal photoelectric effect—their behavior will be defined by a new continuum of corresponding energy levels. According to the theory developed by



**Fig. 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.

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 neighboring 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 to happen. An electron deficiency left in the covalent bond of the silicon crystal may be filled by one of the neighboring 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 (C/s) or amperes (A).

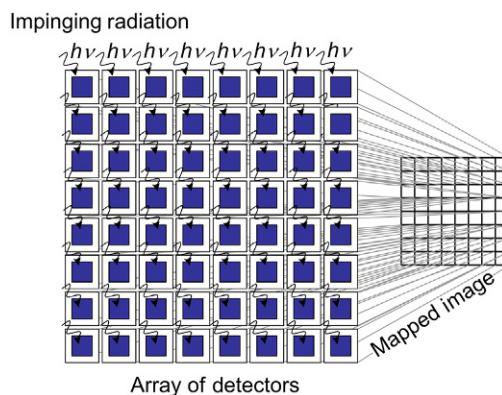
## 1.5 Early developments in photodetector technology

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

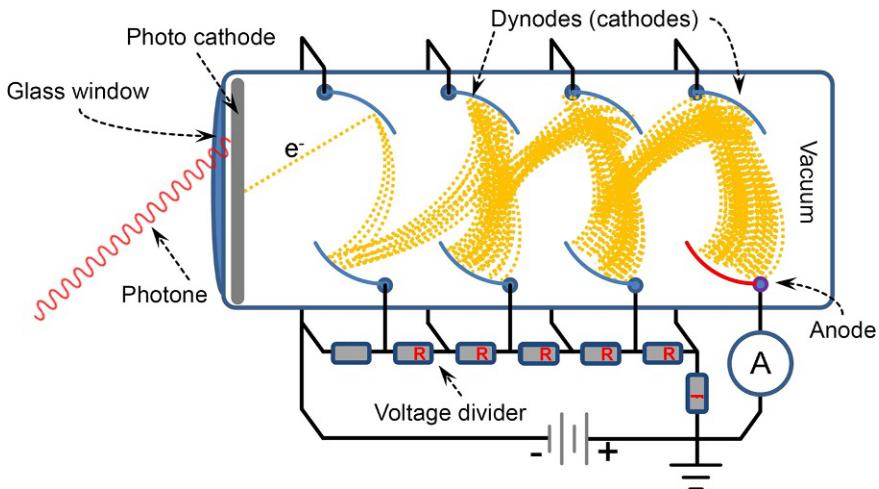
change of these electrical properties delivers indirectly the information regarding the energy of the impinging radiation itself. If a two-dimensional array of photosensitive elements is illuminated, and the information from every element is properly mapped on a two-dimensional surface, an image will 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.

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 ultralow 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<sub>2</sub>K<sub>3</sub>Sb, or K<sub>2</sub>CsSb ([Hamamatsu, 2007](#)) are used, results in the emission of electrons caused by the photoelectrical effect.



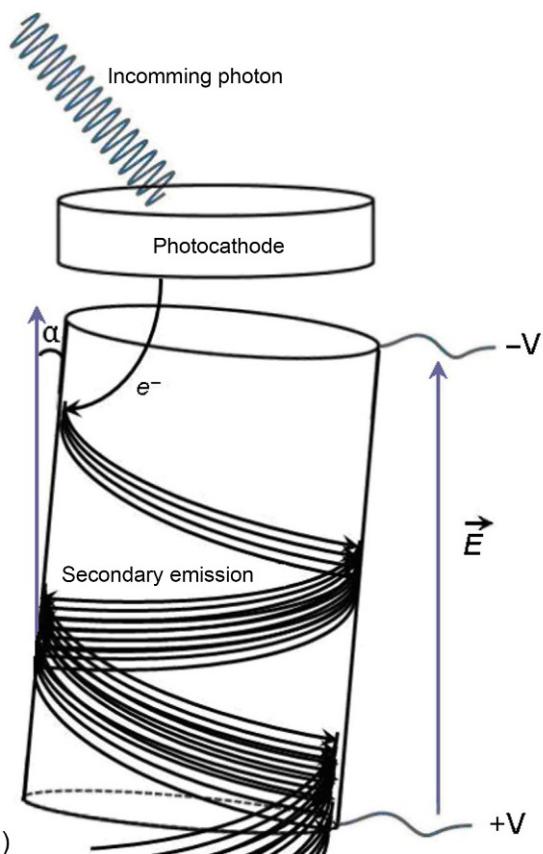
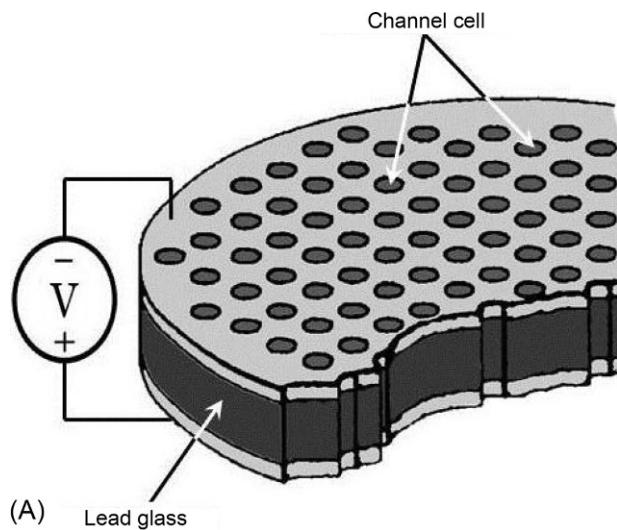
**Fig. 1.8** A graphic representation of the solid-state imaging principle using an array of photodetectors.



**Fig. 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, that is, 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 spatial resolution of several  $\text{mm}^2$  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



**Fig. 1.10** (A) A cutaway view of an MCP and (B) a cross section of the MPC cell causing the electron multiplication process due to secondary emission.

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 microchannel 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  $\alpha$  in Fig. 1.10B. This inclination simultaneously reduces the ion or radiation feedback within the micro-channels. At the first hit, the initial electron multiplication process takes place. The newly created electrons continue to travel across the induced vertical electrical field (see Fig. 1.10A) 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, for example, the aspect ratio of the cell (the ratio of the depth of each microchannel 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  $\mu\text{m}$ . Moreover, they allow the building of portable systems capable of ultralow 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 a very promising alternative for this task with silicon as the material of choice. This new generation of photosensors is discussed in the next chapter.

## 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. *Philos. Mag. Ser. 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., Jordan, P., 1925. Zur Quantenmechanik II. *Z. Phys.* 35, 557–615.
- CIE, 1983. *The Basis of Physical Photometry*, second ed. Commission International de L'Eclairage Publ. No. 18.2, Central Bureau of the CIE, Vienna.

- Cline, D., Hofstetter, H.W., Griffin, J.R., 1997. Dictionary of Visual Science, fourth ed. 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., Hendrickson, A.E., 1990. Human photoreceptor topography. *J. Comp. Neurol.* 292, 497–523.
- Davson, H., 1976. The Eye: Visual Function in Man, second ed. vol. 2A Academic Press, New York, p. 616.
- de Broglie, L., 1924. A tentative theory of light quanta. *Philos. 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. *Proc. R. Soc. London, Ser. 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. *Proc. Am. Assoc. Adv. Sci.*, 109–112.
- Einstein, A., 1905a. Zur Elektrodynamik bewegter Körper. *Ann. Phys.* 322 (10), 891–921.
- Einstein, A., 1905b. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Ann. Phys.* 17 (6), 132–148.
- Einstein, A., 1918. Dialog über Einwände gegen die Relativitätstheorie. *Naturwissenschaften* 6, 697–702.
- Hamamatsu, (Ed.), 2007. Photomultiplier Tubes. Basics and Applications, third ed. Hamamatsu Photonics, Hamamatsu. [http://sales.hamamatsu.com/assets/pdf/catsandguides/PMT\\_handbook\\_v3aE.pdf](http://sales.hamamatsu.com/assets/pdf/catsandguides/PMT_handbook_v3aE.pdf).
- Helmholtz, H., 1924. The Sensation of Vision: Treatise on Physiological Optics. vol. 2 Optical Society of America, Washington, DC (transl. J.P.C. Southall).
- Kaschke, M., Donnerhacke, K.H., Rill, M.S., 2014. Optical Devices in Ophthalmology and Optometry: Technology, Design Principles and Clinical Applications. Wiley-VCH Verlag GmbH & Co. KGaA, Germany, p. 638. ISBN: 978-3-527-41068-2.
- 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., Rechenberg, H., 1982. The Historical Development of Quantum Theory. Springer-Verlag, New York (Chapter 1).
- Michelson, A., Morley, E., 1887. On the relative motion of the Earth and the luminiferous ether. *Am. J. Sci.* 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. Verhandlungen der Deutschen Physikalischen Gesellschaft. Leipzig, In: Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum, pp. 237–245.
- Rayleigh, L., 1900. Remarks upon the law of complete radiation. *Philos. Mag.* 49, 539–540.
- Ripps, H., Weale, R.A., 1976b. Visual adaptation. In: Davson, H. (Ed.), The Eye: Visual Function in Man. vol. 2A. Academic Press Inc., New York, pp. 5–41. second ed.
- Rosenthal, E., Solomon, R.J., Johnson, C., 2004. Waves vs. Photons: A New View of How we Detect Light and Color. Creative Technologies LLC, United States.
- Rutherford, E., 1911. The scattering of  $\alpha$  and  $\beta$  particles by matter and the structure of the atom. *Philos. Mag.* 21 (6), 669–688.

- Schrödinger, E., 1926. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.* 28 (6), 1049–1070.
- Schultze, M., 1866. Zur Anatomie und Physiologie der Retina. *Arch. Mickrosc. Anat.* 2165–2286.
- Stoletov, M.A., 1888. On a kind of electrical current produced by ultra-violet rays. *Philos. Mag.* Ser. 5 26 (160), 317.
- Thomson, J.J., 1897. Cathode rays. *Philos. Mag.* 44, 293.
- Thomson, W., 1901. Lord Kelvin, nineteenth-century clouds over the dynamical theory of heat and light. *Philos. Mag.* Ser. 6 (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. *Philos. Mag.* Ser. 6 7 (39), 237.

## Further reading

- Hölzl, J., Schulte, F.K., 1979. ‘Work function of metals’ in Solid Surface Physics, Springer Tracts in Modern Physics. vol. 85 Springer-Verlag, Berlin Heidelberg, New York, pp. 1–50.
- Ripps, H., Weale, R.A., 1976a. The visual photoreceptors. In: Davson, H. (Ed.), The Eye: Visual Function in Man. vol. 2A. Academic Press Inc., New York, pp. 101–129. second ed.

# Operational principles of silicon image sensors

2

D. Durini<sup>a</sup>, D. Arutinov<sup>b</sup>

<sup>a</sup>National Institute of Astrophysics, Optics and Electronics (INAOE), Tonantzintla, Puebla, Mexico, <sup>b</sup>Central Institute of Engineering, Electronics and Analytics ZEA-2—Electronic Systems, Forschungszentrum Jülich GmbH, Jülich, Germany

*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 (CCDs) 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 three-dimensional (3D) technologies have been developed to address these challenges, and what technical hurdles still need to be addressed.

## 2.2 Silicon phototransduction

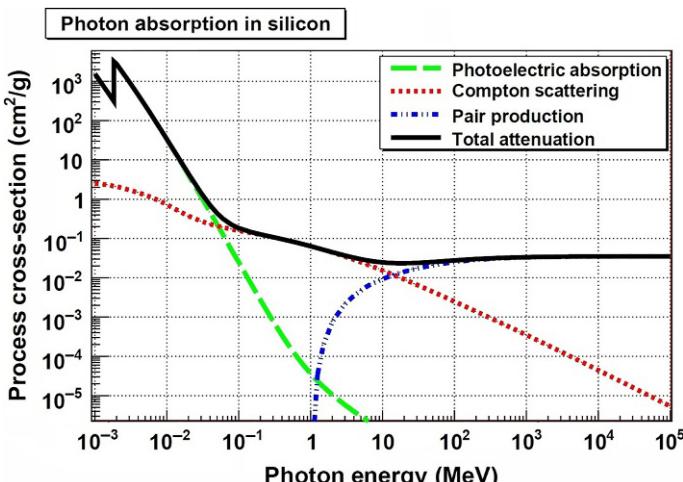
Silicon ( $\text{Si}^{14}$ ) is a fourth group element of the periodic table that forms, if in solid state, a diamond lattice crystal structure. In solid state, it presents covalent bonds of its four valence electrons oscillating around the nucleus in the  $[\text{Ne}]3s^13p^2$  orbits (Sze, 2002). The 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 fulfill 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 centimeter 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, for example, 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



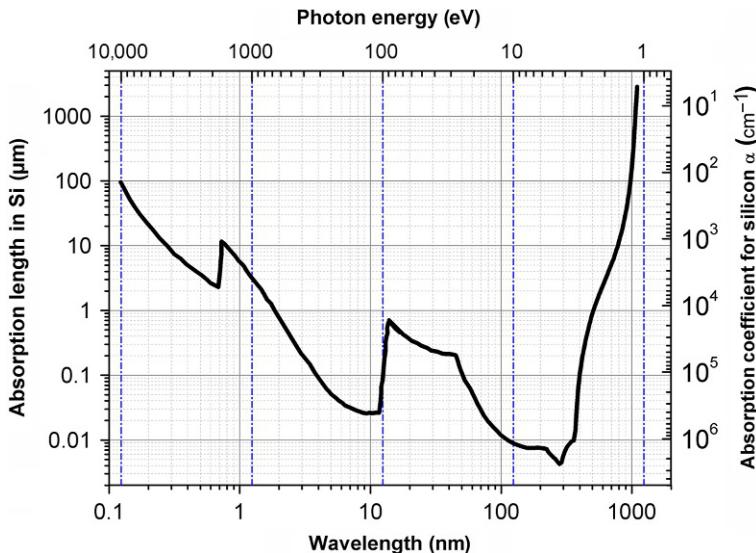
**Fig. 2.1** Process total cross sections as a function of energy of the impinging radiation in silicon, showing the contributions of the three most common processes: photoelectric absorption, Compton scattering, and pair production (Spieler, 1998).

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 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 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 indirect 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 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



**Fig. 2.2** Wavelength ( $\lambda$  in nm) and photon energy (in eV)-dependent impinging radiation absorption depth (in  $\mu\text{m}$ ), obtained as the inverse of the absorption coefficient ( $\alpha$ ) in silicon. Adapted from data taken from Theuwissen, A.J.P., 1996. Solid-State Imaging With Charge-Coupled Devices, Kluwer Academic Publishers, Alphen van den Rijn and Green, M.A., Keevers, M., 1995. Optical properties of intrinsic silicon at 300 K. Prog. Photovolt. 3 (3), 189–192.

*k*-values. Phonons, quantized 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  $\Phi_{ph}$  (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  $\Delta z$  within the silicon crystal is given by Eq. (2.1) (Sze, 2002), where  $\alpha$  is a proportionality constant defined as the absorption coefficient.

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

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)

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

Considering the photon energy (see Eq. 1.4) and  $\Phi_{ph}$ , 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  $\alpha$ , 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  $1/e$ , that is, to  $0.37 \cdot \Phi(z)$ . 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, for example, 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  $i$ th 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  $i$ th 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 fulfillment of the momentum conservation law. If the impinging photon energy reaches the binding energy of the electron in the  $i$ th 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, for example, for energies between  $E_{ph} = 3$  keV and  $E_{ph} = 15$  keV, the silicon depth should be 10  $\mu\text{m}$  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. This 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 if surrounded by majority carriers is the so-called minority carriers’ recombination time (or lifetime):  $\tau_p$  for holes transiting an *n*-type region, or  $\tau_n$  for electrons transition through a *p*-type region (Schroder, 1998). The distance of 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  $D_p$  being the charge carrier mobility ( $\mu_n$  and  $\mu_p$ )-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).

It is important to mention that “excess” EHPs 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 centers) 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 photon caused generation 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 photon caused generation 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 centimeter ([Schroder, 1998](#)); the same mechanisms that can be found in the presence of any mechanically altered silicon region, for example, the walls of shallow trench isolation (STI—a feature of the integrated circuit (IC) 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 tunneling 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 tunneling 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,  $\mathfrak{R}$ , 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  $\eta$  of the detector.  $\eta$  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 is described below.

## 2.3 Principles of CCD and 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

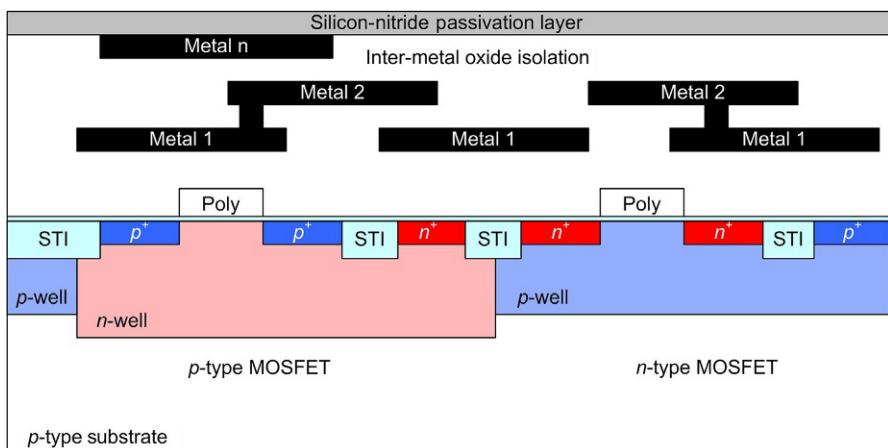
invention appeared in *Bell Systems Technical Journal* in April 1970. In the same issue, a technical paper appeared in which Gilbert Amelio, Michael Tompsett, and the same George E. Smith ([Amelio et al., 1970](#)), experimentally verified the charge transfer principle between the CCD neighboring cells 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 metal-oxide-semiconductor capacitors (MOS-C) 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 neighboring 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 (CTE) 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 can be used for photodetection applications. The invention and its application to photodetection tasks brought Willard Boyle and George E. Smith the Nobel Prize in Physics in 2009. The CCD technology will be thoroughly described in [Chapter 3](#).

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  $50 \times 50$  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 United Kingdom ([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 (ADCs), 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 photo-detection 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 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



**Fig. 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.

*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 from 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, while the readout circuitry is fabricated in a separate 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  $\mu\text{m}$  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 is 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 \(1928\)](#) 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  $\text{SiO}_2$  (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

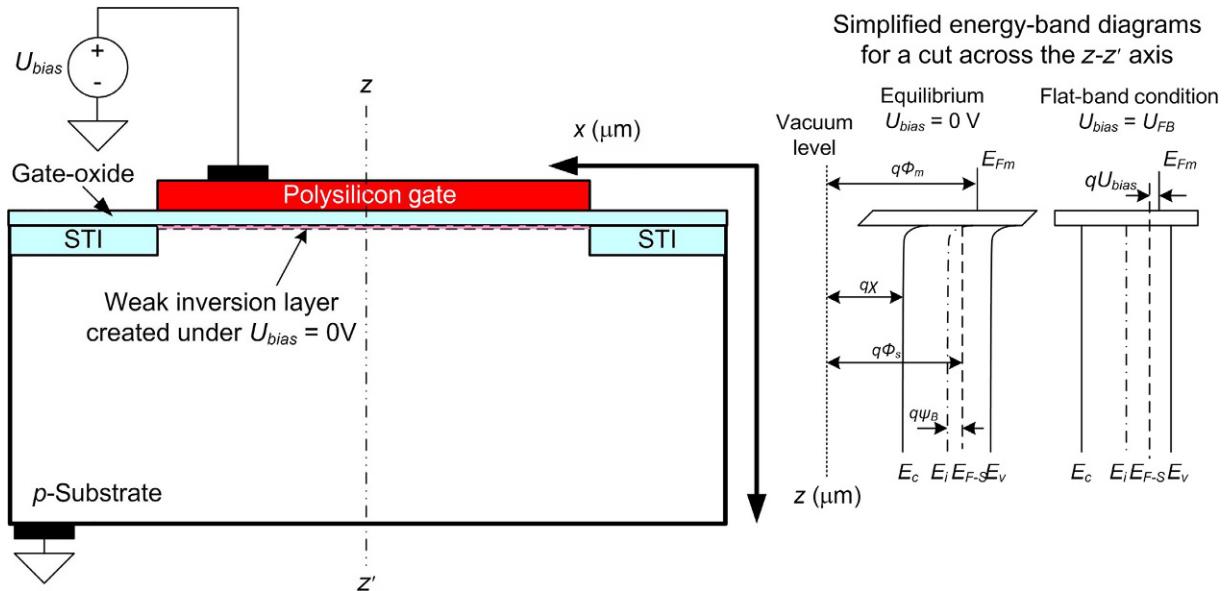
In silicon-based planar-technologies, *n*-doped silicon regions are normally fabricated by ion implantation of phosphorus (P) or arsenic (As) atoms into a *p*-type—previously boron (B) doped—silicon substrate. Alternatively, *p*-regions are formed by ion implantation of boron atoms into an *n*-type—previously P or As doped—silicon. This is followed by a high-temperature impurity activation process in which the donor, for

example, P or As which form part of the fifth group of elements (and have five valence electrons), or acceptor impurities such as B which forms part of the third 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 neighboring 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 neighboring 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<sub>2</sub>-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 much less than 10 nm nowadays (in modern technologies also other elements with high dielectric constants, the so-called “high-k” materials, are sometimes used instead), has been fabricated, it gets covered by a polysilicon layer of some hundreds of nanometers 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  $q(\Phi_m - \Phi_s)$ , that is, 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  $U_{FB} < 0$  V, 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). Fig. 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  $\Phi_m$ , the silicon work function  $\Phi_s$ , the electron affinity  $\chi$ , 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  $U_{bias} > 0$  V is applied to the gate, as shown in Fig. 2.4, the holes are pushed away from the interface and a negative space-charge region (SCR), also known as depletion region, gets formed by the ionized acceptors. The bands bend downwards and the concentration of holes decreases at the silicon-oxide



**Fig. 2.4** Technology cross section and band diagram of a MOS-C structure fabricated on  $p$ -type silicon substrate for a nonideal system in equilibrium, where  $U_{bias} = 0\text{ V}$  and  $\Phi_m > \Phi_s$ , and the flat-band condition where  $U_{bias} = U_{FB}$  and  $\Phi_m = \Phi_s$ .

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) of surface charge density  $Q_B$ , the silicon surface electric field  $E_s$ , and the electrostatic potential  $\psi_s$ , whose dependence of the SCR represents the bending of the bands at the silicon-oxide interface, can be respectively, expressed through Eqs. (2.3)–(2.5) (Sze, 2002), for  $q$  the fundamental charge,  $N_A$  the acceptor concentration density in silicon,  $W_{SCR}$  the SCR width,  $\epsilon_0$  the permittivity in vacuum, and  $\epsilon_{Si}$  the silicon dielectric constant.

$$Q_B = qN_A W_{SCR} \quad (2.3)$$

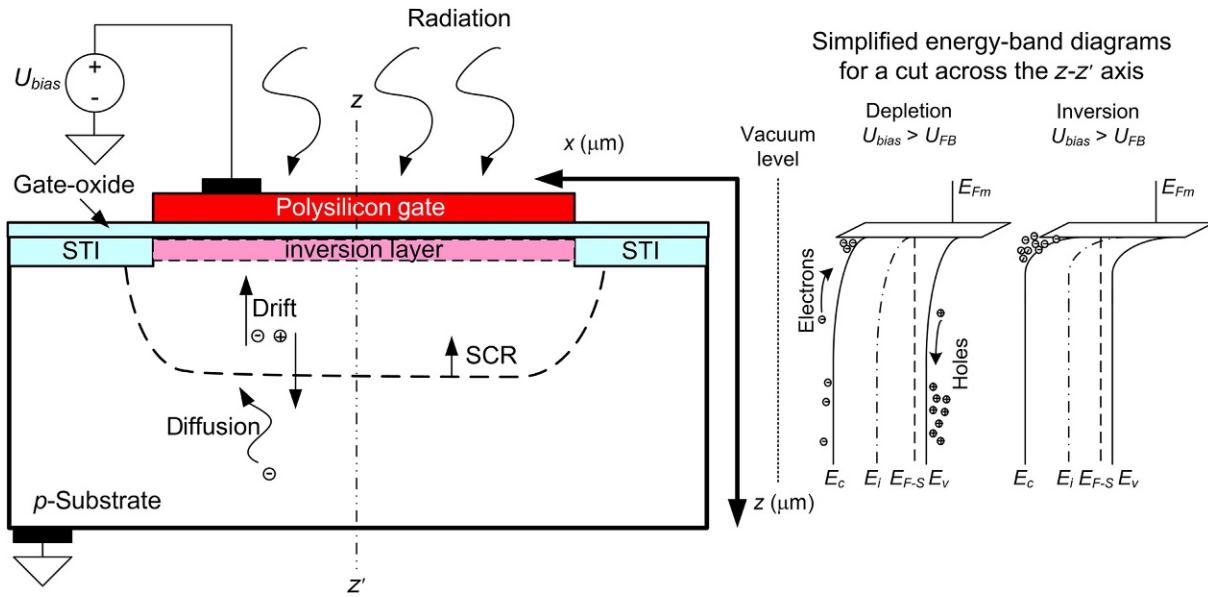
$$E_s = \frac{qN_A W_{SCR}}{\epsilon_0 \epsilon_{Si}} \quad (2.4)$$

$$\psi_s = \frac{qN_A W_{SCR}^2}{2\epsilon_0 \epsilon_{Si}} \quad (2.5)$$

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  $\epsilon_{ox}$  is the oxide dielectric constant and  $d_{ox}$  is the oxide thickness.

$$W_{SCR} = \sqrt{\frac{\epsilon_0 \epsilon_{Si}}{qN_A} (U_{bias} + U_{FB}) + \left(\frac{\epsilon_{Si}}{\epsilon_{ox}} d_{ox}\right)^2} - \frac{\epsilon_{Si}}{\epsilon_{ox}} d_{ox} \quad (2.6)$$

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, that is, 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, that is, 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.



**Fig. 2.5** Technology cross-section and band diagram of a MOS-C structure fabricated on *p*-type silicon substrate for a nonideal system under illumination in deep-depletion and inversion.

Therefore, a two-dimensional (2D) 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.

During this nonequilibrium process the intrinsic level at the interface reaches and eventually crosses through the nonequilibrium state created quasi-Fermi level in silicon, that is,  $\psi_s \leq \psi_B$ . Here, the situation is that of the majority electrons being accumulated 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, that is,  $\psi_s = 2\psi_B + 4k_B T/q$ . In this case, the system has reached a new equilibrium condition. Under illumination, once the system is again out of equilibrium, the amount of collected charge forming this inversion layer corresponds to the amount of charge, photogenerated and thermally generated 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).

$$W_{inv} = \sqrt{\frac{4\epsilon_0\epsilon_{Si}\psi_s}{qN_A}} \quad (2.7)$$

$$E_s = \frac{qN_A}{\epsilon_0\epsilon_{Si}} W_{inv} \quad (2.8)$$

$$U_{th} = U_{FB} + 2\psi_B + E_{ox}d_{ox} = U_{FB} + 2\psi_B + \frac{d_{ox}}{\epsilon_0\epsilon_{ox}} \sqrt{4qN_A\epsilon_0\epsilon_{Si}\psi_B} \quad (2.9)$$

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. 2.9B 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 centers, 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 CTE 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, and consequently the minimum of the electrostatic potential to which the electrons are being attracted, is 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  $30 \text{ e}^-_{\text{rms}}$  and a 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 mainly due to the possibility to fabricate them in standard CMOS technology without any extra fabrication steps.

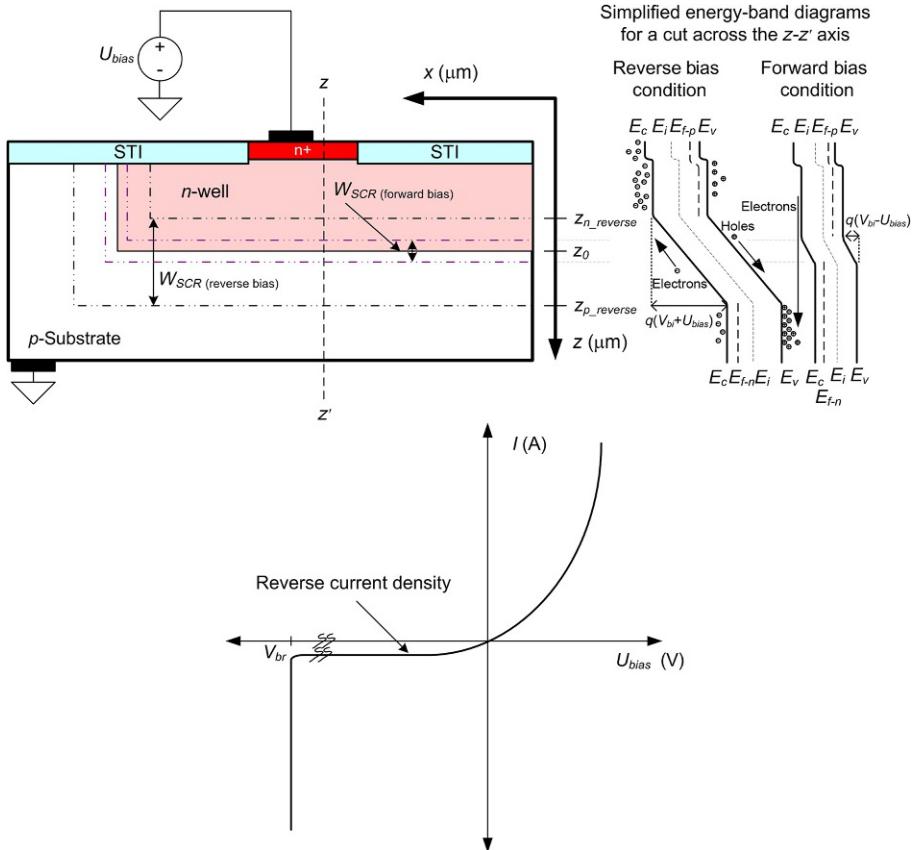
The *p-n* junction-based photodetectors proved nevertheless much more efficient than their MOS-C-based counterparts from the very beginning, and were incorporated into the so-called interlined CCD technology very early and continued to be used in all existing applications until today, as will be explained 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 positive charged mobile holes continue to leave the *p*-side, negative charged acceptors near the junction are left uncompensated due to their inability to move. Similarly, some of the positive charged donors are left uncompensated once the negative charged mobile electrons start leaving the *n*-side of the junction. This causes a negative space charge (acceptors) to build up at the (positive charged) *p*-side of the junction and the positive space charge (donors) to build up on the opposite (negative charged) side. This SCR creates an electric field that is directed from the positive charge toward the negative charge in the direction opposite to the initial condition. 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  $\Phi$  is the induced electrostatic potential.

$$\begin{aligned} E(z) &= -\frac{d\varphi}{dz} = -\frac{qN_A(z+z_p)}{\epsilon_0\epsilon_{Si}}, \quad \text{for } -z_p \leq z \leq 0 \\ E(z) &= \frac{d\varphi}{dz} = \frac{qN_D(z-z_n)}{\epsilon_0\epsilon_{Si}}, \quad \text{for } 0 \leq z \leq z_n \end{aligned} \tag{2.10}$$

At this point, an electric current starts to flow opposite to the direction of the original carrier diffusion currents, that is, the generated electrical field causes a drift current of each carrier type to flow opposed to its diffusion 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



**Fig. 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.

is equal to the total charge inside, the unique space charge distribution and the electrostatic potential  $\Phi$  are given by Eq. (2.11), the so-called Poisson equation for  $p$ - $n$  junctions, where  $\rho_s$  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  $n \gg p$ .

$$\frac{d^2\varphi}{dz^2} \equiv -\frac{dE}{dz} = -\frac{\rho_s}{\epsilon_0 \epsilon_{Si}} = -\frac{q}{\epsilon_0 \epsilon_{Si}} (N_D - N_A + p - n) \quad (2.11)$$

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  $U_{bias} = 0$  V, and the drift currents equal the diffusion ones.

$$W_{SCR} = z_n + z_p = \sqrt{\frac{2(V_{bi} + U_{bias}) \cdot \epsilon_0 \cdot \epsilon_{Si}}{q} \left( \frac{N_A + N_D}{N_A N_D} \right)} \quad (2.12)$$

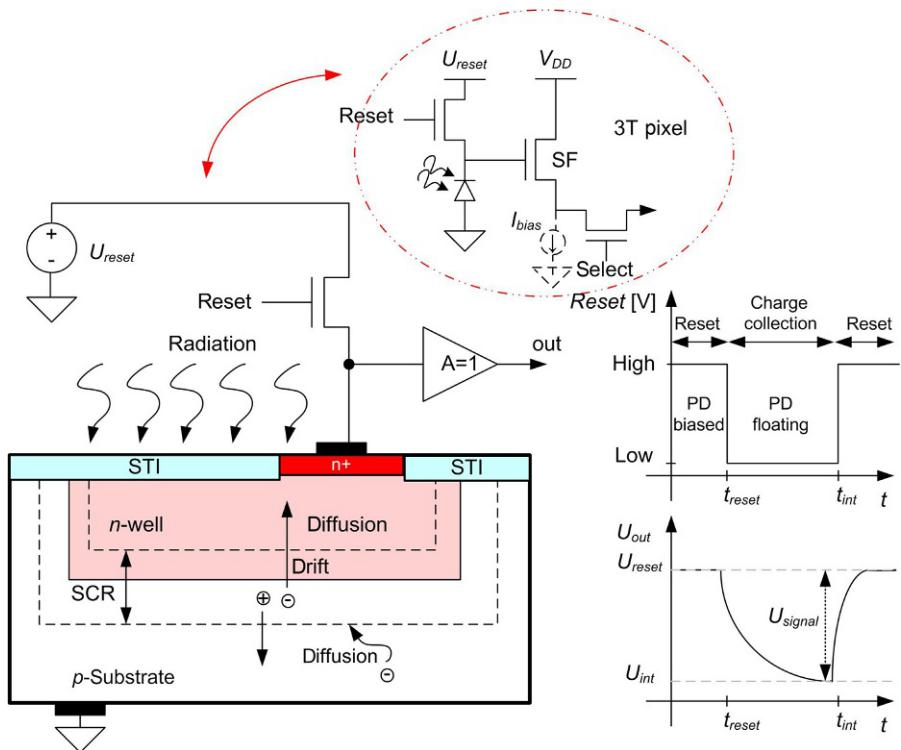
The total variation of the electrostatic potential across the junction, the so-called built-in voltage  $V_{bi}$ , is also defined by the doping concentrations as expressed in Eq. (2.13) (Sze, 2002), where  $k_B$  is the Boltzmann constant,  $T$  temperature, and  $n_i$  the intrinsic carrier concentration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

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

One way to alternate this thermal (dynamic) equilibrium of the *p-n* junction is to apply a bias voltage to it. The 2D 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 (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 of an electron has to occupy this energy level, and  $E_{F-p}$  and  $E_{F-n}$  the doping concentration-dependent new quasi Fermi levels for the *p*-type region and *n*-type region, respectively, valid during the time the junction remains out of equilibrium. 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, that is, 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



**Fig. 2.7** Schematic representation of 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.

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  $U_{out}$  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  $C_{SN}$ .

$$\begin{aligned}
 C_{SN} &= \frac{\epsilon_0 \epsilon_{Si}}{W_{SCR}} = \sqrt{\frac{2(V_{bi} + U_{bias}) \cdot \epsilon_0 \cdot \epsilon_{Si}}{q} \left( \frac{N_A + N_D}{N_A N_D} \right)} \\
 &= \sqrt{\frac{q \cdot \epsilon_0 \cdot \epsilon_{Si}}{2(V_{bi} + U_{bias}) \left( \frac{N_A + N_D}{N_A N_D} \right)}}
 \end{aligned} \tag{2.14}$$

$$U_{signal}(t) = U_{reset} - \frac{Q_{signal}(t)}{C_{SN}} \quad (2.15)$$

## 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](#)) inspired by the work performed by [Noble \(1968\)](#). To accomplish this, they introduced what they called a “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 on  $p$ -substrate-based PD is by using a MOSFET. This allows for charging the PD to a defined  $U_{reset}$  during the reset operation, and else leaving the PD floating, that is electrically isolated from all other nodes, during the charge collection phase—which is accomplished by shutting the gate voltage of the reset MOSFET off. This functionality 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 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 is 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 \(1968\)](#). An additional select switch (here also a MOSFET can be used) is incorporated into the pixel to be able to select it for readout if forming part of a matrix, just as proposed by [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  $128 \times 128$  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 Jr., 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 of the measured entity  $x$ .

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

$$\sigma^2 = (\bar{x} - \bar{\bar{x}})^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 p(x) dx \quad (2.16)$$

The square root of the variance is called a “standard deviation” of the signal and is often represented by the letter  $\sigma$ . For voltage or current signals,  $\sigma$  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  $S_v(f)\Delta f$  is interpreted as being the amount of noise voltage contained in the frequency band defined by  $f$  and  $f + \Delta f$ , where  $\Delta f$  is small. Both dependencies are expressed in Eq. (2.17) (Leach, 1994).

$$\begin{aligned} \bar{v}_n^2 &= \frac{1}{T} \int_{-T/2}^{T/2} v_n^2(t) dt \\ \bar{v}_n^2 &= \int_0^{\infty} S_v(f) df \end{aligned} \quad (2.17)$$

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 (1928) of American Telephone and Telegraph Company (AT&T) in 1927 (since 1934 called Bell Telephone Laboratories), and theoretically analyzed 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 generated by this impedance in the frequency band  $\Delta f$  is given by Eq. (2.16) (Nyquist, 1928). Because  $Z$  is a function of frequency,  $\Delta f$  must be small enough so that the impedance real part  $Re(Z)$  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 in  $\Delta f$  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.

$$\frac{\overline{v_{n\_thermal}^2}}{\overline{i_{n\_thermal}^2}} = \frac{4k_B T \text{Re}(Z) \Delta f}{\text{Re}(Z)} \quad (2.18)$$

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 modeled as a noise current flowing in parallel to the measuring device. In general, the mean-square flicker noise current in the frequency band  $\Delta f$  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 modeled, 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  $1/f$  behavior.

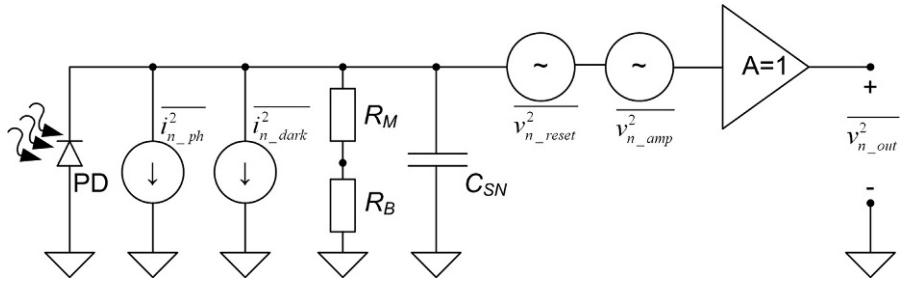
$$\begin{aligned} \overline{i_{n\_flic\_ker}^2} &= \frac{K_f I_{DC}^m \Delta f}{f^n} \\ S_i(f) &= \frac{\overline{i_{n\_flic\_ker}^2}}{\Delta f} = \frac{K_f I_{DC}^m}{f^n} \end{aligned} \quad (2.19)$$

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 is the output signal voltage power and is the mean-square noise output voltage.

$$SNR = 10 \log \left( \frac{v_{so}^2}{v_{no}^2} \right) \quad (2.20)$$

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 in Fig. 2.8. Fig. 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 thermodynamical uncertainty of the charge on a capacitor.

In Fig. 2.8, the complex impedance  $Z$ , expressed by (2.21), consists of the pixel SN 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  $R_M$ .



**Fig. 2.8** Schematic representation of the noise sources present in a PD 3T pixel depicted in Fig. 2.7.

$$Z = R_S \left| \left( \frac{1}{j\omega C_{SN}} \right) \right| = \frac{R_S}{1 + j\omega C_{SN} R_S} = \frac{R_S}{1 + (\omega R_S C_{SN})^2} - j \frac{\omega R_S^2 C_{SN}}{1 + (\omega R_S C_{SN})^2} \quad (2.21)$$

Considering Eqs (2.17), (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  $\xi$  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  $1 - e^{-\frac{t}{R_B C_{SN}}}$  by multiplying the result of Eq. (2.22).

$$\begin{aligned} \overline{v_{n\_kTC}^2} &= \xi \int_0^\infty 4k_B T \operatorname{Re}(Z) df = \xi \int_0^\infty \frac{4k_B T R_M}{1 + (2\pi f R_M C_{SN})^2} df \\ &= \xi \frac{2k_B T}{\pi C_{SN}} \int_0^\infty \frac{dx}{1 + x^2} = \xi \cdot \frac{k_B T}{C_{SN}} \end{aligned} \quad (2.22)$$

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  $\overline{v_n^2 kTC} = Q_n^2 / C_{SN}^2$ , then the ENC, due to  $k_B T/C$ , can be expressed as  $ENC_{k_B T/C} = \frac{\sqrt{\xi \cdot k_B T C_{SN}}}{q}$ .

The “partition noise,” closely related to the  $k_B T/C$  noise used to be embedded in the reset noise model with a scaling factor  $\alpha_{part}$  applied to the conventional thermal noise source (Lai and Nathan, 2005). The  $\alpha_{part}$  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  $\alpha_{part} = 1.3$ , 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).

$$\begin{aligned}\overline{v_{n\_par}^2} &= \frac{2\alpha_{part}(k_B T C_{G\_RST})}{\pi^2 C_{SN}^2} \\ ENC_{part} &= \frac{\sqrt{2\alpha_{part} \cdot k_B T C_{G\_RST}}}{q\pi}\end{aligned}\tag{2.23}$$

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  $Q_{signal}$ , that is, in a nondestructive 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  $\sigma_{SF}^2$  of the SF output signal or its ENC, where  $\overline{v_{n\_SF}^2}$  is the input-referred mean-square source-follower noise voltage.

$$\begin{aligned}\sigma_{SF}^2 &= \frac{C_{SN}^2 \cdot \overline{v_{n\_SF}^2}}{q^2} \\ ENC_{SF} &= C_{SN} \cdot \frac{\sqrt{\overline{v_{n\_SF}^2}}}{q}\end{aligned}\tag{2.24}$$

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  $\sigma_{dark}^2$ . Thus, the number of thermally generated electrons fluctuates about  $\overline{n_{dark}}$  with a standard deviation of  $\sigma_{dark} = \sqrt{\overline{n_{dark}}}$ .

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

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, that is,  $\sigma_{ph}^2 = \overline{n_{ph}}$ . 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  $T_{int}$ ,  $\sigma_{ph}^2$ , is defined by Eq. (2.26).

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

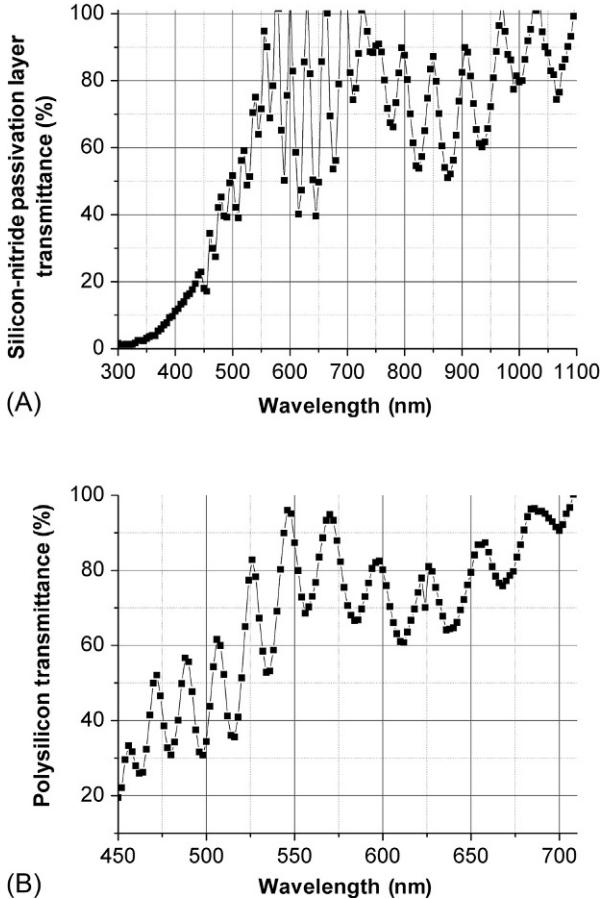
Another type of phenomena also related to the amount of photons actually impinging the silicon photoactive regions are the reflection and absorption characteristics of the layers deposited on top of silicon in every photodetector array. If any electrical signals are to be carried to the photodetector (e.g., digital signals required for the reset transistor,  $U_{bias}$  or  $V_{DD}$ ) or out of it (e.g.,  $U_{out}$ ), 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  $0.35\text{ }\mu\text{m}$  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. 2.9A (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 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. 2.9B, 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 the detection of light in the UV part of the spectra with wavelengths between 100 and 450 nm.

Thinning the silicon substrates to below  $50\text{ }\mu\text{m}$  to achieve back-side illumination (BSI) 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.

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  $n_{ph}$  at the end of  $T_{int}$ , as shown in Eq. (2.27) (Durini, 2009).

**Fig. 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.



$$\begin{aligned}
 SNR &= 10 \cdot \log \left( \frac{\overline{n_{ph}}^2}{\sigma_{dark}^2 + \sigma_{ph}^2 + \sigma_{kTC}^2 + \sigma_{part}^2 + \sigma_{amp}^2} \right) = \\
 &= 10 \cdot \log \left( \frac{\overline{n_{ph}}^2}{\overline{n_{dark}} + \overline{n_{ph}} + \frac{k_B T C_{SN}}{q^2} + \frac{2\alpha_{part} \cdot k_B T C_{G\_RST}}{\pi^2 q^2} + \frac{C_{SN}^2}{q^2} \cdot \frac{v_{n\_amp}^2}{\overline{n_{amp}}}} \right)
 \end{aligned} \tag{2.27}$$

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, for example, 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 antireflection coatings (ARC) or even completely removing the inter-metal isolation stack on top of the photoactive area 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  $\mu\text{m}$  should be completely depleted of majority charge carriers.

- *Decrease the dark current and the  $\bar{n}_{dark}$ :* the “dangling” bonds on the silicon surface or on the walls of neighboring STI structures should be at all times satisfied in order to reduce their generation rates and at best “screen” their noise contributions. This can be achieved by using highly doped “burying” or “passivation” layers diffused on these regions and keeping the stored signal charge away from these regions rich in SRH generation-recombination centers. The SCR reaching these regions should also be avoided. Another strategy that could additionally be followed is to increase 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  $8^\circ\text{C}$ , 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 occupied 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 (FWC), 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 (also called *double delta sampling* DDS) can be performed, where the output signal related to one  $T_{int\_1}$  is subtracted from the reset signal from the following  $T_{int\_2}$ . The assumption here is that the reset

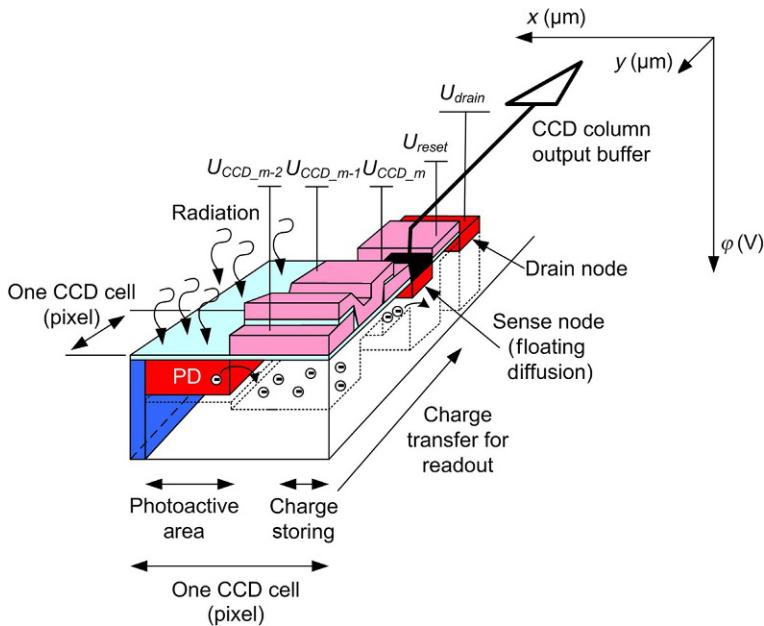
function-related noise components remain similar between two reset operations under similar operating conditions, which is unfortunately not always the case.

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](#). 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  $0.7 \text{ e}^-_{\text{rms}}$  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 1288, 2013](#)), have been reported in the 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 (FD) SN, the reset gate biased through  $U_{\text{reset}}$ , and the drain diffusion (forming all together the reset transistor or the “digital signal regeneration stage” as proposed by [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 FWC 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

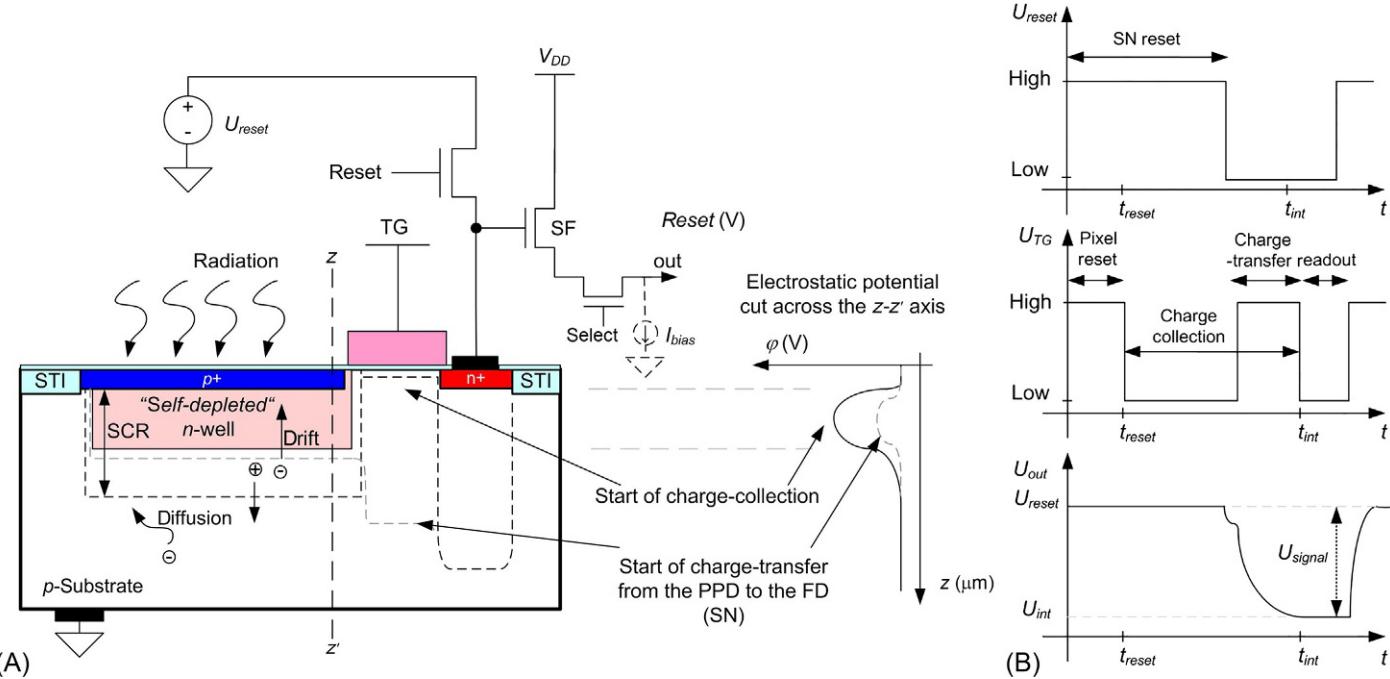


**Fig. 2.10** Schematic representation of technology cross-section of an interline CCD cell (pixel).

the majority carrier. This measure lowered additionally 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 4 T active pixel structure (because of the four transistors in the pixel structure: the reset transistor, the select transistor, the source-follower main transistor SF, and the transfer-gate TG). This 4T active pixel structure, first 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), is 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. The high concentration of



**Fig. 2.11** (A) Schematic representation of technology cross-section of a  $p^+$   $np$  junction-based pinned photodiode 4 T 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.

holes in this layer additionally boosts the recombination rate of electrons thermally generated at the silicon-oxide interface. The self-depleted  $n$ -well remains in this structure isolated from most of the SRH generation-recombination sources. To ensure this, a certain minimal distance should be additionally incorporated between the  $n$ -well and the STI walls, which should at best also be covered with  $p^+$  layers. The polysilicon TG is used to electrically decouple the PPD from the SN FD.

As can be observed in the time-diagrams depicted in Fig. 2.11B, 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 FD and  $U_{TG-high}$  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 fulfill 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 near the TG. Once they get there, they will be equally drifted and transferred to  $U_{reset}$  sustaining the 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. 2.11B.

For the charge-collection operation, the FD remains in reset mode while the photoactive region of the pixel gets electrically decoupled from the SN by shutting the TG 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 positive 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. 2.11A. 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. To avoid losing the electrons photogenerated in this region due to recombination mechanisms the  $p^+$ -pinning layer should be made as thin as possible. 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. 2.11A). This will increase the FWC of the PPD.

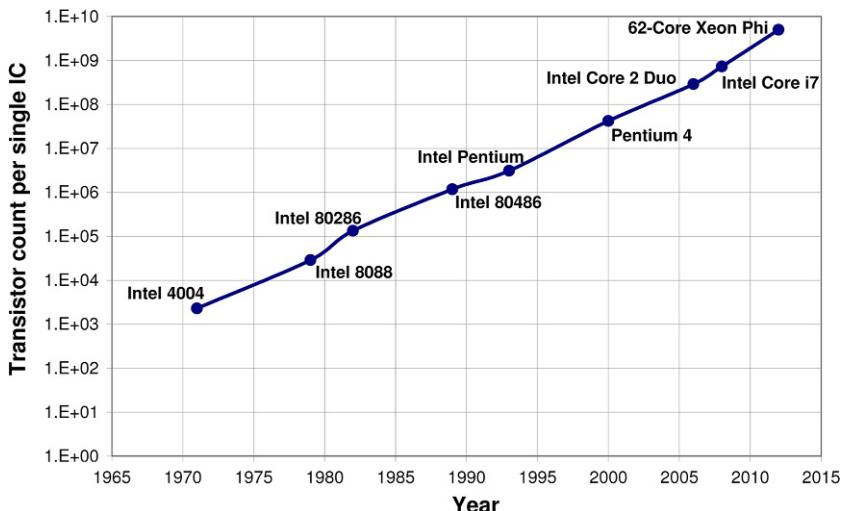
At a certain point defined by the image sensor frame rate, the reset transistor is shut OFF, and the FD is left floating. Once enough time has passed for the reset transistor to get settled and if the correlated double sampling operation is to be performed, at this point the reset pixel output signal is to be read out through the SF stage and saved somewhere at a capacitor located 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 TG (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 SN enables charge-collection and simultaneous pixel nondestructive readout, that is (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  $k_B T_C$ -noise, and desired pixel 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 and increase the charge-to-voltage conversion factor of the pixel, tailoring 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, it's 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. This pixel structure profited from most of the just mentioned advantages, but still suffered 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



**Fig. 2.12** Transistor count of the Intel processors versus production year. Note the logarithmic scale of the vertical axis.

Data from Intel Corporation, 2013. Laptop, Desktop, Server and Embedded Processor Technology—Intel (online). <http://www.intel.com>.

market. A well-known trend described by [Moore \(1965\)](#), who is the cofounder of Intel, states that the number of transistors per unit area of an IC roughly doubles every 2 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 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, for example, those with minimum feature sizes below 90 nm, the gate-oxide thickness is in the order of 1 nm, 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 and 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 analog or transduction functionalities in ICs 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 microchannel plates (MCPs), photomultiplier tubes (PMTs), 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 suboptimum photo-detection 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 are only a handful of CMOS foundries nowadays (Aptina Imaging, acquired by ON Semiconductor in August of Eckardt, 2014, Taiwan Semiconductor Manufacturing Company Ltd. (TSMC), Tower Jazz, ST Microelectronics, 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  $\mu\text{m}$ . 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  $e^{-}_{\text{rms}}$ , 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 [Chapters 10 and 15](#)), biology and medicine (as described in [Chapters 12 and 13](#)), 3D imaging and ranging that is enabling a vast amount of possibilities in the entertainment and autonomous driving markets (as described in [Chapter 11](#)) as well as many other industrial and scientific applications, or even broadcast and HDTV as explained in [Chapter 14](#), 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 13 that deal with BSI 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 (ROICs) in integrated solutions.

## 2.9 Single-photon counting

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 are 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, are also an attractive goal. Nevertheless, a 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 having the performance and not necessarily the cost as the driving element for further developments.

Near single-photon counting with nanosecond and even picosecond time resolution has been one of the main breakthroughs of the last decade. This was achieved in the form of single-photon counting avalanche diodes (SPADs) as proposed by [Cova et al. \(1981\)](#), based on the pioneering work of [Haitz \(1965\)](#), with biasing voltages well above the  $p-n$  junction breakdown voltage  $V_{br}$  and special quenching circuits. Avalanche quenching is a mechanism especially implemented into the SPAD pixels to detect and immediately quench the avalanche process of the SPAD cell in order to avoid damaging the device due to the exponential increase of the SPAD current caused by the electron impact ionization mechanism. The latter can be accomplished passively by adding an additional quenching resistor to the structure, or actively by including an active quenching circuit that can detect the increase in the anode current

and quench the process in a more efficient way if compared to passive quenching ([Bronzi et al., 2013](#)). This second approach has been followed for development of SPAD arrays with pixel independent readout mechanisms that can be used in time-correlated single-photon counting (TC-SPC) measurements as those used in 3D imaging and ranging that uses pulsed active illumination in the so-called time-of-flight (ToF) approach, or many applications in biology or medicine such as fluorescence lifetime imaging (FLIM), to name just a couple of examples.

SPAD arrays are realized nowadays in advanced CMOS technologies ranging from 0.35  $\mu\text{m}$  CMOS ([Bronzi et al., 2013](#)), over 180 nm CMOS mainly provided by ST Microelectronics ([Zhang et al., 2018](#)), 90 nm CMOS ([Karami et al., 2012](#)), and all the way down to 45 nm 3D stacked BSI CMOS ([Lee et al., 2018](#)). These developments were mainly struggling with fill-factor issues, as in a single pixel the SPAD photo-detection structure as well as the active quenching circuit and digital counters and time-to-digital converters (TDCs) are normally to be placed. Two approaches to solve this problem have been followed since: the first one seeks to reduce the technology node in order to diminish the area of the pixel required for the logic and quenching circuits, and the second is pushing the development of 3D integration stacking technology that will be described more in detail later in this chapter, which allows for a separation between the SPAD arrays and their respective circuits on different silicon tiers. Additionally, problems intrinsically related to the SPAD physics such as avalanche processes started by thermally generated electrons (dark counts), electrons previously captured by traps in the silicon bulk and then released at any given moment in time (an effect normally called after pulsing), or avalanche processes started by visible photons in neighboring SPAD cells originated during a starting avalanche event (cross-talk) are some of the issues addressed through various manufacturers and designers over the past years. These and other issues of this very promising technology have been thoroughly explained in ([Schaart et al., 2016](#)).

A lot has been done in this field in the last 5 years, where also an increasing amount of CMOS processing technologies has been further optimized for these tasks. To illustrate the performance these sensors have recently reached, in ([Zhang et al., 2018](#)) a  $252 \times 144$  SPAD pixel sensor called *Ocelot* is reported, fabricated in the 180 nm CMOS technology featuring 1728 12-bit TDCs with a 48.8 ps resolution. In this development, each 126 pixels are connected to six TDCs, which enables effective sharing of resources and a fill-factor of 28% with a pixel pitch of 28.5  $\mu\text{m}$  ([Zhang et al., 2018](#)). On the other hand ([Lee et al., 2017](#)), reports on the world's first BSI 3D-stacked SPAD structure achieving a dark count rate (DCR) of 55.4 cps/ $\mu\text{m}^2$ , a maximum photon detection efficiency (PDE) of 31.8% for 600 nm of impinging (red) light, and a timing jitter of 107.7 ps at 2.5 V of excess bias (over the breakdown voltage in order to operate the SPAD in the Geiger mode). The development of these first SPAD structures in a 45 nm BSI 3D stacked technology opens a path toward high-performance SPAD-based sensors that should be solving many currently present issues in this technology.

Based on the same concept, in an attempt to overcome the so-called Geiger limitation, where a single-SPAD cell remains “blind” (unable to detect an additional photoelectron) during the quenching process, an array of SPAD cells with a short-circuited common output was proposed giving birth to so-called silicon photomultipliers (SiPMs), proposed as an alternative to the PMT technology, originally

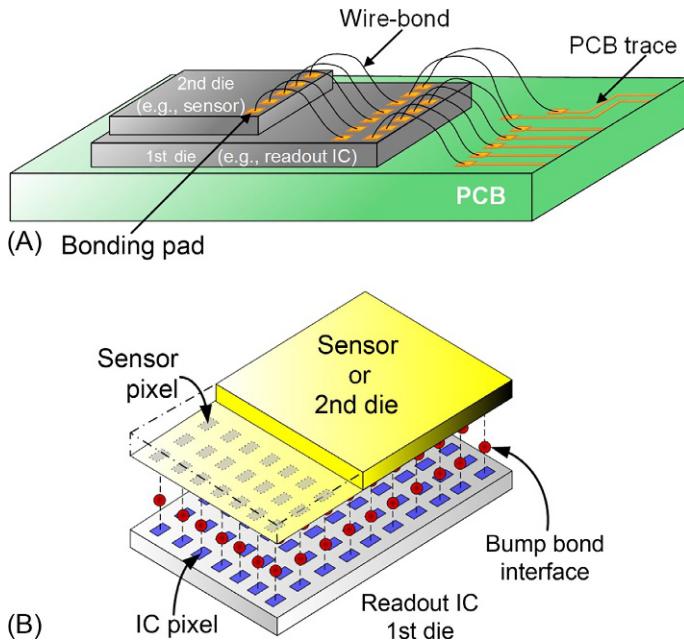
called *micro-pixel avalanche photodiodes* (MAPD) with individual surface quenching resistors in the late 1990s by Russian scientists [Sadygov \(1998\)](#) and [Golovin et al. \(1999\)](#). The idea was not new also outside from Russia, as described by [Dautet et al. \(1993\)](#), and the current developments, issues, and future developments are very well described by [Dinu \(2016\)](#).

The quite attractive and more developed CCD-based counterpart of SPADs and SiPMs using avalanche processes for near single-photon counting performance was proposed by [Hyncek \(1992\)](#) working at Texas Instruments; 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 individual CCD cells, implemented in addition to the standard CCD array, keeping the multiplication noise low, from one cell to the other across a CCD register until multiplication values of several hundreds are occasionally reached. [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.

Another concept was conceived by Eric Fossum and collaborators in 2004 and published in 2005 ([Fossum, 2005](#)), which was abandoned and eventually retook again in 2011, named Quanta image sensor (QIS) ([Fossum et al., 2016](#)). The QIS are conceived as an alternative to detection mechanisms based on avalanche multiplication that could nevertheless enable single-photon counting with very small pixel pitches and importantly lower power consumption. In the single-bit QIS, the output of each field is a binary bit plane, where each bit represents the presence or absence of at least one photoelectron in a photodetector. The specialized sub-diffraction-limit (SDL) photodetectors in the QIS are referred to as “jots” (from Greek for “smallest thing”) and a QIS may have a gigajot or more read out at 1 kfps for a data rate exceeding 1 Tb/s. The idea of this approach is to count photons with SDL pixels with a pitch between 200 nm and 1  $\mu\text{m}$  as they arrive at the sensor ([Fossum et al., 2016](#)). A series of bit planes is generated through high-speed readout and a kernel or “cubicle” of bits  $(x, y, t)$  is used to create a single output image pixel thought of as a jot data cube having two spatial dimensions ( $x$  and  $y$ ) with the third dimension being time. The current QIS devices are being fabricated in 65-nm BSI CIS process yielding read noise as low as 0.22  $\text{e}^-_{\text{rms}}$  and a conversion gain as high as 420  $\mu\text{V/e}^-$  using power efficient readout electronics currently as low as 0.4 pJ/b fabricated in the same process ([Fossum et al., 2016](#)). A room-temperature QIS recently implemented in a slightly modified standard CMOS images sensor (CIS) 3D stacked BSI fabrication process yields a FWC of only a few electrons and deep subelectron read noise of less than 0.3  $\text{e}^-_{\text{rms}}$  ([Fossum and Anagnos, 2018](#)).

## 2.10 Hybrid and 3D detector technologies

As an alternative for dealing with all the problems related to pixel fill-factors and the heterogeneity of requirements for the CMOS processes used to fabricate the different parts of an image sensor (front-end photodetectors and pixels requiring higher biasing voltages, thicker oxides or lower substrate resistivities, and the readout and data



**Fig. 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.

interface electronics requiring higher speed and higher functionalities) accompanying the “More Moore” developments, the American physicist Richard Feynman proposed in his lectures given in Tokyo in 1985 an IC stack. 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 Memorial Lectures, 2008](#)).

3D 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. 2.13A](#). 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 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. 2.13B). 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. 2.13B. 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.

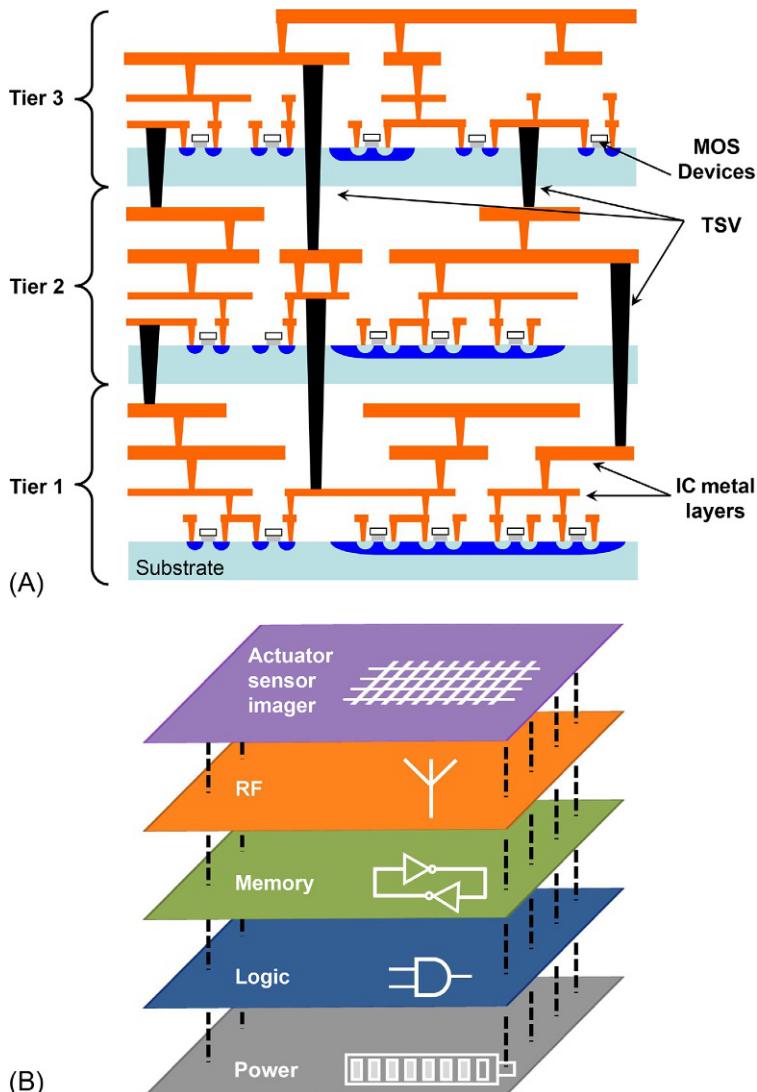
The flip-chip technique is also used in the development of imaging devices. The basic idea is to split the imager into two separate parts: the sensor part and the readout IC part, and interconnect them using this technique. The readout IC contains pixel-related active readout and signal processing electronics. 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 µm 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  $170 \times 170 \mu\text{m}^2$  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, e.g., Bower et al., 2008). Fig. 2.14A shows several vertically stacked tiers 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



**Fig. 2.14** (A) A cross-section of a diagram of a 3D integrated circuit consisting of three tiers (dice) stacked on top of each other and interconnected using through silicon vias (*black pillars*) and (B) technology mixing possibilities of 3D integration.

step. A typical via diameter starts from less than a micrometer 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. 2.14B. 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 micrometers, 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.

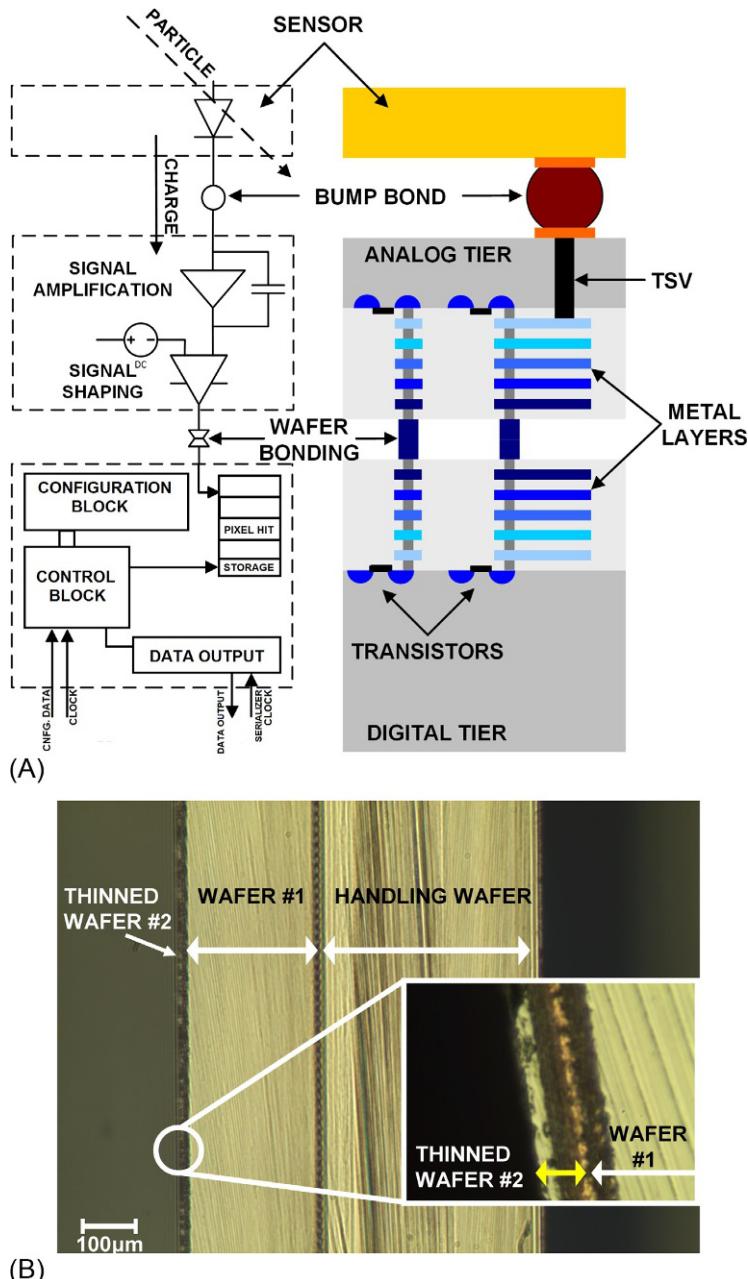
As it will be explained in depth by Robert Gove in Chapter 7, first commercially available 3D integrated CMOS image sensors have been hitting the market since 2017 pursuing smaller pixel pitches and incrementing the signal and data processing capability on-chip. The first three-layer stacked CMOS image sensor die in volume

production, the 19.3 megapixels IMX400, was designed into Sony's Xperia XZs phone (Haruta et al., 2017). Sony has also used stacking to achieve a 1  $\mu\text{m}$  pixel pitch introducing a back-illuminated CIS (BI-CIS) with hybrid bonding of two substrates with Cu-Cu metal bonding and an interlayer dielectric (ILD) oxide bonding. The interconnect pitch between the CIS and ISP layers is 3  $\mu\text{m}$  (enabling direct bonding of groups of pixels as well as bonding on the periphery of the die fabricated on each tier). The device is a stacked BSI-CIS with 22.5 megapixels (Kagawa et al., 2017). On their side also OmniVision and TSMC feature a three-layer stack based on 55-nm CMOS process enabling the design of a high-performance 24-megapixels sensor with 0.9  $\mu\text{m}$  pixels—PureCel<sup>R</sup>—S (Venezia et al., 2017).

Over the past decade, 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, while the 3D integration of the fabricated tiers was performed by Tezzaron Semiconductor. One of the prototypes developed, the FE-TC4 (Fig. 2.15A), represents a two tier pixelated sensor readout IC, consisting of 854 pixels with a pixel size of 50  $\mu\text{m} \times$  166  $\mu\text{m}$ . The 3D pixel functional diagram and its technology cross-section (out of scale) diagram are both shown in Fig. 2.15A. 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 backside of the tier providing the analog pixel functionality is thinned down to 12  $\mu\text{m}$  (see Fig. 2.15B) 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.

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) as reported in Boulenc et al. (2017) and Eckardt (2014), 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.

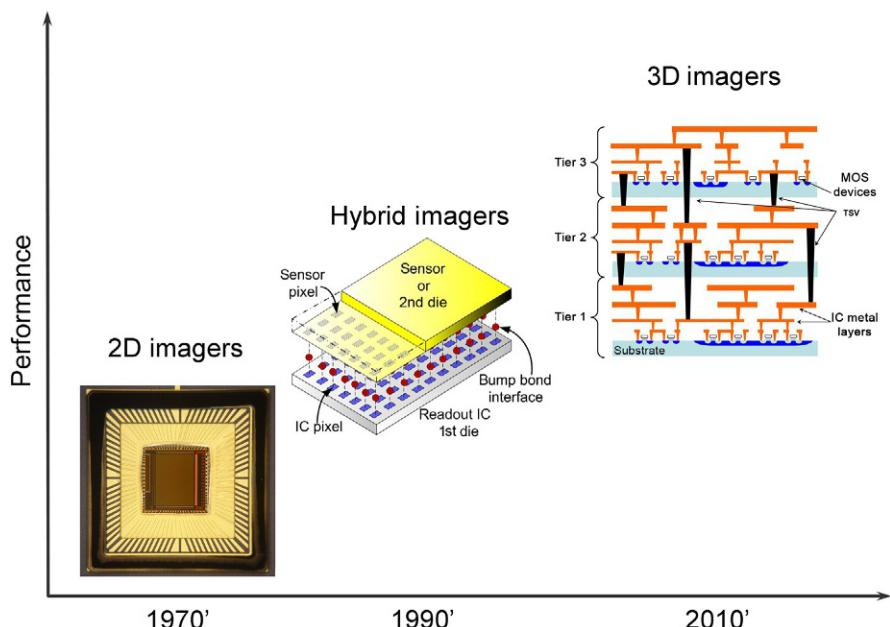


**Fig. 2.15** (A) Graphical representation of the pixel of the ATLAS pixel detector readout IC prototype FE-TC4 developed in 3D integration technology and (B) a photograph of the cross-section of the FE-TC4. The bonding interface is zoomed ([Eckardt, 2014](#)).

## 2.11 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 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 postprocessing, which is an issue in most cases. The trend is, nevertheless, that in future one single foundry also provides postprocessing expertise.

Although much improved and already comparable with CCD performances, the 2D CIS still suffer from miss-match related photoresponse nonuniformity (PRNU) and dark signal nonuniformity (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, for example,



**Fig. 2.16** From 2D image sensors to 3D imagers.  
Photograph by Stefan Bröcker.

time-delayed integration (TDI) readout and column-wise 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 already used in commercial CIS developed, for example, by SONY or OmniVision and TSMC, brings a series of new challenges where image sensor performances are concerned that are yet to be solved. In all cases, there is much to be done in the future.

## References

- Amelio, G.F., Tompsett, M.F., Smith, G.E., 1970. Experimental verification of the charge coupled device concept. *Bell Syst. Tech. J.* 49 (4), 593–600.
- 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., Mekkaoui, A., 2013. Experience with 3D integration technologies in the framework of the ATLAS pixel detector upgrade for the HL-LHC. *Nucl. Inst. Methods Phys. Res., Sect. A*.
- Bisogni, M.G., et al., 1998. Performance of 4096 pixel photon counting chip. In: CERN-EP-98-162, CERN, Geneva, Switzerland.
- Boulenc, P., Robbelein, J., Wu, L., Haspeslagh, L., De Moor, P., Borremans, J., Rosmeulen, M., 2017. High speed TDI embedded CCD in CMOS sensor. In: Proc. SPIE 10562, International Conference on Space Optics—ICSO 2016, 105622P (25 September 2017). <https://doi.org/10.1117/12.2296149>.
- Bower, C., Garrou, P., Ramm, P., 2008. *Handbook of 3D Integration: Technology and Application of 3D Integration Circuits*. Wiley-VCH, Weinheim.
- Boyle, W.S., Smith, G.E., 1970. Charge coupled semiconductor devices. *Bell Syst. Tech. J.* 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., Paschen, U., 2013. Large area CMOS SPADs with very low dark counting rate. In: Proc. SPIE 8631, Quantum Sensing and Nanophotonic Devices X, 86311B.
- Burkey, B.C., Chang, W.C., Littlehale, J., Lee, T.H., 1984. The pinned photodiode for an interline-transfer image sensor. In: 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  $0.7\text{e}^-_{\text{rms}}$  temporal-readout-noise CMOS image sensor for low-light-level imaging. In: Solid-State Circuits Conference Digest of Technical Papers (ISSCC), IEEE International, pp. 384–386.
- Chung, S.-M., Kim, Y.-H., 2007. Fine pitch and low temperature bonding using solder bumps and the assessment of solder joint reliability. *Mater. Trans.* 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., Andreoni, A., 1981. Towards picosecond resolution with single-photon avalanche diodes. *Rev. Sci. Instrum.* 52, 408.
- Dautet, H., Deschamps, P., Dion, B., MacGregor, A., MacSween, D., McIntyre, R.J., Trottier, C., 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*, second ed. Springer-Verlag, Berlin Heidelberg.
- Dick, J., 2013. Inside today's hot products: what teardowns can reveal. In: Chipworks, The Content Strategy Conference (ConFab) 2013.
- Dinu, N., 2016. Silicon photomultipliers. In: Nabot, B. (Ed.), *Photodetectors. Materials, Devices and Applications*, Woodhead Publishing Series, Elsevier, Cambridge, UK, pp. 255–314 (Chapter 8).
- Durini, D., 2009. *Solid-State Imaging in Standard CMOS Processes*. Ph.D. Thesis, University Duisburg-Essen, Germany.
- Durini, D., Hosticka, B., 2007. Photodetector structures for standard CMOS imaging applications. In: Proc. IEEE PRIME 2007, Bordeaux, France, pp. 193–196.
- Eckardt, A., 2014. High resolution earth observation cameras. In: Proc. 7th Fraunhofer IMS Workshop CMOS Imaging From Photon to Camera, May 20–21, 2014, Duisburg, Germany.
- EMVA 1288, 2013. [http://www.emva.org/cms/upload/Standards/Stadard\\_1288/EMVA1288-3-Irc.pdf](http://www.emva.org/cms/upload/Standards/Stadard_1288/EMVA1288-3-Irc.pdf).
- Fossum, E.R., 1995. CMOS image sensors: electronic camera-on-a-chip. In: IEEE IEDM Tech. Digest, pp. 17–25.
- Fossum, E.R., 1997. CMOS image sensors: electronic camera-on-a-chip. *IEEE Trans. Electron Devices* 44 (10), 1689–1698.
- Fossum, E.R., 2005. Some thoughts on future digital still cameras. In: Nakamura, J. (Ed.), *Image Sensors and Signal Processing for Digital Still Cameras*. CRC Press, Boca Raton, FL, pp. 305–314.
- Fossum, E.R., Anagnos, K., 2018. The quanta image sensor (QIS): making every photon count. *Laser Focus World*, 13–15. December 2018.
- Fossum, E.R., Ma, J., Masoodian, S., Anzagira, L., Zizza, R., 2016. Review: the quanta image sensor: every photon counts. *Sensors* 2016 (16), 1260. <https://doi.org/10.3390/s16081260>.
- Fry, P., Noble, P., Rycroft, R., 1970. Fixed pattern noise in photomatrices. *IEEE J. Solid State Circuits* SC-5, 250–254.
- Golovin, V., Tarasov, M., Bondarenko, G., 1999. Avalanche Photodetector. Russian Patent No. 2142175, Application Reg. No. 98117319/28, Nov. 27, 1999.
- Green, M.A., Keevers, M., 1995. Optical properties of intrinsic silicon at 300 K. *Prog. Photovoltaics* 3 (3), 189–192.
- Guidash, R.M., et al., 1997. A 0.6 μm CMOS pinned photo diode color imager technology. In: IEDM Technical Digest, pp. 927–929.
- Haitz, R.H., 1965. Mechanisms contributing to the noise pulse rate of avalanche diodes. *J. Appl. Phys.* 36 (10), 3123–3131.
- Haruta, T., Nakajima, T., Hashizume, J., Nomoto, T., et al., 2017. A 1/2.3-inch 20 Mpixel 3-layer stacked CMOS image sensor with DRAM. In: Proc. of ISSCC 2017 (4.6).
- Hecht, E., 1989. *Optik*. translated by F. Siemsen (original title: Optics), Addison-Wesley, Hallbergmoos.

- Horton, J., Mazza, R., Dym, H., 1964. The scanistor—a solid-state image scanner. *Proc. IEEE* 52, 1513–1528.
- Hosticka, B.J., 2007. Analog circuits for sensors. In: *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. Electron Devices* 39 (8), 1972–1975.
- Intel Corporation, 2013. Laptop, Desktop, Server and Embedded Processor Technology—Intel (online). <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., Heyes, P., 2001. The LLCCD. Low light imaging without the need for an intensifier. *Proc. SPIE* 178.
- Johnson, J.B., 1928. Thermal agitation of electricity in conductors. *Phys. Rev. Am. Phys. Soc.* 32, 97–109.
- Kagawa, Y., Fuji, N., Aoyagi, K., Kobayashi, Y., Nishu, S., Takeshita, S., Taura, J., Takahasi, H., Nishimura, Y., Tatani, K., Kawamura, M., 2017. An advanced CuCu hybrid bonding for novel stacked CMOS image sensor. In: *Proc. 2017 International Image Sensor Workshop (IISW 2017)*.
- Karami, M.A., Gersbach, M., Charbon, E., 2012. A new single-photon avalanche diode in 90 nm standard CMOS technology. In: *Proc. SPIE*, 7780.
- Kittel, C., 1996. *Introduction to Solid State Physics*, seventh ed. John Wiley & Sons Inc., New York, p. 673.
- Kosonocky, E.F., Carnes, J.E., 1971. Charge-coupled digital circuits. *IEEE J. Solid State Circuits* SC-6 (5), 314–322.
- Lai, J., Nathan, A., 2005. Reset and partition noise in active pixel image sensors. *IEEE Trans. Electron Devices* 52 (10), 1–4.
- Leach Jr., W.M., 1994. Fundamentals of low-noise analog design. *Proc. IEEE* 82 (10), 1514–1538.
- Lee, P., Gee, R., Guidash, M., Lee, T., Fossum, E.R., 1995. An active pixel sensor fabricated using CMOS/CCD process technology. In: *Proc. 1995 IEEE Workshop on CCDs and Advanced Image Sensors*, Dana Point, CA, 20–22 April.
- Lee, M.J., Ximenes, A.R., Padmanabhan, P., Wang, T.J., Huang, K.C., Yamashita, Y., Yaung, D.N., Charbopn, E., 2017. A back-illuminated 3D-stacked single-photon avalanche diode in 45nm CMOS technology. In: *Proc. 2017 IEEE International Electron Devices Meeting (IEDM)*, Dec. 2–6, 2017, San Francisco, USA, pp. 405–408.
- Lee, M.-J., Ximenes, A., Padmanabhan, P., Wang, T.-J., Huang, K.-C., Yamashita, Y., Yaung, D.-N., Charbon, E., 2018. High-performance back-illuminated three-dimensional stacked single-photon avalanche diode implemented in 45-nm CMOS technology. *IEEE J. Sel. Top. Quantum Electron.* 24 (6), 1–1.
- Lilienfeld, J.E., 1928. Device for Controlling Electric Current. US Patent No. 1,900,018, filed March 1928, patented 7 March 1933.
- Loukianova, N.V., Folkerts, H.O., Maas, J.P.V., Verbugt, D.W.E., Mierop, A.J., Hoekstra, W., Roks, E., Theuwissen, A.J.P., 2003. Leakage current modeling of test structures for characterization of dark current in CMOS image sensors. *IEEE Trans. Electron. Devices* 50 (1), 77–83.
- Magnan, P., 2003. Detection of visible photons in CCD and CMOS: a comparative view. *Nucl. Inst. Methods Phys. Res. A* 504, 199–212.

- McKelvey, J.P., 1982. Solid State and Semiconductor Physics. Krieger Publishing Co., Malabar, p. 512.
- Mendis, S., Kemeny, S., Fossum, E.R., 1993. A  $128 \times 128$  CMOS active pixel image sensor for highly integrated imaging systems. In: IEEE IEDM Tech. Digest, pp. 583–586.
- Moore, G.E., 1965. Cramming more components onto integrated circuits. In: Electronics. vol. 38.8. [ftp://download.intel.com/museum/Moores\\_Law/Articles-Press\\_Releases/Gordon\\_Moore\\_1965\\_Article.pdf](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., Fossum, E.R., 1995.  $128 \times 128$  CMOS photodiode-type active pixel sensor with on-chip timing, control and signal chain electronics. In: 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. Phys. Rev. Am. Phys. Soc. 32, 110–113.
- Pain, B., Yang, G., Cunningham, T., Hancock, B., 2000. Photodiode based CMOS APDs with zero lag, low noise, high-linearity design. In: NASA Tech. Briefs, ABP.
- Rossi, L., Fischer, P., Rohe, T., Wermes, N., 2006. Pixel Detectors From Fundamentals to Applications. Springer Verlag, Berlin and Heidelberg.
- Sadygov Z. (1998) Avalanche Detector, Russian Patent 2102820, Application Reg. No. 96119669/25, Jan. 20, 1998.
- Schaart, D.R., Charbon, E., Frach, T., Schulz, V., 2016. Advances in digital SiPMs and their application in biomedical imaging. Nucl. Inst. Methods Phys. Res. A809, 31–52.
- Scholze, F., et al., 2000. Determination of the electron-hole pair creation energy for semiconductors from the spectral responsivity of photodiodes. Nucl. Inst. Methods Phys. Res., Sect. A 439, 208–215.
- Schroder, D.K., 1998. Semiconductor Material and Device Characterization, second ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Schuster, M.A., 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, second ed. John Wiley & Sons, Inc., Hoboken, NJ.
- Teranishi, N., Kohono, A., Ishihara, Y., Oda, E., Arai, K., 1982. No image lag photodiode structure in the interline CCD image sensor. In: Proc. IEEE Electron Device Meeting (IEDM), pp. 324–327.
- The ATLAS Collaboration, et al., 2008. The ATLAS experiment at the CERN large hadron collider. J. Instrum. 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 Electron. 52, 1401–1406.
- Tompsett, M.F., Amelio, G.F., Smith, G.E., 1970. Charge coupled 8-bit shift register. Appl. Phys. Lett. 17, 111–115.

- Venezia, V.C., Shih, C., Yang, W., Zang, Y., Lin, Z., Grant, A., Rhodes, H., 2017. 1.0um pixel improvements with hybrid bond stacking technology. In: Proc. 2017 International Image Sensor Workshop, R03 (IISW 2017).
- Wadsworth, M., McGrath, R.D., 1984. Charge coupled devices for X-ray imaging. In: IEDM Proceedings, pp. 20–24.
- Walden, R., Krambeck, R., Strain, R., Mckenna, J., Schryer, N., 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 Jr., N., 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.
- Zhang, C., Lindner, S., Antolovic, I.M., Pavia, J.M., Wolf, M., Charbon, E., 2018. A 30-frames/s, 252 × 144 SPAD flash LiDAR with 1728 dual-clock 48.8-ps TDCs, and pixel-wise integrated Histogramming. IEEE J. Solid-State Circuits. <https://doi.org/10.1109/JSSC.2018.2883720> (accessed December 2018).
- Zhong, Z.W., 2009. Fine and ultra-fine pitch wire bonding: challenges and solutions. Micro-electron. Int. 26 (2), 10–18.

## Further reading

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

This page intentionally left blank

# Charge-coupled device (CCD) image sensors

3

M. Lesser

Steward Observatory, UA Imaging Technology Laboratory, University of Arizona,  
Tucson, AZ, United States

## 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 we describe the CCD imager and its use for high-performance imaging. Our 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 (QE), backside charging, and antireflection (AF) 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 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  $E_\lambda$  of wavelength  $\lambda_{\text{cutoff}}$  is greater than the bandgap of silicon in order to excite an electron from the silicon valence band to the conduction band,

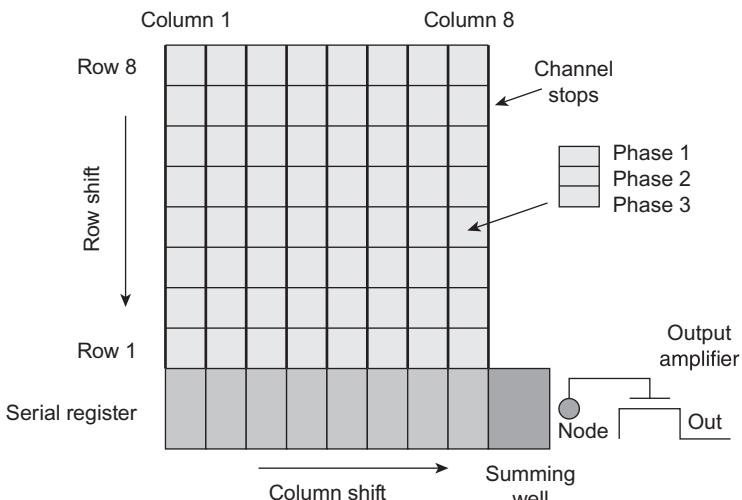
$$E_\lambda = \frac{hc}{\lambda} > E_{\text{bandgap}} \quad (3.1)$$

where  $h$  is Planck's constant and  $c$  is the speed of light. For silicon  $E_{\text{bandgap}} = 1.1 \text{ eV}$ , so  $\lambda_{\text{cutoff}} = \frac{hc}{1.1 \text{ eV}} = 1.12 \mu\text{m}$ . Silicon detectors are therefore excellent detectors for nearly all wavelengths shortward of about  $1 \mu\text{m}$ , in the near-infrared (IR) region. 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 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 (2D) digital image.

Pixels are composed of phases each of which has 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



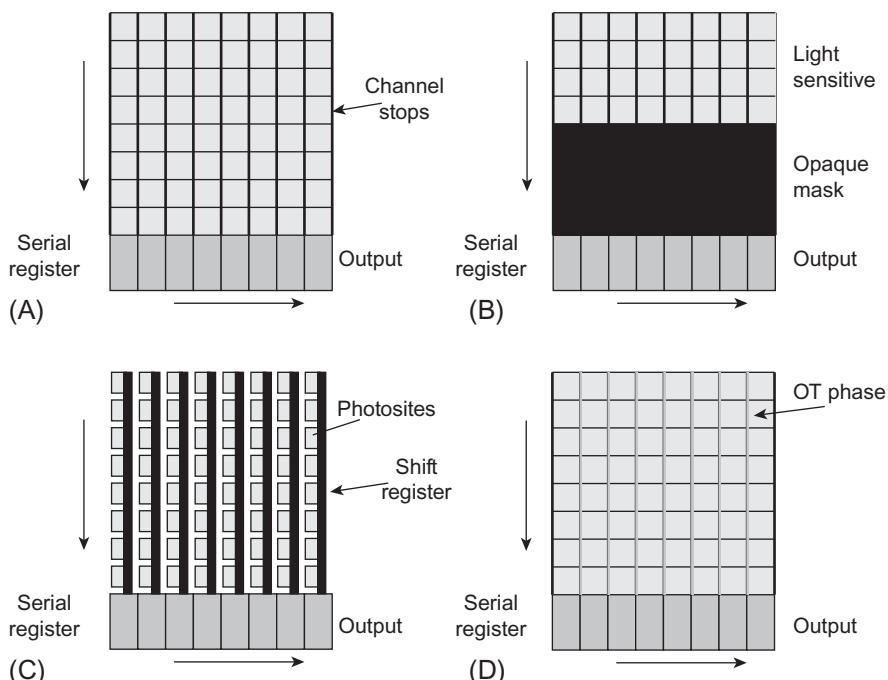
**Fig. 3.1** Layout of a typical three-phase CCD.

simple CCD. There 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  $2\text{--}30\ \mu\text{m}$ , with the smaller sizes being more common on the most recent sensors.

The CCDs can be divided into two main types: linear arrays and area arrays. Linear arrays consist of one (or a few) column 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



**Fig. 3.2** Common area array CCD architectures. (A) Full frame CCD, (B) Frame transfer CCD, (C) Interline transfer CCD, (D) Orthogonal transfer CCD.

during integration. This allows a very rapid shift ( $10^{-6}$ – $10^{-4}$  s) from image store to frame store after integration. If the shift 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 readout 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 that 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.

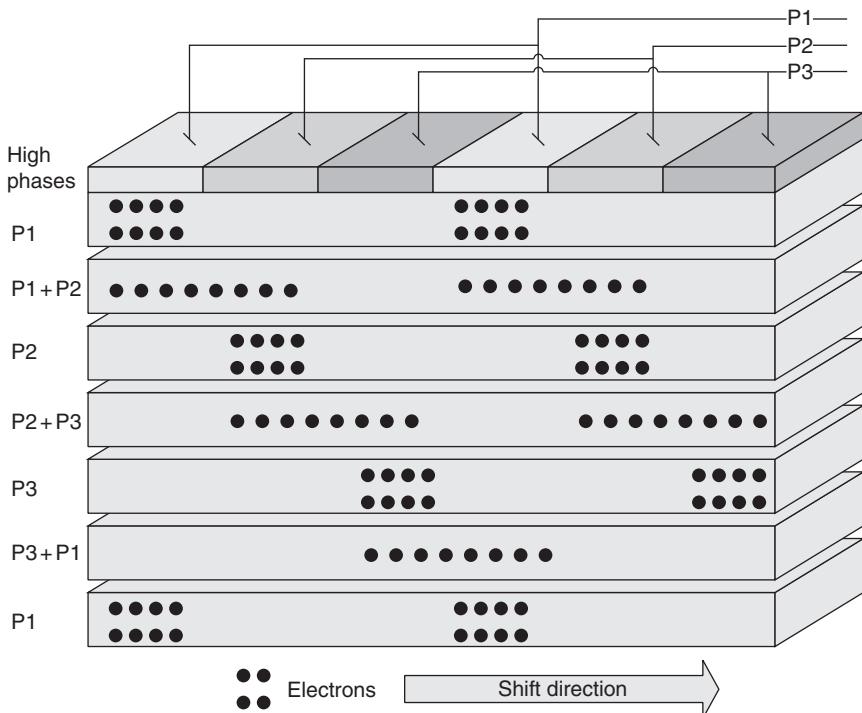
The orthogonal transfer array (OTA) is a monolithic device composed of (nearly) independent cells each of which is an orthogonal transfer CCD (Burke et al., 2004). The advantage of the OTA over the OTCCD is that the same detector can both provide the feedback signal and perform the data observation. 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 (SNR). 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. OTA's have on-chip logic to address the OTCCD cells so that each cell can have independent timing. This allows some cells to be readout while others are integrating. The integrating cells can shift small amounts in X and Y based on the feedback signal obtained from the cells being readout 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.

#### 3.2.3.1 Integration during exposure

During integration (or light collection) the potential minimum 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 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.



**Fig. 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.

### 3.2.3.2 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.

### 3.2.3.3 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.

### 3.2.3.4 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 readout 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 SNR of the image is increased with binning,

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

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

## 3.2.4 Time delay and integration operation

An alternative operating mode to the more common “integrate and shift” exposure sequence is the so-called time delay and integration or TDI mode. This method of CCD readout combines the integration and shift operations when the scene being imaged can be moved across the sensor. Often TDI mode is used for high-speed

imaging when a scene is synchronously moved along CCD columns at the same rate as the parallel clocks are shifted. By altering the shift time though clocking parameters and/or the speed of the scene movement, an arbitrary exposure time may be achieved to optimize the SNR. The effective exposure time is the sum of the integration times of all the pixels which imaged the same image scene element.

An advantage of TDI mode is that each element of the scene is imaged by multiple physical pixels as the scene shifts across different rows of the CCD. This causes any pixel-to-pixel variations in photoresponse to be averaged, resulting in more uniform images. This technique can significantly improve photoresponse uniformity by averaging out effects due to issues such as cosmetic defects, dust on the sensor or optics, and interference fringing. TDI can also improve observing efficiency because the fraction of time the sensor is integrating may be higher than the traditional integrate and shift technique.

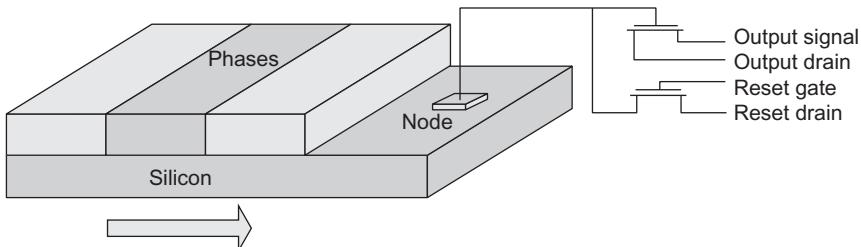
A common use of TDI mode for astronomical imaging takes advantage of the Earth's rotation. Instead of tracking a telescope to compensate for the Earth's rotation, the telescope can be fixed and the CCD shifted at the same rate as the projected sky crosses the sensor. A continuous image is readout of the CCD with no shuttering required. Each column of the resultant image has been created by the average photoresponse of all the pixels in the column. A variation of this mode is to move the telescope at a non-sidereal rate while adjusting the clocking speed in order to obtain different effective exposure times. TDI mode for astronomical imaging also reduces the need to precision telescope tracking, reducing the overall system cost.

### 3.2.5 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

$$V = Nq/C \quad (3.3)$$

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 ( $1.6 \times 10^{-19}$  C). Typically a single



**Fig. 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.

electron produces from 1 to 50  $\mu\text{V}$  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  $\sim 100$  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 SNR 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.

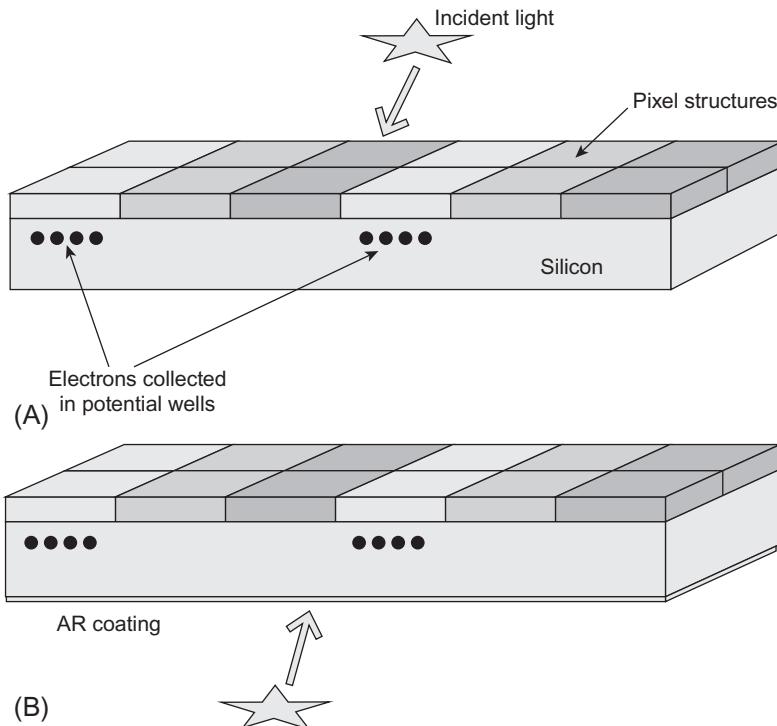
### 3.3 Illumination modes

The 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 on the opposite or back side ([Fig. 3.5](#)).

#### 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 as they are fabricated from the silicon wafer (no additional processing steps). However, the front side gates absorb almost all blue and ultraviolet (UV) light and so they are not directly useful for imaging at wavelengths  $\lambda < 400 \text{ nm}$ . In addition, the physical gate structure causes many reflections, including complex 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.



**Fig. 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).

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 reemission 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 add considerably to the 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 detectors 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 AF-coated back-illuminated CCD may have a peak QE > 98% in the visible.

Optical absorption and multiple reflections from frontside structures are 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 back-side 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.

### ***3.3.2.1 Backside thinning***

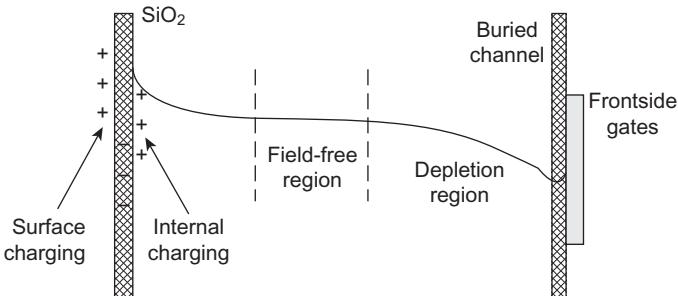
A back-illuminated CCD must be relatively thin in order for photogenerated electrons to diffuse to the frontside pixel wells and be collected under the sites where they were created. Most CCDs are built on epitaxial silicon with layers that are 10–20  $\mu\text{m}$  thick. The device must be thinned 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](#)).

### ***3.3.2.2 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 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 and Iyer, 1998](#); [Lesser, 2000](#)), 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.

### ***3.3.2.3 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 proper indices and be non-absorbing in the spectral region of interest ([Lesser, 1987, 1993](#)). With absorbing substrates that have indices



**Fig. 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.

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.

### 3.3.2.4 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

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

where  $x_{ff}$  is the field-free thickness and L is the distance from the backside surface 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  $>1000 \Omega \text{ cm}$ , 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 Quantum efficiency

The QE is the measure the 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.

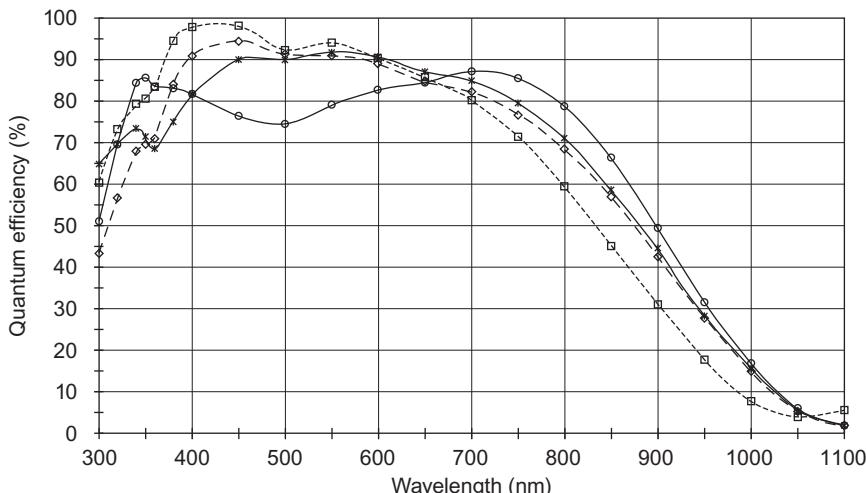
The absorptive quantum efficiency  $QE_\lambda$  is the fraction of incident photons which is absorbed in the detector and is given by

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

where  $R_\lambda$  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,  $\alpha_\lambda$  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

1. reducing surface reflectivity (reduce  $R_\lambda$  with AF coatings)
2. increasing the thickness of absorbing material (increase  $t$ )
3. increasing the absorption coefficient (decrease  $\alpha_\lambda$  by material optimization)

Because nearly all CCDs are made with standard silicon semiconductor processing techniques, only options 1 and 2 are viable. The 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.



**Fig. 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.

Back-illuminated CCDs are in fact being fabricated considerably thicker than it was a few years ago. This is to increase the QE especially in 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 as thick as 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](#)).

### 3.4.1.1 Quantum yield

Related to QE is quantum yield (QY) which is the term applied to the phenomenon in which one energetic interacting photon may create multiple electron-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 a 5.9-keV Fe-55 X-ray source which produces  $\sim$ 1620 electrons per incident photon. For photon energies  $> 3.1$  eV, or about  $0.4\text{ }\mu\text{m}$ ,

$$QY = \frac{E_\lambda}{E_{e-h}} \quad (3.6)$$

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

The 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 accurately control 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.

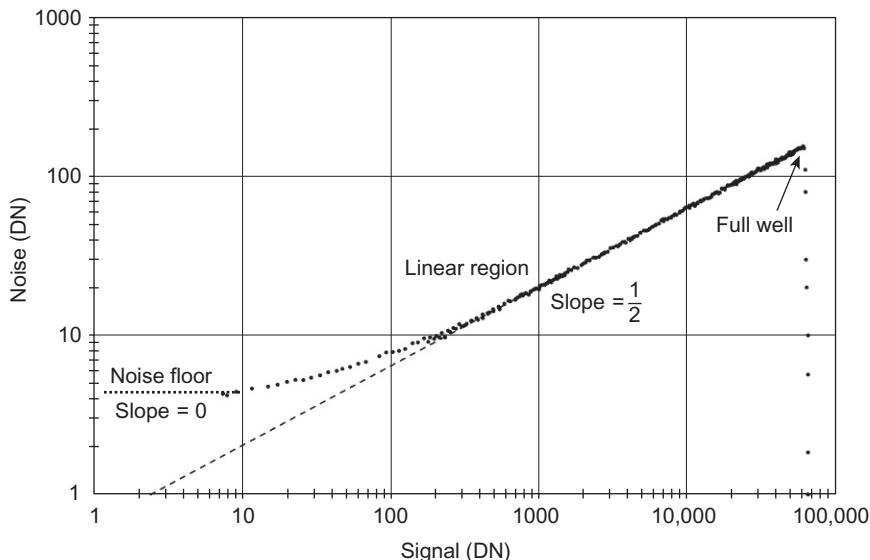
Nearly all high-end cameras use correlated doubled sampling (CDS) to reduce read noise by reducing the uncertainty in absolute charge level at the output node. When the output node is reset, its final value is uncertain due to kTC or thermal noise. The CDS technique eliminates this uncertainty by sampling 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  $\sigma$  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 using 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 curve 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):



**Fig. 3.8** A photon transfer curve. The three regions described in the text can be seen in this PTC for a  $2k \times 2k$  pixel CCD.

- slope=0 at low signal is the read noise floor
- slope= $\frac{1}{2}$  (linear portion) is where photon shot noise is dominant
- significant deviation from slope  $\frac{1}{2}$  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.

### 3.4.4 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, an Fe-55 X-ray event (1620 electrons) in the far corner of a  $4k \times 4k$  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. An  $Fe^{55}$  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 vs 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 nonlinearity 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 determined 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

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

where  $D_{FM}$  is a silicon “quality” parameter in nA/cm<sup>2</sup> at 300 K,  $A_{pix}$  is the pixel area,  $E_g$  is the silicon bandgap,  $T$  is the 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  $-100^\circ\text{C}$  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  $-40^\circ\text{C}$  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 accurately make. 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 uncertainty 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 are important when characterizing absolute dark signal.

#### 3.4.6.1 Dark noise

Dark noise is the uncertainty 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 so that dark signal is not the dominant noise source. This may occur at very low temperatures ( $-100^{\circ}\text{C}$  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 nonuniformity or PRNU), dark signal variations (dark signal nonuniformity or 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 SNR 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 CCD facilities and also the need for low-power sensors with integrated image processing capabilities. However, there are several areas where CCDs will likely remain dominant for many years.

First and foremost large pixel image sensor 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  $10\text{k} \times 10\text{k}$  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 ( $> 20 \mu\text{m}$ ) 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.

## References

- Burke, B., Tonry, J., Cooper, M., Luppino, G., Jacoby, H., Bredthauer, R., Boggs, K., Lesser, M., Onaka, P., Young, D., Doherty, P., Craig, D., 2004. The orthogonal-transfer array: a new CCD architecture for astronomy. Proc. SPIE 5499, 185–192.
- Dereniak, E., Crowe, D., 1984. Optical Radiation Detectors. John Wiley and Sons.
- Holland, S., Groom, D., Palaio, N., Stover, R., Wei, M., 2003. Fully depleted, back-illuminated charge-coupled devices fabricated on high-resistivity silicon. IEEE Trans. Electron Devices 50 (3), 225–238.
- Hynecek, J., 2001. Impactron—a new solid state image intensifier. IEEE Trans. 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., Heyes, P., 2001. The LLLCCD: low light imaging without the need for an intensifier. Proc. SPIE 4306, 178–186.
- Leach, R., Lesser, M., 1987. Improving the blue and UV response of silicon CCD detectors. Publ. Astron. Soc. Pac. 99, 668.
- Lesser, M., 1987. Antireflection coatings for silicon CCDs. Opt. Eng. 26, 911.
- Lesser, M., 1990. CCD thinning, coating, and mounting research for astronomy. In: Jacoby, G. (Ed.), CCD's in Astronomy. BookCrafters, p. 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., 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., Fattah, M., Hseng, H.-F., 1994. Delta-doped CCDs for high and stable in the UV and Visible. Proc. SPIE 2198, 907.

- Theuwissen, A., 1995. Solid-State Imaging With Charge-Coupled Devices. Kluwer Academic Publishing.
- Tony, J.L., Burke, B.E., Schechter, P.L., 1997. The orthogonal transfer CCD. PASP 109, 1154–1164.
- Zacharias, N., Dorland, B., Bredthauer, R., Boggs, K., Bredthauer, G., Lesser, M., 2007. Realization and application of a 111 million pixel backside-illuminated detector and camera. Proc. SPIE 6690, 8.

This page intentionally left blank

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

4

A. Lahav<sup>a</sup>, A. Fenigstein<sup>b</sup>, A. Strum<sup>b</sup>

Revised by S. Rizzolo<sup>c</sup>

<sup>a</sup>Tower Semiconductor Ltd, Migdal HaEmek, Israel, <sup>b</sup>TowerJazz, Newport Beach, CA, United States, <sup>c</sup>Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, France

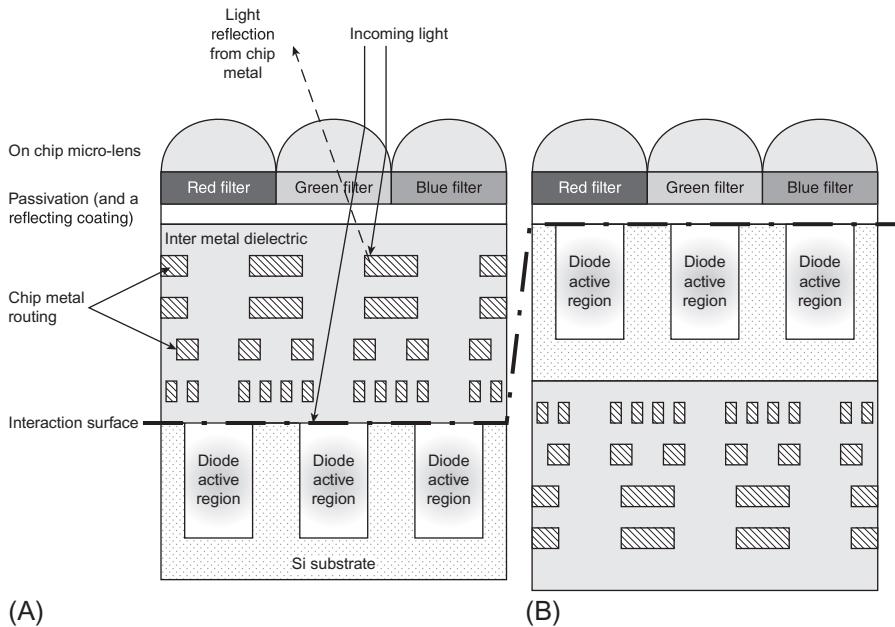
## 4.1 Introduction

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 (QE), 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  $\mu\text{m}$  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 describe briefly the challenges for scaled down pixels in a modern complementary metal-oxide-semiconductor (CMOS) image sensor chip. We review various technologies used to overcome these challenges and show why BSI is the most natural solution for such sensors.

CMOS image sensors (CSIs) consist of an array of light-sensitive pixels. Each pixel consists of a photodiode (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 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 frontside illuminated (FSI) sensor (see Fig. 4.1A), the light reaches the PD active region through the passivation, metallization, and inter-dielectric layers. Several are the loss mechanisms associated with the coupling of light from the frontside of the sensor. First, the reflection of light from the passivation layer, second, the reflection of light from the metal control lines which surround the PD, and third, the shift



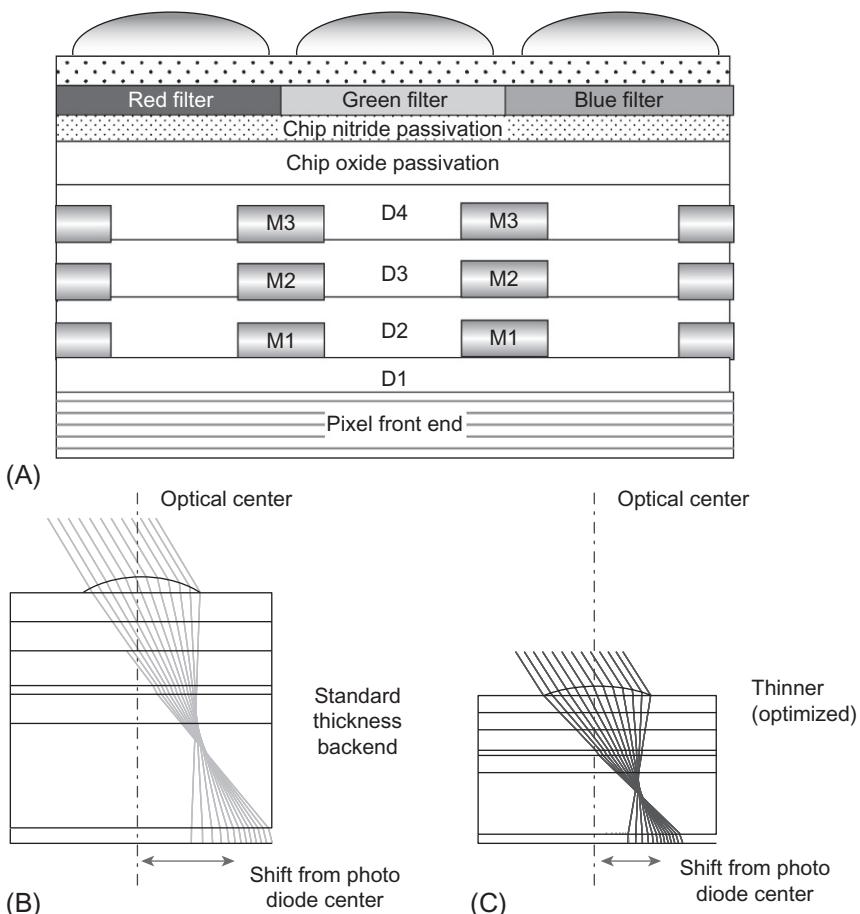
**Fig. 4.1** Illustration of (A) frontside vs (B) backside illuminated sensor.

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, QE reduction, and contribute to large cross talk, and as a result, reduction of the sensor's signal-to-noise ratio (SNR). Moreover, the stratification of the different layers above the PD can cause a chip nonuniformity which leads to a difference in the light collection from the edges to the center of the image sensor.

A BSI sensor contains the same components as a FSI sensor but the sensor's metals are located behind the PD (see Fig. 4.1B). 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 FSI sensor

The typical back-end scheme of a CMOS image sensor is shown in Fig. 4.2A. 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 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



**Fig. 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.

the PD 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 cross-talk 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. 4.2B and 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.

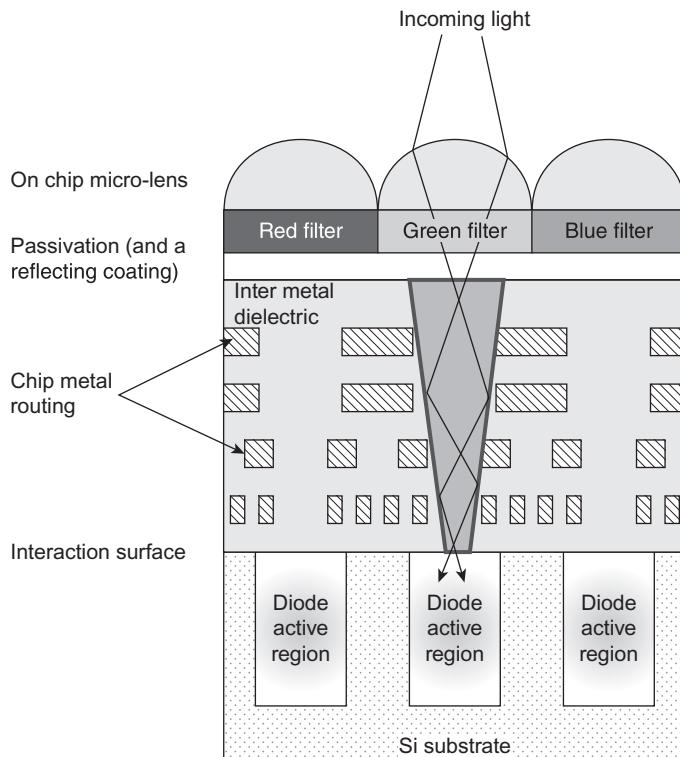
A larger optical opening above the PD is key for achieving higher QE and better angular response. It can be achieved by using narrow control lines for the pixel. A pixel designer would use (in most cases) the narrowest line width that could be manufactured in a given set of fabrication tools. Therefore, 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  $\mu\text{m}$  down to 0.18  $\mu\text{m}$  (Agranov et al., 2005).

For the 0.13- $\mu\text{m}$  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 (Edelstein et al., 1997). 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 (Cohen et al., 2006). 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  $\mu\text{m}$ , creative solutions and nonstandard backend modules were developed. One example is recess of the metallization stack above the pixels array (Adkisson et al., 2010). 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 (Rhodes, 2008). 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 (Gambino et al., 2006; Agranov et al., 2009) which is illustrated in Fig. 4.3. A “light pipe” (or “light guide”) is fabricated by

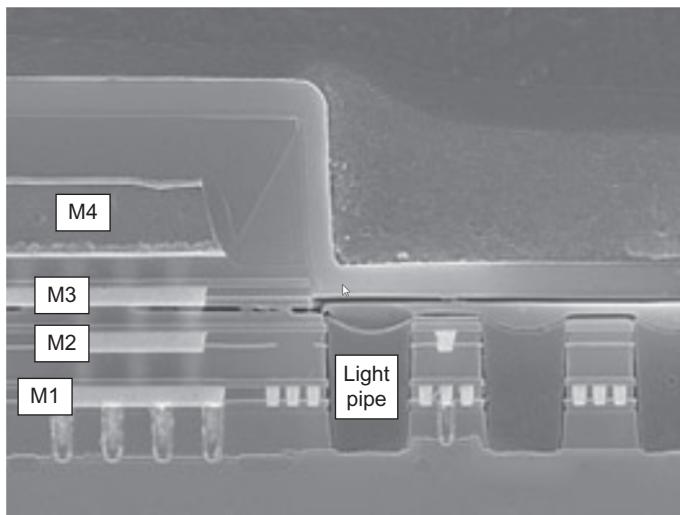


**Fig. 4.3** Illustration of the light pipe concept. Deep vias etched through the inter-metal dielectric layer above each diode. The vias are filled with a slightly higher refractive index than their surroundings. The  $\mu$ -lenses are modified to ensure good coupling of the photons into the light pipe.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

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) (Reznik et al., 2008) or with a special polymer having a high refractive index (Gambino et al., 2007). 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 (Chipworks, 2012). 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 cross talk 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 nonstandard metallization backend processes to decrease the optical path, light guides, and a recessed backend.



**Fig. 4.4** An Aptina sensor taken from a LENOVO phone (courtesy of Chipworks, 2012). 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.

These solutions needed re-optimization over and over again for each new pixel dimension.

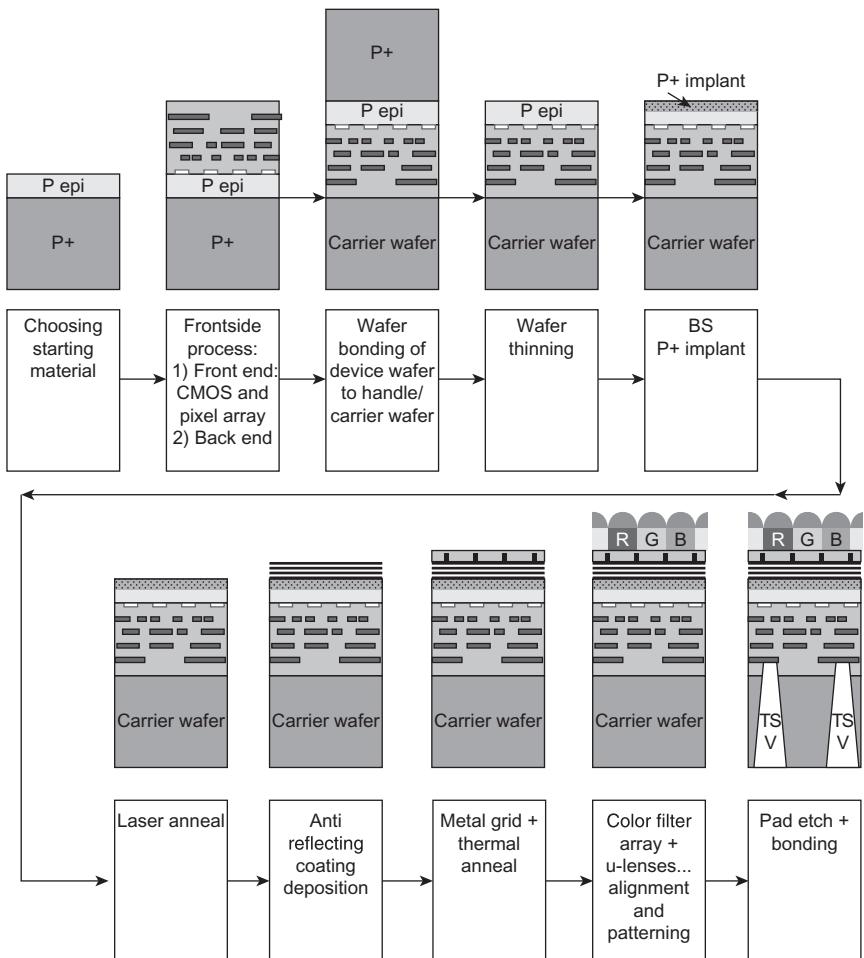
The alternative approach is to try to couple the light from the back side of the sensor. This approach could in principle save the effort of customizing the backend process and allow using advanced CMOS fabrication technology. The other benefit from BSI is an increased QE due to optimization of an antireflecting 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.

## 4.3 Basics of 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 ([Wuu, 2009](#); [Rhodes, 2009](#)). 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 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 is discussed as well.

### 4.3.1 Starting material selection

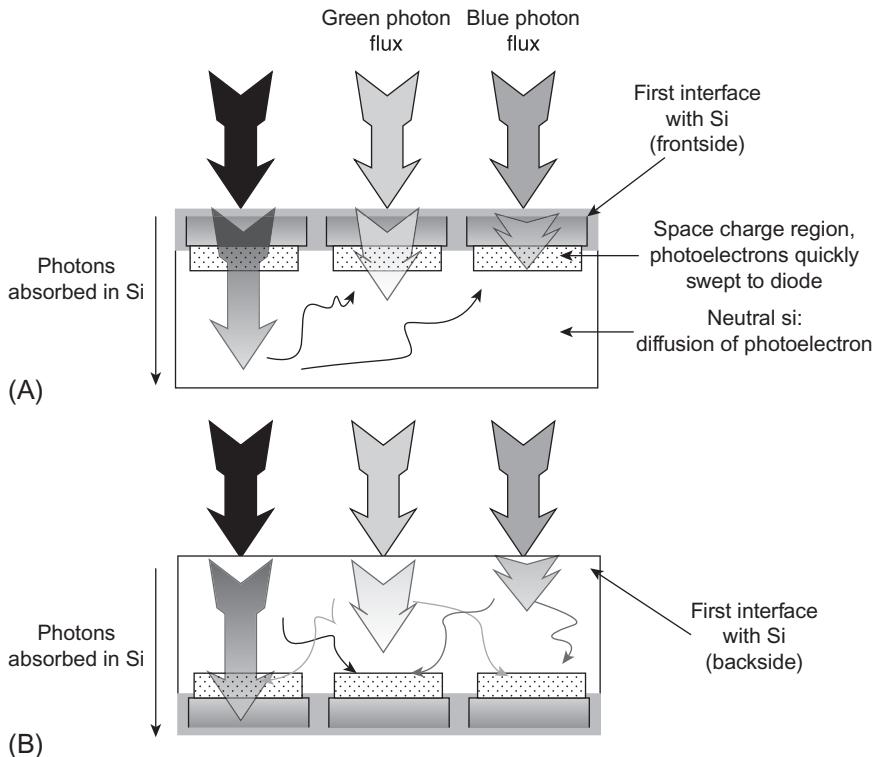
The sensor fabrication process begins by starting material selection. This is a very important technological choice which will have many integration consequences in later fabrication steps. OVT and TSMC chose to start with a P on P+-doped wafer. The P on P+-doped wafer is available with a proven supply chain, relatively cheap and widely used in the CIS industry for FSI devices. The other choice is to use silicon on insulator (SOI) wafers ([Pain, 2009](#)) which were used, for example, by Sony in an early BSI process generation ([Iwabuchi and Maruyama, 2006](#)) and by TowerJazz for devices targeted at high-end markets ([Edelstein et al., 2011](#)). 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 it is discussed in detail later. The main disadvantages are the cost and availability of such wafers, compared to bulk wafers which make it less ([Getman et al., 2007](#)) attractive for high volume production for the mobile phone market.



**Fig. 4.5** Typical BSI complete process flow (inspired by TSMC and OmniVision from IISW 2009 ([Wuu, 2009](#); [Rhodes, 2009](#))).

### 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. The increased electrical cross talk reduces the modulation transfer function (MTF) for monochrome sensors and degrades the post color interpolation SNR ([Yadid-Pecht](#)



**Fig. 4.6** Illustration of photon absorption in FSI pixel compared to BSI.

and Etienne-Cummings, 2004). 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 (Shah et al., 2003; Tournier et al., 2011). In BSI pixels, the problem becomes severe since a significant amount of green photons ( $\lambda=550$  nm,  $\sim 30\%$  absorption in the first  $0.5\ \mu\text{m}$ ) and the blue photons ( $\lambda=470$  nm,  $\sim 70\%$  absorption in the first  $0.5\ \mu\text{m}$ ) are strongly absorbed (Nakamura, 2005; pveducation, 2014).

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 (Michelot et al., 2011) 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 nonlinear response at low illumination levels. An alternative similar solution is to use a deep P+ implant to try to isolate the diodes (Prima et al., 2007). 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

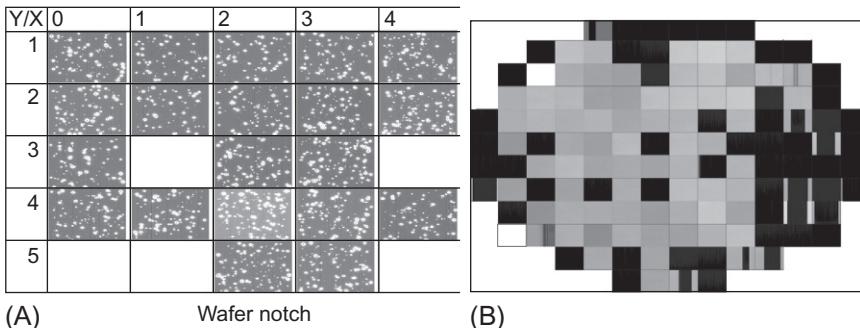
blemish. As is very well known, proper anneals require a long, high-temperature process that is not compatible with modern CMOS technology nodes.

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 were solved in the early days of silicon imagers was gettering of the metal elements during fabrication (Prima et al., 2007; Reiner and Coppens, 1978). 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 (Teranishi, 2012; Chepis et al., 1997). 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 (Capedelli, Di Turi, and Faralli, 2006a).
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 (Uya et al., 2010).

We briefly point out two examples of these techniques which 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 (Takizawa, 2009). Sony suggests using gettering centers on the SOI substrate part. A path for the metal contamination in the active epitaxial layer to the silicon is created by damaging the SOI oxide layer which separates the epitaxial from the Si substrate. Implementation of such a technique requires very strong cooperation of many parties including SOI and epitaxial 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 (Capedelli et al., 2006b). 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 (Tournier et al., 2011) by ST Microelectronics in their 1.4  $\mu\text{m}$  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



**Fig. 4.7** Dark image taken from a full wafer with BSI sensor manufactured by TowerJazz (A) gray image wafer map for early process version with no proper gettering technique and (B) gray image wafer map for process which includes gettering technique.

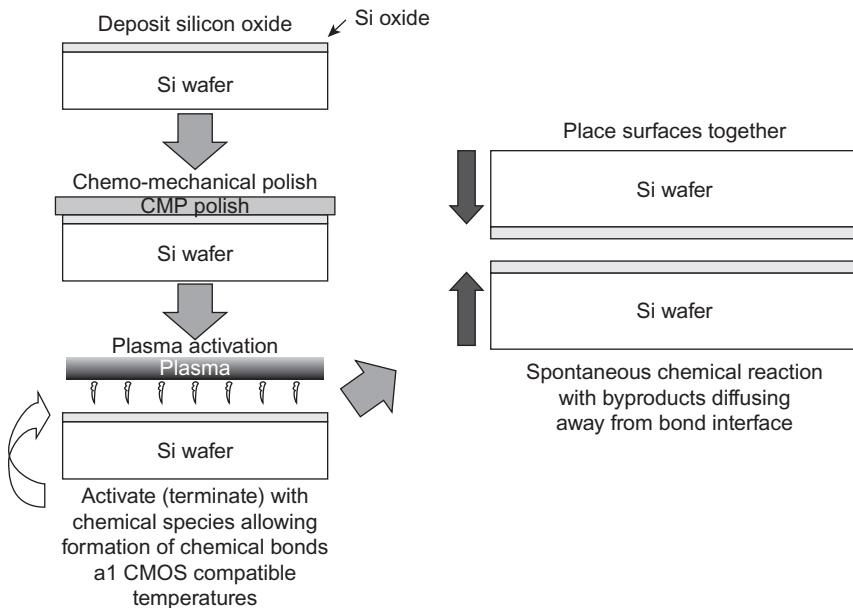
subject, in Fig. 4.7, gray level images (50% saturation at 30 fps) taken at wafer level for two different BSI products which ran at TowerJazz on SOI starting material are illustrated. In Fig. 4.7A the process does not include a special gettering process and exhibits many bright points. Clean images can be seen in Fig. 4.7B 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 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 (Gaudin and Riou, 2010; Lagahe-Blanchard and Aspar, 2009; Blanchard and Radu, 2011). Fig. 4.8 illustrates a modern wafer to wafer bonding technique known as direct bonding<sup>TM</sup> from Ziptronix (Blanchard and Radu, 2011; Enquist, 2012). 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



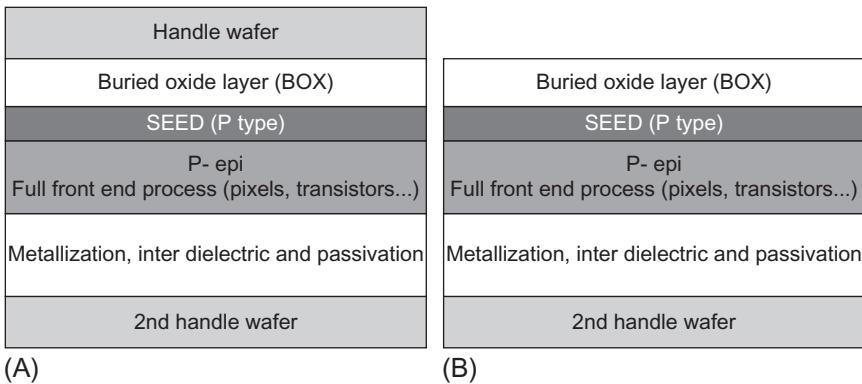
**Fig. 4.8** Modern wafer-to-wafer bonding (ZiBona process flow).

Taken from Ziptronix presentation (Direct Bonding TM).

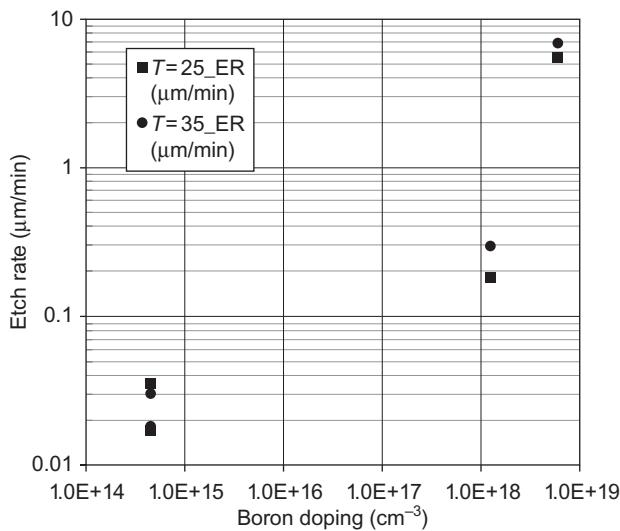
being chosen, the wafer thinning is quite simple and involves removing the silicon and stopping on the BOX. Immediately after, the BOX itself is removed (now the active silicon is used as an etch stop) and the wafer is ready for the next steps of fabrication. These steps are well described in Fig. 4.9.

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 and 3  $\mu\text{m}$ , for pixel sizes under 2  $\mu\text{m}$ , so the thinning of a P/P+ wafer usually starts with mechanical grinding from initial thickness of about 700  $\mu\text{m}$  down to 20  $\mu\text{m}$ , and only then comes the special etch, that 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  $10^{18} \text{ cm}^{-3}$  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 margins 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



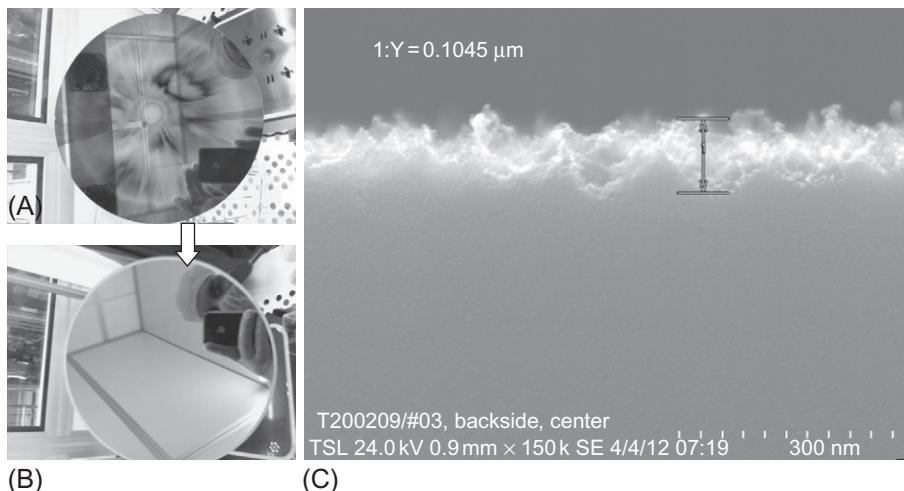
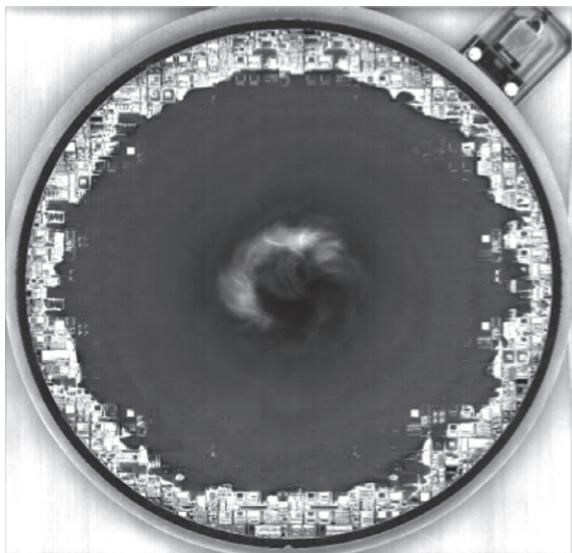
**Fig. 4.9** (A) SOI wafer after wafer bonding and flipping and (B) SOI wafer after removing the first handling wafer and stopping on the SOI bottom oxide.



**Fig. 4.10** Etch rates differences with respect to the doping level of a bare silicon wafer.

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 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

**Fig. 4.11** A 200 mm P/P+ wafer after HNA etch without edge protection. The P active layer thickness is 5  $\mu\text{m}$  and the HNA process started after grinding of the silicon down to 20  $\mu\text{m}$ . At the edge of the wafer there is no active silicon left and the first inter-dielectric layer of the metallization is fully exposed.



**Fig. 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.

edge requires nonstandard exposure since the photoresist 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  $\pm 0.1 \mu\text{m}$  achieved on 200 mm bulk wafers.

#### **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 ([Pain, 2009](#)). We discuss briefly these three common methods.

##### **4.3.5.1 Laser annealing**

This method involves the following steps:

- The backside of the wafer is implanted by shallow p-type implant (for n-type PD 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 nanosecond ( $<200 \text{ ns}$ ). The heat generated by the pulse is enough for local annealing of implant damages and activating of the traps ([Huet et al., 2011](#)).
- Since the pulse is very short, the temperature in the metallization area of the wafer does not change.

This method is widely used for commercial manufacturing of BSI CSI. 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 QE in short wavelengths

##### **4.3.5.2 Passivation using low-temperature growth of highly doped epitaxial-silicon layers**

This method was invented by the California Institute of Technology (CALTECH) and Jet Propulsion Labs (JPL) ([Huet et al., 2011](#); [Hoek et al., 1992](#)) and was widely used for scientific applications ([Pain et al., 2005](#)) of BSI close-coupled device (CCD) and

CMOS sensors. The main advantage of this method is the high QE 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.

#### ***4.3.5.3 Passivation using dielectric with negative fixed charge***

Recently ([Meynarts, 2010](#)), Guy Meynarts of CMOSIS, suggested using  $\text{Al}_2\text{O}_3$  as a passivation layer. The  $\text{Al}_2\text{O}_3$  is deposited over  $\text{SiO}_2$  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  $\text{Al}_2\text{O}_3$  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:

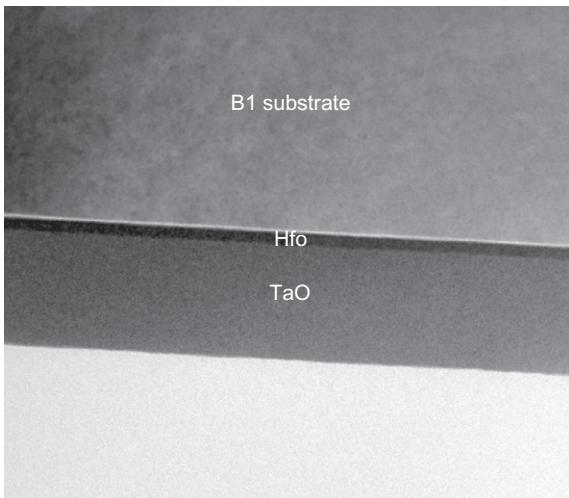
- an antireflecting coating
- a metal grid
- a color filters array (CFA) and  $\mu$ -lenses

#### ***4.3.6.1 Antireflecting 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  $\text{HfO}$ ,  $\text{SiON}$ , and  $\text{SiO}_2$ , but more exotic material systems like  $\text{MgF}_2$  and  $\text{Ta}_2\text{O}_5$  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](#).

#### ***4.3.6.2 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 nonuniformity (PRNU).



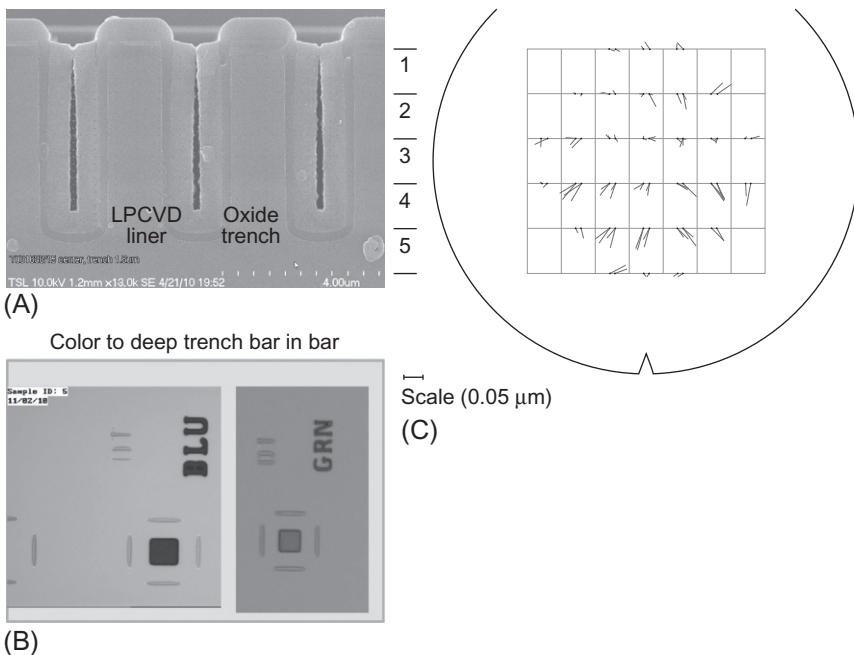
**Fig. 4.13** Antireflecting coating example of 1.55  $\mu\text{m}$  generation from Sony BSI. Courtesy of Chipworks (Sony, 2012. ISX014, CMOS Image Sensor (SOC) With Built-in High-Performance Image-Processing Engine and Camera System Functions. <http://www.sony.net/Products/SC-HP/cx-news/vol70/pdf/isx014.pdf>).

#### 4.3.6.3 Color filters array (CFA) and $\mu$ -lenses

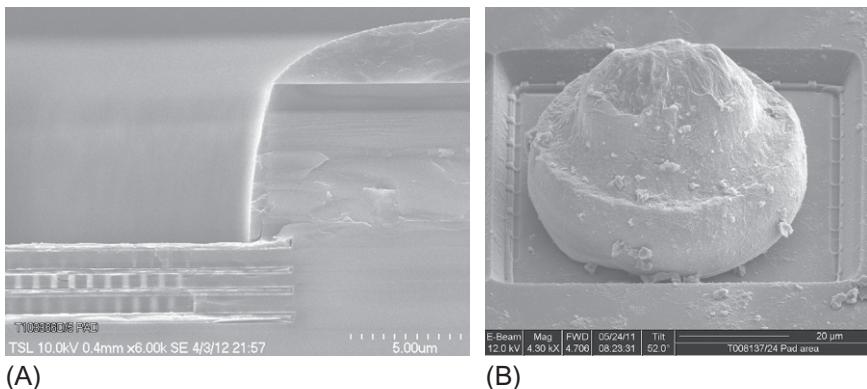
Color filter array and  $\mu$ -lenses patterning is very similar to the front-end case. The main difficulty here is to align the optical stack and the metal grid to the pixel active diode area. Typically, the Si thickness after thinning may be over 2  $\mu\text{m}$  for a high volume small pixel sensor and over 5  $\mu\text{m}$  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 BSI sensor to the outside world are discussed here. 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 to be used by some mass market players (Fontaine, 2011), 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 antireflective dielectric coating, through the epitaxial active layer and the first frontside inter-dielectric, stopping at the M1 pads as shown in Fig. 4.15.

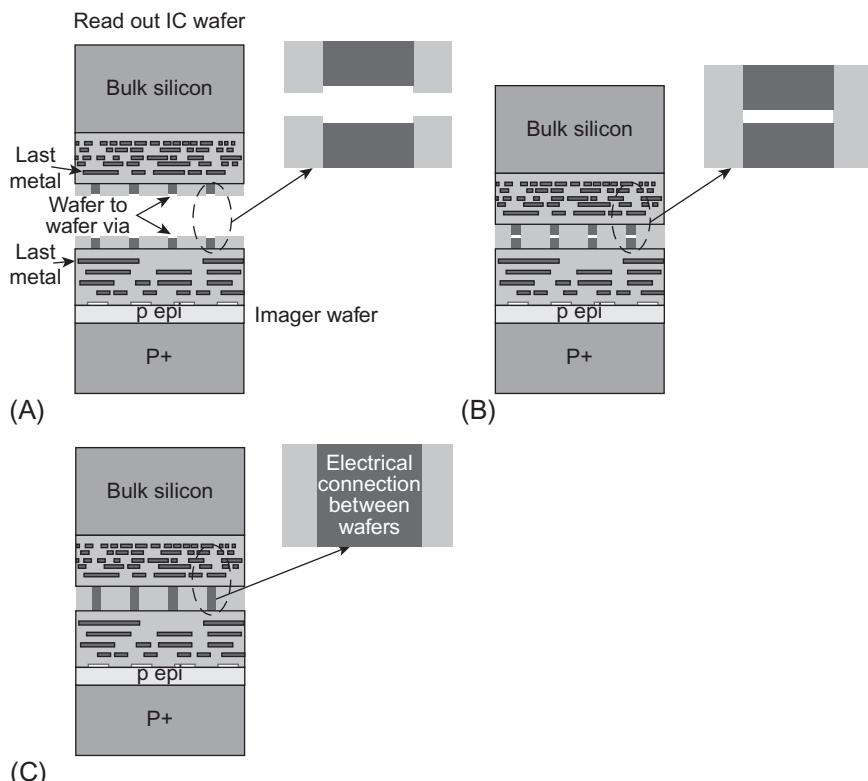


**Fig. 4.14** Alignment of the color filter array to the frontend using deep trenches (A) SEM pictures of the deep trenches alignment mark, (B) *blue* and *green* color filter alignment to the deep trench mark, and (C) typical wafer map which measure the misalignment between the deep trench mark and the color filter.



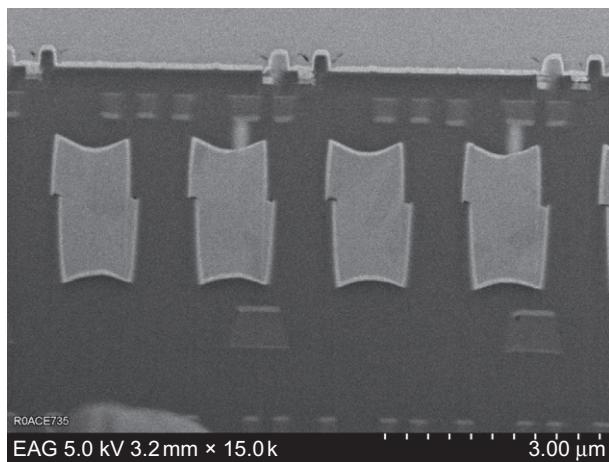
**Fig. 4.15** Pad opening scheme (A) cross section showing deep etch from the sensor backside which stops at M1 of the pad structure and (B) arbitrary unit bonding in the center of the backside pad.

A discussion on the most advanced method for interfacing a BSI 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 wire bond as described above. The process of electrically bonding the two wafers is schematically described in Fig. 4.16. Sony has already announced (Sony, 2012) an image sensor chip which is using some readout circuits in the carrier



**Fig. 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.

**Fig. 4.17** Cross section of a wafer-to-wafer electrical test structure with  $2\text{ }\mu\text{m}$  separation between the vias. Courtesy of Ziptronix (Huet, F. R. K., Boniface, C., Venturini, J., Ait Fqir Ali-Guerry, Z., Beneyton, R., Marty, M., Dutartre, D., 2011. High performance and high yield junction formation with full device exposure laser thermal annealing. In International Image Sensor Workshop).



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  $2\text{ }\mu\text{m}$  separation is shown in Fig. 4.17. Some prototypes of a 3-Meg imager were reported by Ziptronix ([Ziptronix, 2013](#)). 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 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 ([Lahav and Fenigstein, 2008](#)). 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 ([Yang and El Gamal, 1999](#)).

## 4.5 Conclusion

In this chapter we reviewed some of the major challenges and trends in the fabrication of BSI sensors. We started by describing the advantages of a BSI sensor over an FSI sensor. We showed that the possibility to use advanced materials to create a highly optimized antireflecting coating is the main driving force for BSI technology for scientific and high-end applications. On the other hand, for pixels sizes  $<2\text{ }\mu\text{m}$  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.

## References

- J. Adkisson, J. Gambino, Z. He, 2010. CMOS Imager Array With Recessed Dielectric. US Patent 7781781B2.
- Agranov, G., Gilton, T., Mauritzson, R., Boettiger, U., Altice, P., Shah, J., Ladd, J., Fan, X., Brady, F., McKee, J., Hong, C., Li, X., Patrick, I., 2005. Optical-electrical characteristic of small, sub-4  $\mu\text{m}$  and sub-3  $\mu\text{m}$  pixels for modern CMOS image sensors. In: IEEE CCD and AIS Workshop, pp. 206–209.
- Agranov, G., Mauritzson, R., Ladd, J., Dokoutchaev, A., Fan, X., Li, Z., Yin, R., Lenchikov, V., Nagaraja, S., Gazeley, W., Bai, J., Lee, H., D'Anna, A., De-Amicis, G., 2009. Pixel continues to shrink. Pixel development for novel CMOS image sensors. In: International Image Sensor Workshop.
- Blanchard, C., Radu, I., 2011. Smart stacking TM technology: An industrial solution for 3D layer stacking. In: EEE International Conference on IC Design Technology.
- R. Capedelli, L. Di Turi, D. Faralli, 2006a. Process for Manufacturing a SOI Wafer With Improved Gettering Capability. US Patent App. 11/347, 801.
- R. Capedelli, et al. 2006b. 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).
- D. J. S. Chepis et al., 1997. Method and Structure for Front-Side Gettering of Silicon-on-Insulator Substrates. US Patent US005646053A.
- Chipworks, 2012. <http://www.chipworks.com/blog/technologyblog/2012/05/23/a-survey-of-recent-image-sensor-pixel-structures/>. (Accessed 29 January 2013).
- Cohen, M., Roy, F., Herault, D., Cazaux, Y., Gandolfi, A., Reynard, J.P., Cowache, C., Bruno, E., Girault, T., Vaillant, J., et al., 2006. Fully optimized Cu based process with dedicated cavity etch for 1.75  $\mu\text{m}$  and 1.45  $\mu\text{m}$  pixel pitch CMOS image sensors. In: International Electron Devices Meeting, 2006. IEDM'06, pp. 1–4.
- Edelstein, D., Heidenreich, J., Goldblatt, R., Cote, W., Uzoh, C., Lustig, N., Roper, P., McDevitt, T., Motsiff, W., Simon, A., Dukovic, J., Wachnik, R., Rathore, H., Schulz, R., Su, L., Luce, S., Slattery, J., 1997. Full copper wiring in a sub-0.25  $\mu\text{m}$  CMOS ULSI technology. In: International Electron Devices Meeting, 1997, IEDM '97. Technical Digest, pp. 773–776.
- Edelstein, R., Katz, O., Lavi, B., Aberman, I., Rosenthal, S., Shadmi, M., Arad, S., Golan, N., Shach Caplan, M., Massalha, M., Lagahe-Blanchard, C., Marinier, L., Fontanière, R., Castex, A., Broekaart, M., Martinez, M., Milhet, N., Rigny, A., Pelissier, C., 2011. Process integration aspects of back illuminated CMOS imagers using smart stacking TM technology with best in class direct bonding. In: International Image Sensor Workshop.
- Enquist, P., 2012. 3D integration applications for low temperature direct bond technology. In: Low Temperature Bonding for 3D Integration (LTB-3D), 2012 3rd IEEE International Workshop on IEEE.

- Fontaine, R., 2011. A review of the 1.4  $\mu\text{m}$  pixel generation. In: International Images Sensor Workshop.
- Gambino, J., Ackerson, K., Guthrie, B., Leidy, B., Abadeer, W., Mongeon, S., Meatyard, D., Adkisson, J., Rassel, R.J., Ellis-Monaghan, J., Jaffe, M., 2007. Reliability of CMOS image sensor with polymer lightpipe. In: International Image Sensor Workshop.
- Gambino, J., Leidy, B., Adkisson, J., 2006. CMOS imager with copper wiring and lightpipe. In: International Electron Devices Meeting (IEDM).
- Gaudin, G., Riou, G., 2010. 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.
- Getman, A., Uvarov, T., Han, Y.I., Kim, B., Ahn, J.C., Lee, Y.H., 2007. Crosstalk, color tint and shading correction for small pixel size image sensor. In: International Image Sensor Workshop, pp. 166–169.
- Hoenk, M.E., Grunthaner, P.J., Grunthaner, F.J., Terhune, R.W., Fattah, M., Tseng, H.F., 1992. Growth of a delta-doped silicon layer by molecular beam epitaxy on a charge-coupled device for reflection-limited ultraviolet quantum efficiency. *Appl. Phys. Lett.* 61 (9), 1084–1086.
- Huet, F.R.K., Boniface, C., Venturini, J., Ait Fqir Ali-Guerry, Z., Beneyton, R., Marty, M., Dutartre, D., 2011. High performance and high yield junction formation with full device exposure laser thermal annealing. In: International Image Sensor Workshop.
- Iwabuchi, S., Maruyama, Y., 2006. A back-illuminated high-sensitivity small-pixel color CMOS image sensor with flexible layout of metal wiring. In: Solid-State Circuits Conference, ISSCC 2006.
- Lagahe-Blanchard, C., Aspar, B., 2009. Wafer stacking: Key technology for 3D integration. In: 2009 IEEE International SOI Conference, pp. 1–4.
- Lahav, A., Fenigstein, A., 2008. The megapixel race: a chip designer's point of view. *EDN (Electrical Design News)* 53 (12), 57.
- G. Meynants, 2010. Method of Manufacture of a Backside Illuminated Image Sensor. US Patent App. 12/939, 599.
- Michelot, J., Roy, F., Prima, J., Augier, C., Barbier, F., Ricq, S., Boulenc, P., Essa, Z., Pinzelli, L., Gatefait, M., Broquin, J., 2011. Back illuminated vertically pinned photodiode with in depth charge storage. In: International Image Sensor Workshop.
- Nakamura, J., 2005. Image Sensors and Signal Processing for Digital Still Cameras. CRC.
- Pain, B., 2009. Backside illumination technology for SOI-CMOS image sensors CMOS image sensors. In: Symposium on Backside Illumination of Solid-State Image Sensors.
- Pain, B., Cunningham, T., Nikzad, S., Hoenk, M., 2005. A back-illuminated megapixel CMOS image sensor. In: IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors.
- Prima, J., Roy, F., Coudrain, P., Gagnard, X., Segura, J., Cazaux, Y., Herault, D., Virolet, N., Moussy, N., Giffard, B., Gidon, P., 2007. A 3 mega pixel back-illuminated image sensor in 1T5 architecture with 1:45  $\mu\text{m}$  pixel pitch. In: International Image Sensor Workshop. no. 1.
- pveducation, 2014. Optical-properties-of-silicon. <http://pveducation.org/pvcdrom/appendices/optical-properties-of-silicon> (Accessed 4 February 2014).
- Reiner, T., Coppen, P.J., 1978. Characterization of leakage current with phosphorus gettering. In: Proceedings of 1978 International Conference on the Application of Charge Coupled Devices.
- Reznik, H., Edelstein, R., Shach-Caplan, M., Dulberg, F., Kamenetsky, V., Karaste, K., Rantala, J.T., 2008. Integration and characterization of spin on dielectric materials in image sensor devices. In: Proceedings of SPIE.

- H. E. Rhodes, 2008. Local Interconnect Structure for a CMOS Image Sensor and Its Manufacturing Method. US Patent 7345330.
- Rhodes, H., 2009. Mass production of BSI image sensors: performance results. In: Symposium on Backside Illumination of Solid-State Image Sensors.
- Shah, J., Jernigan, M.E., Hornsey, R.I., 2003. Characterization and deblurring of lateral crosstalk in CMOS image sensors. *IEEE Trans. Electron Devices* 50 (12), 2361–2368.
- Sony, 2012. ISX014, CMOS Image Sensor (SOC) With Built-in High-Performance Image-Processing Engine and Camera System Functions. [http://www.sony.net/Products/SC-HP/cx\\_news/vol70/pdf/isx014.pdf](http://www.sony.net/Products/SC-HP/cx_news/vol70/pdf/isx014.pdf).
- Takizawa R. 2009. 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/US810199?pg=PA1&dq=Takizawa+image+pickup+soi&f=false](http://www.google.com/patents/US810199?pg=PA1&dq=Takizawa+image+pickup+soi&hl=en&sa=X&ei=RgMIUavLO6mC4gS8pYCQAg&ved=0CDMQ6AEwAA#v=onepage&q=Takizawa+image+pickup+soi&f=false) (Accessed 29 January 2013).
- Teranishi, N., 2012. Dark current and white blemish. In: Image Sensors.
- Tournier, A., Leverd, F., Favenne, L., Perrot, C., Pinzelli, L., Gatefait, M., Cherault, N., Jeanjean, D., Carrere, J.-P., Hirigoyen, F., Grant, L., Roy, F., 2011. Pixel-to-pixel isolation by deep trench technology: Application to CMOS image sensor. In: International Image Sensor Workshop.
- S. Uya et al, 2010. Backside Illuminated Imaging Device, Semiconductor Substrate, Imaging Apparatus and Method for Manufacturing Backside Illuminated Imaging Device. US Patent 7781715 (Accessed 24 October 2010).
- Wuu, S., 2009. BSI technology with bulk Si wafer. In: Symposium on Backside Illumination of Solid-State Image Sensors.
- Yadid-Pecht, O., Etienne-Cummings, R., 2004. CMOS Imagers: From Phototransduction to Image Processing. Springer.
- Yang, D., El Gamal, A., 1999. Comparative analysis of SNR for image sensors with enhanced dynamic range. In: Proceedings of SPIE.
- Ziptronix, 2013. <http://www.ziptronix.com/> (Accessed 4 January 2013).

This page intentionally left blank

# CMOS circuits for high-performance imaging

5

Bhaskar Choubey<sup>a</sup>, Waqas Mughal<sup>b</sup>, Luiz Gouveia<sup>c</sup>

<sup>a</sup>Chair of Analogue Circuits and Image Sensors, Siegen University, Siegen, Germany, <sup>b</sup>School of Engineering, University of Glasgow, Glasgow, United Kingdom, <sup>c</sup>IMSE-CNM - Seville Institute of Microelectronics, Sevilla, Spain

Complementary metal-oxide semiconductor (CMOS) image sensors were initially projected to be low cost but poor image quality replacement for charge-coupled devices (CCDs) for imaging systems. However, research and development, particularly in process refinements, as well as novel interface circuits have significantly improved the performance of these devices to a point where they are found in many high-performance imaging applications such as biomedical imaging, high-end digital cameras, or scientific instrumentation. In this chapter, we will review circuits used in some of these high-performance applications. This chapter covers broad areas of imaging performance.

In the first section, we review approaches aimed at increasing the number of pixels and hence improving the resolution of the image sensor. CMOS imaging historically lagged the CCDs on account of their poor noise performance. Section 5.2 studies various noise sources and the approaches followed to reduce them. Most commercial image sensors require video frame rate abilities. However, scientific imaging often requires frames rates of thousands of frames per second. Furthermore, with increasing number of pixels, the speed of the sensor as well as the architecture of its interconnects have to be enhanced to meet the state-of-the-art most common specifications. Techniques to achieve this are being studied in Section 5.3. This section also provides a brief analysis of circuits required to achieve global-shutter readout and an introduction on time-delay-integration (TDI) sensors. One of the largest markets for image sensors is that of hand-held devices that also includes mobile phones, wherein battery life is of paramount importance, that is, the image sensor should consume as low power as possible. Hence, Section 5.4 studies techniques developed to reduce the power consumption of image sensors. The limited dynamic range (DR) 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 number of other high-performance sensors including those with large formats as well as low-light sensitivity.

## 5.1 High-resolution image sensors

CMOS image sensor designers have been pushed by the market forces to increase the spatial resolution in order to improve the image quality as well as to enable the digital zoom. The latter is of particular significance in mobile cameras, wherein 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 of the sensor resolution. For example, HDTV 1080p has a resolution of about two million pixels ( $1920 \times 1080$ ). Similarly, new ITU (International Telecommunication Union) approved high-definition standards UHDTV-1 (4K HDTV) and UHDTV-2 (8K HDTV) aim to transmit 8.3 and 33.2 million pixels, respectively (Kitamura et al., 2011; Huang et al., 2009).

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 optics and the image sensor, can be characterized using different measurement methods that include the use of resolution test charts and the modulation transfer function (MTF) (Holst, 2007). 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 sensor size, are involved in this race.

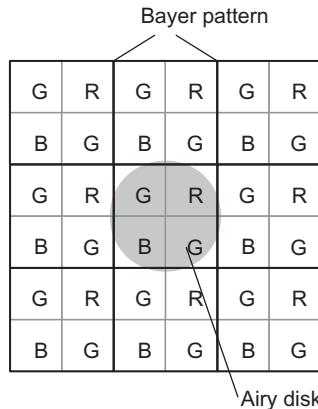
### 5.1.1 Pixel size limits

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 lens 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

$$D_A = 2.44\lambda F/\# \quad (5.1)$$

where  $\lambda$  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  $3.7 \mu\text{m}$ . 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 micrometer (Chao et al., 2011). It must be noted, however, that this reduced dimension can provide opportunities for oversampling and compensation for the Bayer pattern of color filter, normally used to obtain color 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 colors, the distance between color is twice the pixel size as shown in Fig. 5.1. However, the optical resolution is not



**Fig. 5.1** Bayer pattern as typically used in CMOS image sensor. Each “color pixel” is built from one “red,” one “blue,” and two “green” pixels. The airy disk covers the Bayer pattern of minimum size pixels.

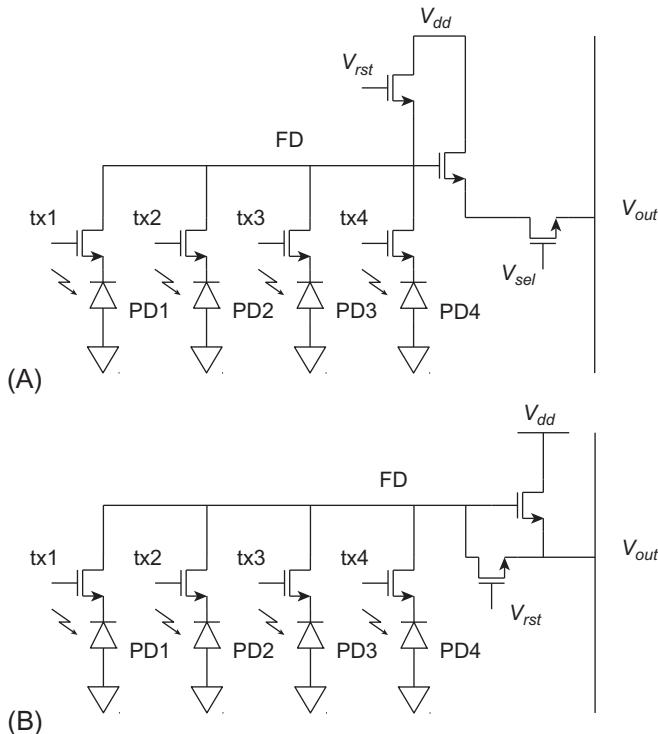
the only limiting factor for the pixel size. In general, reducing the pixel size degrades the capacity of the photodetector to store charge. This leads to poor sensitivity and therefore, poor signal-to-noise ratio (SNR) and DR (Chen et al., 2000).

### 5.1.2 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 nondetector components with neighbor pixels. In a 4T 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) (Kitamura et al., 2011), 1.75T (four photodiodes) (Kim et al., 2006) as in Fig. 5.2, or 1.5T (six photodiodes). The row-selection transistor can also be eliminated if the configuration in Fig. 5.2B is used (Takahashi et al., 2004). In these pixels, all floating diffusions (FDs) voltages rather than the ones on the selected row are set to a low value by setting  $V_{rst}$  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.

Another approach to increase the fill factor of the pixel is to use microlenses and light guides (light pipes). Microlenses 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 photodetector reducing light crosstalk and scattering.

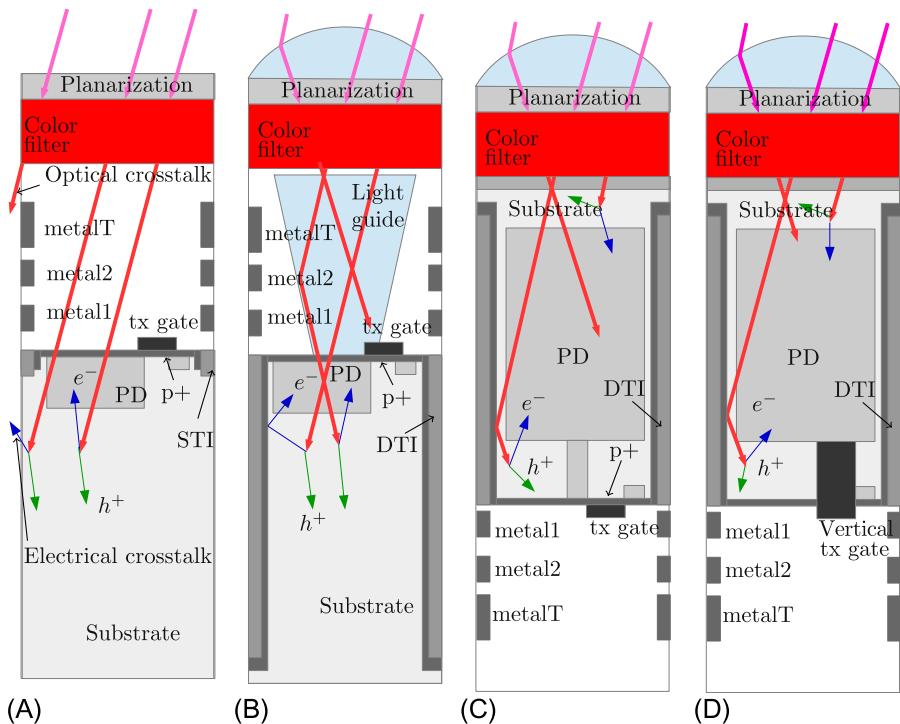
Until recently, most CMOS image sensors were categorized as front side illumination (FSI) sensors. In these sensors, the light had to cross the space between metal and polysilicon stacked layers before it reached the photodetector resulting in significant



**Fig. 5.2** 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.

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.

Crosstalk effects exist in both FSI and BSI designs and are defined as the interference of one pixel state on the state of another pixel. Two main sources of crosstalk effects can be categorized as optical or electrical crosstalk. The source of the former is due to light focused on one pixel being partially captured by neighboring pixels and the latter happens when the electron-hole pair generated on one pixel diffuses or drifts to other pixels. Optical crosstalk can be minimized using the microlenses, light guides as above as well metallic and stacked grids (Cheng et al., 2015). For electrical crosstalk, better interpixel isolation can be achieved by using techniques as dopant implants, deep trench isolation (DTI) (Tournier et al., 2011), and metal-oxide-semiconductor (MOS) capacitor deep trench isolation (CDTI) (Nguyen et al., 2016). BSI also allows the implementation of vertical transfer gate (VTG) (Roy et al., 2017) which increases the well capacity of the pixel, increasing its DR (Fig. 5.3).



**Fig. 5.3** Frontside illumination (FSI) and backside illumination (BSI) pixels cross-section. (A) Conventional CMOS FSI pixel, (B) CMOS FSI pixel with microlenses, light guide, and deep trench isolation (DTI), (C) CMOS BSI pixel with deep photodiode (PD), and (D) CMOS BSI pixel with vertical transfer gate.

An alternative technique to enhance pixel count without physically shrinking the pixel is to capture all three required wavelengths for color 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 (Lyon and Hubel, 2002). Photons corresponding to blue color 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 color photons. Therefore, a pixel can be built with three photodiodes stacked on the wafer; one for each color. However, such pixels require a nonstandard CMOS fabrication process and impurities deposition on specific depths.

### 5.1.3 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

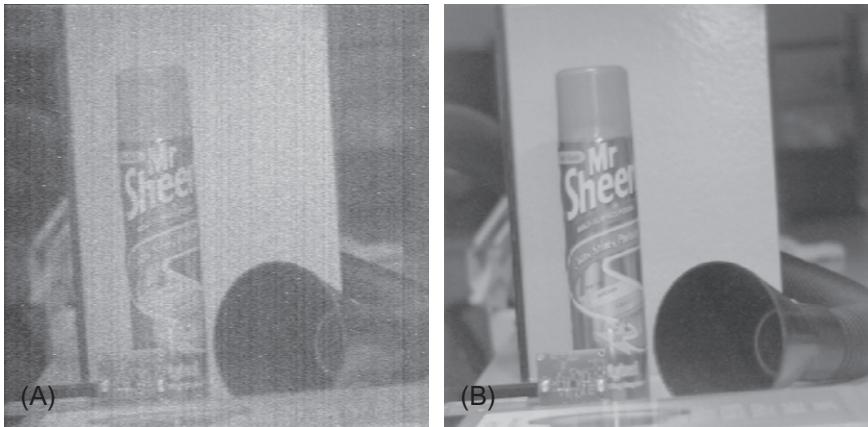
the capability of cropping an image to a smaller region of interest. This is easily performed on CMOS image sensors due to their flexible addressing scheme. The optical resolution is the same as for the whole image, but it allows for smaller images and the capture of more frames per second. Increasing 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  $2 \times 2$  binning would increase the well capacity nearly fourfold at the expense of equivalent resolution losses.

## 5.2 Low-noise 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 (Tian et al., 2001). 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, digitization of the pixel response also introduces quantization 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 nontemporal 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 FPN in images acquired by these detectors (Gamal et al., 1998; Joseph and Collins, 2001). Fig. 5.4 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 intercolumn variations are also visible in the image. A prominent noise observable is that of columnar strips. More importantly, human subjects have better perception of FPN than temporal noise (Driggers et al., 2001) and it is, therefore, of paramount importance to reduce this to less than the contrast ability of the human eye of 1%.



**Fig. 5.4** (A) An image showing the effects of the FPN captured by Fuga camera from fill factory and (B) the same image after correction for FPN.

### 5.2.1 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. 5.5. 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 ([Choubeyp and Collins, 2008](#)) and an optical term produced by the photon falling on diode surface ([Joseph and Collins, 2001](#)):

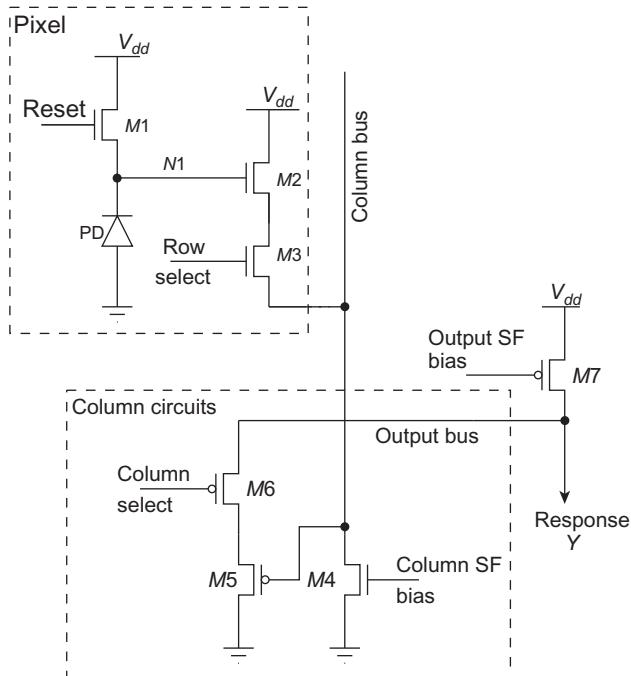
$$I_{PD} = I_{dark} + Q_D G_A A_D G_L L_o \quad (5.2)$$

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 nonphotosensitive 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  $I_{dark}$  depends on the diode area and often its geometry. Therefore, any variation in the photodiode area leads to FPN, which will affect the performance of the image sensors in low-intensity images. These variations are often referred to as the dark current nonuniformity (DCNU) in order to differentiate it from intensity-dependent FPN. It is also worth noting that the diode current is temperature dependent, which leads to changes in FPN values with temperature and in general, the FPN is expected to increase with any increase in temperature ([Joseph and Collins, 2009](#)).

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 ([Tian et al., 2001](#))

$$S_I(f) = q i_D [\text{A}^2/\text{Hz}] \quad (5.3)$$



**Fig. 5.5** A three-transistor CMOS active pixel sensor (3T APS) pixel with a typical column circuitry to enable two-stage addressing mechanism ([Joseph and Collins, 2001](#)).

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 with this noise. The reset noise power obtained can be

$$V_n^2 = kT/C_D \quad (5.4)$$

where  $k$  is the Boltzmann constant,  $T$  is the absolute temperature, and  $C_D$  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  $kT/2C_D$ . 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 ( $M_6$ ). Using a first-order model, the output of the image sensor after two stages of readout can be expressed as ( Joseph and Collins, 2001; Choubey, 2012):

$$\begin{aligned} V_Y = & V_{dd} - V_{T,M1} - \frac{t_{int}}{C_D} (I_{dark} + Q_D G_A A_D G_L L_o) \\ & - \sqrt{\frac{\beta_{M4}}{\beta_{M2}}} (V_{GS,M4} - V_{TM4}) - \{V_{T,M2} + V_{T,M5}\} \\ & - \sqrt{\frac{\beta_{M7}}{\beta_{M5}}} (V_{GS,M7} - V_{TM7}) \end{aligned} \quad (5.5)$$

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

Terms related to transistor  $M_7$  are common to all pixels, while those related to transistors  $M_4$ – $M_6$  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 MOS channel (Pelgrom et al., 1989). 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, of the transistors have been observed to be Gaussian distribution with zero mean and a variance dependent on the device area (Pelgrom et al., 1989):

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

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

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

From Eq. (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  $M_4$ – $M_6$  are shared by all pixels in a column and therefore, they will only introduce column-to-column noise. With column parallel analog-to-digital converters (ADCs) used to increase the speed of operation of image sensors, mismatch due to column readout circuits are removed; however, analog 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, analog 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 (Joseph and Collins, 2009; Lin et al., 2010).

### 5.2.2 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 (Loose et al., 2001; Ricquer and Dierickx, 1995) with limited results. Consequently, postprocessing 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 modeled with an additive and a multiplicative term:

$$Y_i = O_i + G_i X_i \quad (5.8)$$

Here, the subscripts identify an individual element of the array and the parameters can be identified using Eq. (5.5) as:

$$\begin{aligned} O_i &= V_{DD} - V_{T,M1} - \frac{t_{int}}{C_{PD}}(I_{dark}) \\ &- \sqrt{\frac{\beta_{M4}}{\beta_{M2}}}(V_{GS,M4} - V_{TM4}) - \{V_{T,M2} + V_{T,M5}\} \\ &- \sqrt{\frac{\beta_{M7}}{\beta_{M5}}}(V_{GS,M7} - V_{T,M7}) \end{aligned} \quad (5.9)$$

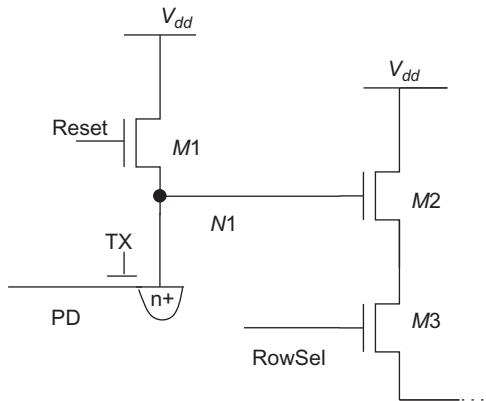
$$G_i = \frac{t_{int}}{C_D} Q_D G_A A_D G_L \quad (5.10)$$

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 regard 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) (Mendis et al., 1997). 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. 5.6.

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

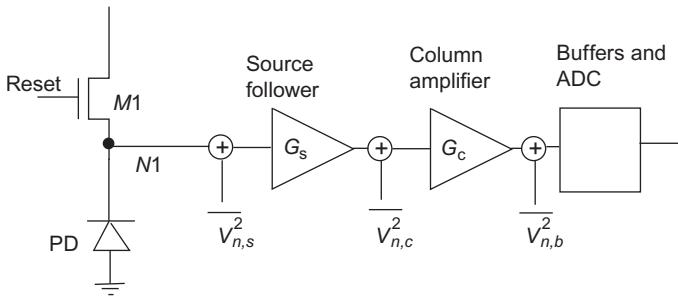
**Fig. 5.6** A four-transistor CMOS APS (4T APS) with buried photodiode to support CDS.



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 analog memory can be used in the column circuits to store these two or they can be directly digitized for simpler difference operation in the digital domain. It is worth noting, however, that with reduction in the feature size of pixels, multiplicative FPN sources have been increasing requiring further correction. This is often achieved using an offline correction in the digital domain.

The CDS, however, is not suitable for pixels requiring continuous operation. Scene-based FPN 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 makes a statistical assumption about the scene and utilizes 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 ([Joseph and Collins, 2001](#)). These techniques, however, require exposure to several scenes and their continuous operation can introduce ghosting after-effects due to expectation of non-stationarity of the objects. Alternatively, global motion between frames captured by an image sensor has also been used to reduce the FPN ([Hardie et al., 2000; Lim and Gamal, 2004](#)). 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 CDS 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 minimized 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



**Fig. 5.7** Noise sources in the signal chain of an image sensor.

digitization. These stages of amplification and digitization 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. 5.7.

The input referred noise can be expressed as:

$$\overline{V_{n,i}^2} = \frac{1}{G_s} \overline{V_{n,i}^2} + \frac{1}{G_c} (\overline{V_{n,i}^2} + \overline{V_{n,c}^2}) \quad (5.11)$$

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 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 subelectron noise floors.

### 5.3 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 (3D) 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 congestion or removing the various bottlenecks in the signal chain of an image sensor.

### 5.3.1 Signal chain

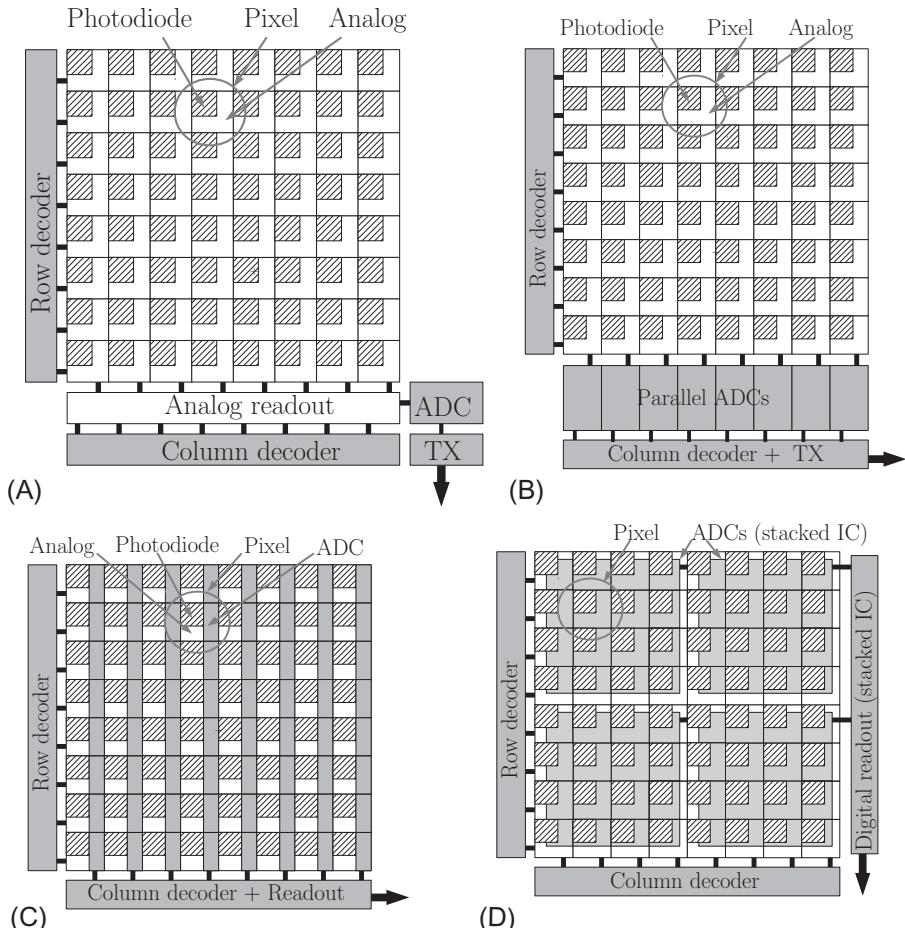
Light falling on several pixels is converted into a train of digital output by an image sensor. Most, if not all image sensors are essentially systems with a large number of parallel inputs and one or few outputs. An ideal and very fast image sensor would function like the human eye where a large number of parallel outputs are recorded from the sensor (the retina). However, with 2D sensor design with limited 3D 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 analog pixel output and digital sensor output. [Fig. 5.8](#) shows various architectures utilized to reduce the effect of this bottleneck.

Early image sensors including CCDs and early CMOS image sensors provided the output through a single analog 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 ([Zhou et al., 1997](#)) as shown in [Fig. 5.8A](#). 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. 5.8B](#). Ideally, one should use a dedicated ADC for each column ([Krymski et al., 2003](#)); however, different ratios of the columns per ADC have been implemented as well ([Takayanagi et al., 2005](#)). 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.

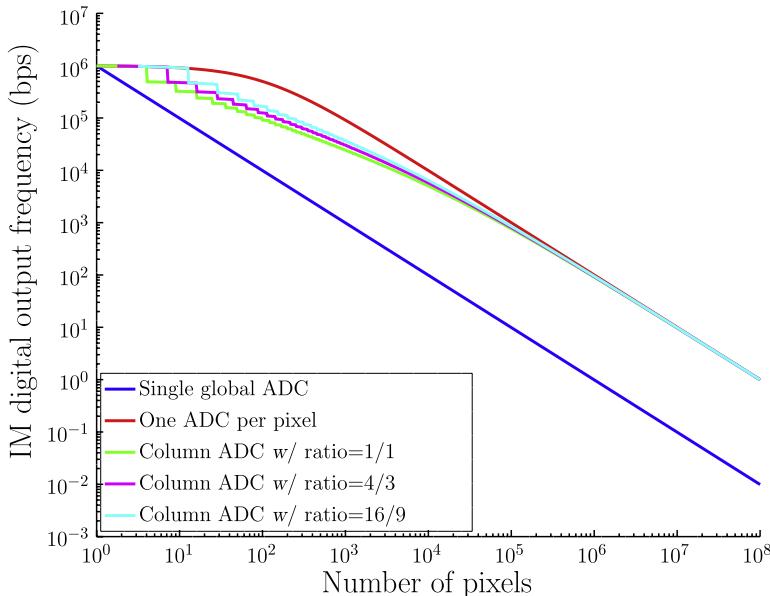
Further improvements on the throughput of all the above architectures can be achieved by allowing for multiple parallel outputs. The sensor array can be split into top and bottom halves with readout circuitry being placed on both sides ([Matsuo et al., 2009](#)). A greater number of outputs can be implemented by partitioning the pixel array



**Fig. 5.8** Different architectures to increase speed. (A) Typical image sensor with one ADC for the whole array and with analog multiplexing of pixel's output. (B) Column parallel design with an ADC per column to reduce the signal delay. (C) Architectures with ADC inside each pixel to further increase speed. (D) Stacked design architecture with an ADC on a different die corresponding to a number of neighbor pixels.

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. 5.8C. This allows for higher throughput since the parasitic load of other pixels is removed from the analog signal chain. However, pixel-level ADCs need to be as small, simple and power starving as possible for area 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



**Fig. 5.9** Comparison of different placement of ADCs on the signal chain. Same ADC conversion and readout times applied. Used expressions from [Leñero-Bardallo et al. \(2014\)](#). For small arrays, the ADC processing time is dominating while the main limiting factor for larger arrays is the readout time for column and pixels ADCs topologies (the difference is in the ratio between ADC processing and readout times).

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.

An alternative approach is to implement ADCs on a stacked IC configuration, allowing for 3D integration. Although the ideal ratio of one ADC per pixel is still impractical for current pixels with few micrometer pitch, some implementations used different ADC/pixel ratios. Sensors were designed where one ADC were to serve a set of 99 ([Kiyoyama et al., 2011](#)), 255 ([Kiyoyama et al., 2013](#))—where only the ADC array chip was implemented—190 ([Takahashi et al., 2017](#)), and 4368 ([Arai et al., 2016](#)) neighboring pixels (Fig. 5.9).

### 5.3.2 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 ADCs topologies have been proposed and used in image sensors. While the specifications change according to the location of the digital conversion in

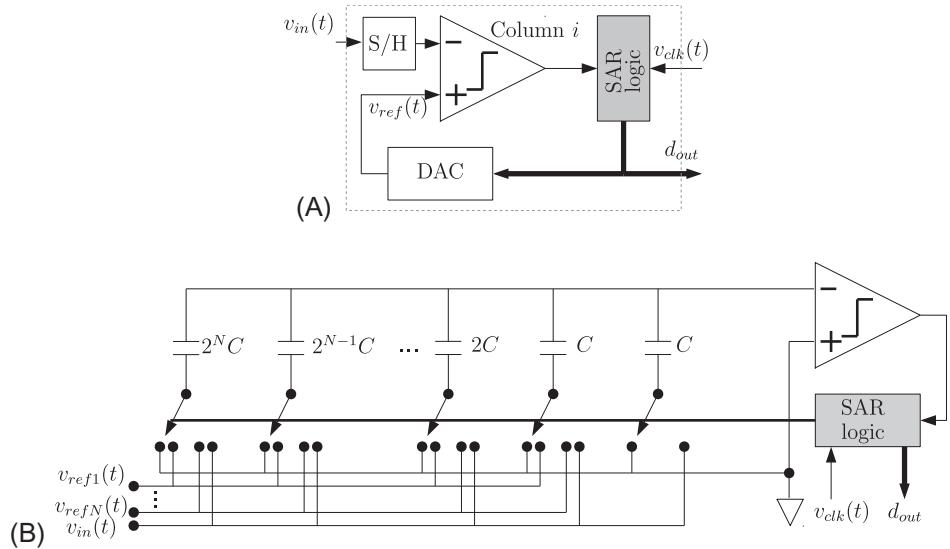
the signal chain, power and area consumption are critical for column and pixel-based ADCs while speed is the main issue in a global ADC.

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. An SAR ADC, as shown in Fig. 5.10A, works by first sampling the pixel output and then changing the comparator reference accordingly. The reference is set using a digital-to-analog converter (DAC) whose input is a digital register. The conversion covers all the  $N_b$  register bits, from the most significant bit to the least significant bit. Therefore, a conversion lasts for approximately  $N_b \times t_{DAC}$  seconds, where  $t_{DAC}$  is the DAC settling time. SAR ADCs can also be used as a column ADC (Krymski et al., 2003), although its large area due to the capacitor scaling, shown in Fig. 5.10B, is a limitation. A 14-bit SAR ADC (with minor modifications) was reported to finish a conversion in 1.7  $\mu$ s (Matsuo et al., 2009). This can be contrasted with the needs of a UHDTV-2 image sensor, which requires every row readout to be completed in 1.9  $\mu$ s.

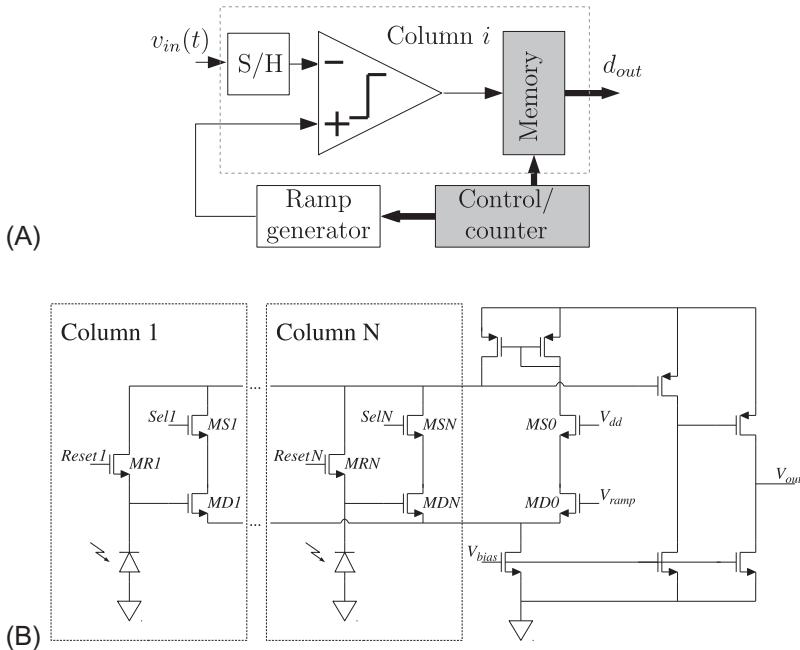
Fig. 5.11A shows yet another popular ADC topology used in image sensors (Wang et al., 2010). The ramp or SS ADCs normally provides a slower conversion rate; however, it requires a small area and generally provides improved linearity characteristics (Takayanagi and Nakamura, 2012). While in the SAR ADC, the reference is adjusted accordingly to the register value; in the SS ADC, the reference applied to the comparator is an SS ramp (Snoeij et al., 2007b). 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 SS ramp. The conversion speed is then a function of the counter's clock frequency  $f_{clk} = 1/T_{clk}$  and is given by  $2^{N_b-1} \times T_{clk}$  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. SS ADCs have also been used as in-pixel ADCs 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 neighboring or column pixels (Bidermann et al., 2003), as shown in Fig. 5.11B. Although most of applications use a linear ramp, other functions can be used as in case of some wide dynamic range (WDR) image sensors (Kim and Song, 2012).

Several other ADC architectures can and were used in image sensor designs such as integrating or dual-slope ADCs (DS ADCs) (Le-Thai et al., 2017), two-step (Lyu et al., 2014) or multiple-slope (Snoeij et al., 2007a), cyclic, and sigma-delta. Some illustrative diagrams are shown in Fig. 5.12. Hybrid designs are also used to address the limitations of each previous architectures by combining two or more of them (Jeon et al., 2017). Table 5.1, adapted from Wang (2016), provides an overall comparison of the most used ADC topologies.

Another alternative for signal conversion is to use time-based ADCs (Kitchen et al., 2004), the majority of which are based on the 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



**Fig. 5.10** (A) A block diagram of successive approximation ADCs (SAR ADCs). (B) A typical circuit implementation.

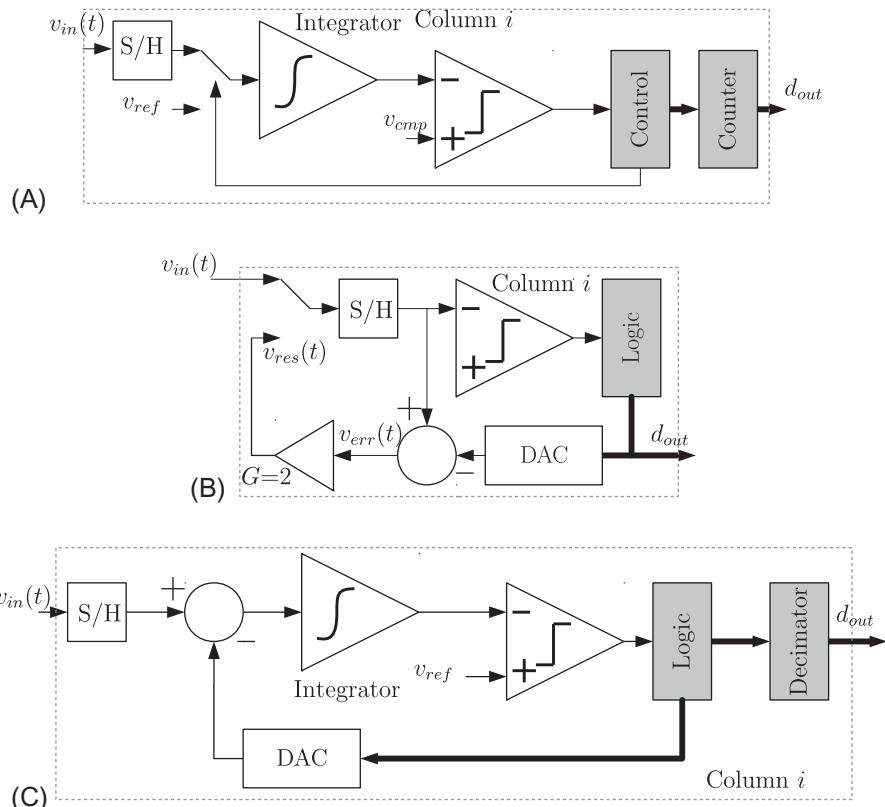


**Fig. 5.11** (A) Block diagram of single-slope ADC (SS ADC). (B) The simplicity of the SS ADC allows for in-pixel or shared-pixel circuit implementations.

(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 DS/multiple-slope ADCs (Snoeij et al., 2007b). Others implement different ADC topologies including sigma-delta converters (Ignjatovic et al., 2012).

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 or static random-access memory (Bidermann et al., 2003). Serial readout of the imaging data has been used in a number of sensors; however, parallel readouts offer higher throughput (Kleinfelder and El Gamal, 2001). Signaling standards developed for high-speed data transmission such as low-voltage differential signaling (Willems et al., 2012) and scalable low-voltage signaling (Toyama et al., 2011) have been used in image sensors. Both of these utilize 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) signaling 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 analog memories to provide temporary frame storage (Tochigi et al., 2012). Although the overall average throughput remains the same, the use of these



**Fig. 5.12** (A) Block diagram of dual-slope ADC (DS ADC). (B) Block diagram of cyclic ADC and (C) first-order sigma-delta ADC (SD ADC).

**Table 5.1** Main ADC architectures used in image sensors.

Topology	Resolution	Speed	Power	Area
SAR	Medium	Medium	Low	High
Single-slope	Medium	Low	Low	Low
Dual-slope	High	Low	Medium	Medium
Cyclic	Medium	Medium	Medium	Medium
Sigma-delta	High	Low	Low	Medium
Hybrid	High	High	Medium	Low

Source: Adapted from Wang, T., 2016. A Study on High-Speed Low-Noise Readout Architectures and Column A/D Converters for CMOS Image Sensors (Ph.D. thesis). Shizuoka University.

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 Tpixels/s ( $10^{12}$  pixels per second) for 128 frames against a 780 Mpixels/s on continuous mode. In [Kondo et al. \(2015\)](#), a stacked architecture was used to implement such memories using micro bumps for every 1.75T pixel cluster achieving 10 kfps (20 Gpixels/s).

### **5.3.3 Pixel improvement**

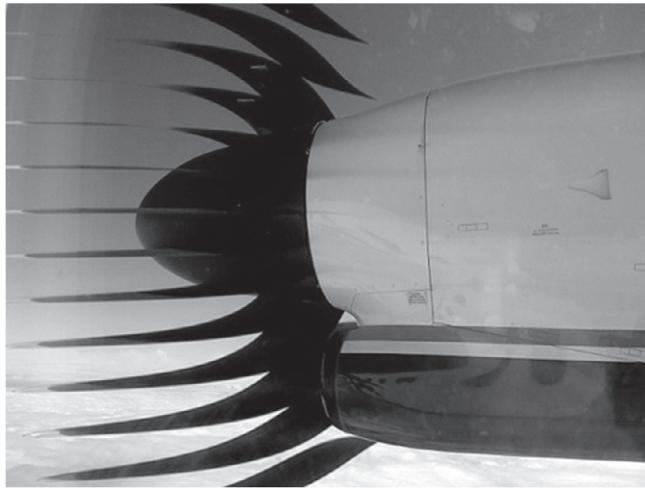
As for the pixel itself, the photodetector area for high-speed applications is usually larger than for high-resolution digital photo-still image sensors. The large pixel size is required to allow a measurable charge integration during the short available integration time ([Takayanagi and Nakamura, 2012](#)).

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 (FD) ([Takeshita et al., 2010](#)). Furthermore, the layout of the photodiode can also result in charge transfer improvements. Examples include T-shaped detectors ([Krymski, 2007](#)) and horn-shaped photodiodes ([Takeshita et al., 2010](#)) 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 ([Spickermann et al., 2011](#)). This reduces the charge transfer time from the photodiode to the FD 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.

### **5.3.4 Global shutters**

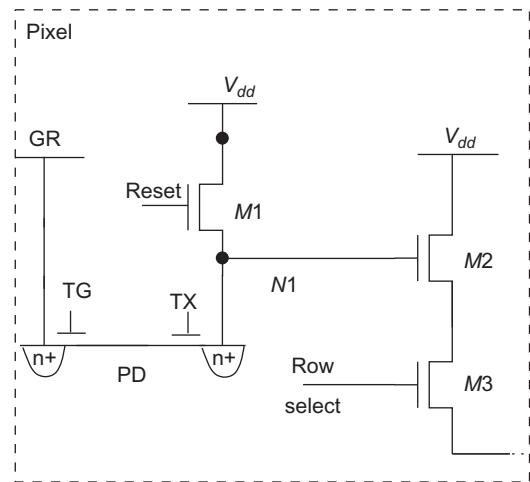
CMOS APS is inherently a rolling shutter pixel system, where every row is read and reset in turn. This, however, introduces smear in fast moving scenes as shown by the example in [Fig. 5.13](#). A number of alternative CMOS pixels have been proposed which are inherently globally shuttered ([Belenky et al., 2007](#)). It is also possible to add additional switches and diffusions to obtain global shutter in an APS.

[Fig. 5.14](#) shows a five-transistor circuit which provides a global shutter by adding an additional diffusion node ([Fox and Nixon, 2000](#)). The circuit provides a global reset through the  $TG$  switch for all photodiodes. At the end of the integration frame, the integrated charge is transferred to the readout node through the switch  $TX$ . However, this disables the CDS and is, therefore, prone to temporal noise. The addition of further switches and diffusions can be made to introduce CDS in this pixel ([Tochigi et al., 2013; Yasutomi et al., 2011](#)).



**Fig. 5.13** A scene showing the effects of rolling shutter on fast moving scenes.

**Fig. 5.14** A five-transistor global-shutter APS circuit.



### 5.3.5 TDI sensors

Linear image sensors' designs present only one row of pixels (1D array) is used to generate 2D images. These imagers are used when the sensor (or the image) is moving with a predictable speed along the perpendicular direction ("along-track") of the pixels row. TDI sensors are similar to linear image sensors but using a number of pixel rows instead of only one (Lepage et al., 2009). This addition is used to implement a multisampling acquisition and therefore, allows for faster scanning (object speed) using similar integration times (frame rates) or for low-light environments. In general, the

SNR of a TDI imager improves as the square root of the number of pixel rows ([Nie et al., 2016](#)).

TDI sensors are preferably implemented using CCD technology rather than CMOS. This is due to the operation of the former allows for natural and noiseless integration of TDI scanning and accumulation principle. In CMOS imagers, the accumulator can be implemented analogically ([Nie et al., 2014](#)) or digitally ([Nie et al., 2016](#)). However, these adders end up noise, leading to significantly stressed readout noise in CMOS implementations of TDI. The dominant noise sources in CMOS TDI systems are generally the pixel source follower, pixel dark current, bias transistor 1/f and random telegraph signal (RTS) as well as bitline sampling kTC noise ([Levski and Choubey, 2016](#)). Analog accumulators induce their own noise in every stage while digital accumulators add ADC noise to the signal chain. Poorly designed CMOS TDI also suffer from poorer MTF and hence optical performance ([Lepage et al., 2009](#); [Levski and Choubey, 2016](#)). An alternative solution to TDI imagers could be to build them in CMOS-CCD process utilizing the advantages of both ([Boulenc et al., 2017](#); [Eckardt et al., 2014](#); [Lee et al., 2017](#)). These offer the potential of low-noise charge accumulation, though at additional cost. Applications of TDI imagers include multispectral imaging, satellite imaging, military reconnaissance, document scanning, machine vision ([Yu et al., 2016](#)) and aerial, X-ray ([Jo et al., 2017](#)), and biomedical ([Zhang et al., 2016](#)) imaging.

## 5.4 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. In these applications, battery life is limited and therefore, image sensors are expected to consume as low power as possible ([Cho et al., 2003](#)). Furthermore, the demands of multimega 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 photodiode. 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 analog and digital circuits on the same chip with the sensors. However, being primarily an analog 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 DR of an image sensor. Technology such as silicon-on-insulator has also been suggested to produce low-power image sensors by reducing leakage current (Brouk et al., 2007; Suntharalingam et al., 2007).

Alternative approaches at architecture or algorithmic level are possible to reduce the power dissipation of image sensors (Fish and Yadid-Pecht, 2008). 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 center. Previous frames can be used to predict any region of interest or even the whole frame. Efficient partition of computation between analog and digital blocks can further optimize 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 (Gao et al., 2005). 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 minimize power consumption of inactive regions of the chip. Multithreshold voltage circuits or dynamic threshold circuits can also be used to further reduce power dissipation in image sensors (Spivak et al., 2012). Pixel-level ADCs can be used to achieve pixel-level parallelism, which in turn can reduce power dissipation (McIlrath, 2001). Another attractive analog design technique to reduce power consumption is to operate some if not all transistors in the subthreshold region (Perenzoni and Gonzo, 2010; McIlrath, 2001). PWM by directly encoding the photocurrent into free-running pulses has also been shown to reduce power dissipation (Wang et al., 2006; Hanson et al., 2010). PWMs offer the advantage of being applicable in even low-voltage technologies as they are generally independent of power supply.

## 5.5 WDR sensors

A typical CMOS APS as well as CCDs have a limited DR of 40–70 dB, which is in turn due to their limited well capacity. However, illumination available in nature and captured by the human eye spans over 180 dB of intensity variation. Even typical real-world scenes can have dynamic intrascene ranges extending to five orders of

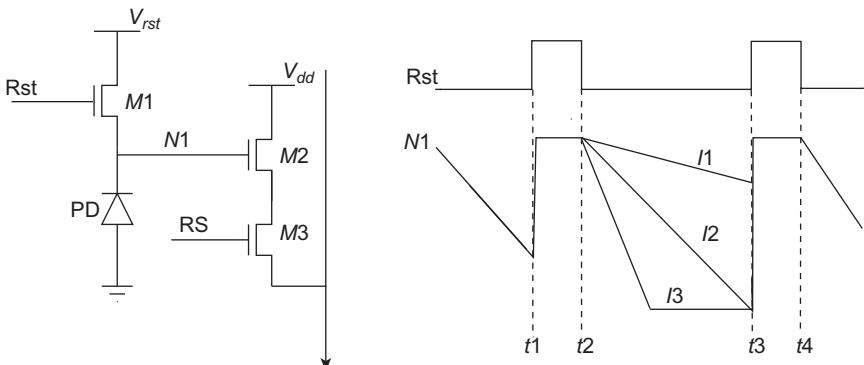


**Fig. 5.15** An image showing the effects of limited DR of APS. Panel (A) shows loss of dark information with a short integration time and panel (B) shows saturation in brighter areas with high integration time.

magnitude, from 1 lux in shadow to 1–105 lux of bright sunlight. The limited DR 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. 5.15. A number of circuits have, therefore, been proposed and used for extending the DR of CMOS APSs as well as for new modes of imaging leading to capture of a WDR of intensities (Yadid-Pecht, 1999; Choubey, 2011a).

Before discussing these approaches, let us revisit an APS to understand its limited DR. Fig. 5.16 shows an APS and a typical signal diagram for three different photo-currents. For very high input intensity,  $I_3$ , and for all currents higher than  $I_2$ , the pixels saturate to a constant output. This limiting level of photo-generated charge 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 DR of the pixel. In addition to these two values, the DR of a pixel is also affected by the residual temporal noise. It is also worth noting, however, that the lowest and the highest level



**Fig. 5.16** CMOS APS and its operation.

of intensities faithfully recorded can be changed by altering the integration time of the pixel. However, this does not change the DR as any change in integration time changes both levels by the same factor.

### 5.5.1 Logarithmic sensors

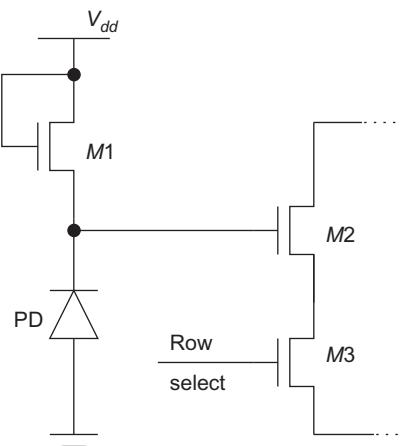
A simple technique to capture the WDR of intensities is to use a logarithmic amplifier inside each pixel to compress the input photocurrent. Fig. 5.17 shows a simple implementation of this technique using an n-type metal-oxide-semiconductor (nMOS) transistor in weak inversion to build the required logarithmic 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:

$$V_{S,M1} (\text{or } V_{G,M2}) = V_{G,M1} - V_{T,M1} - \frac{n k T}{q} \log \left( \frac{I_{DS,M1}}{I_{DSO,M1}} \right) \quad (5.12)$$

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.

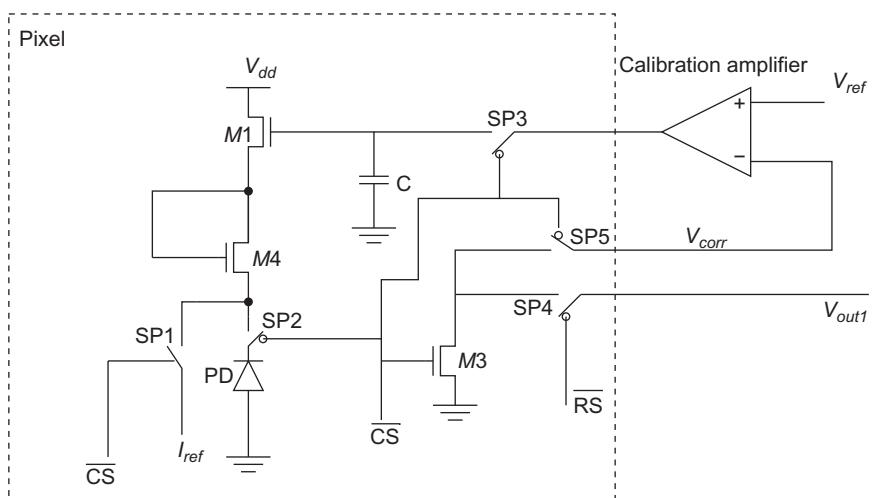
This circuit has the advantage of simultaneous capture and compression of intensities, thereby reducing complexity in later chains of readout including the number of bits required for data conversion. Such pixels also provide true random addressability on account of their continuous operation. However,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



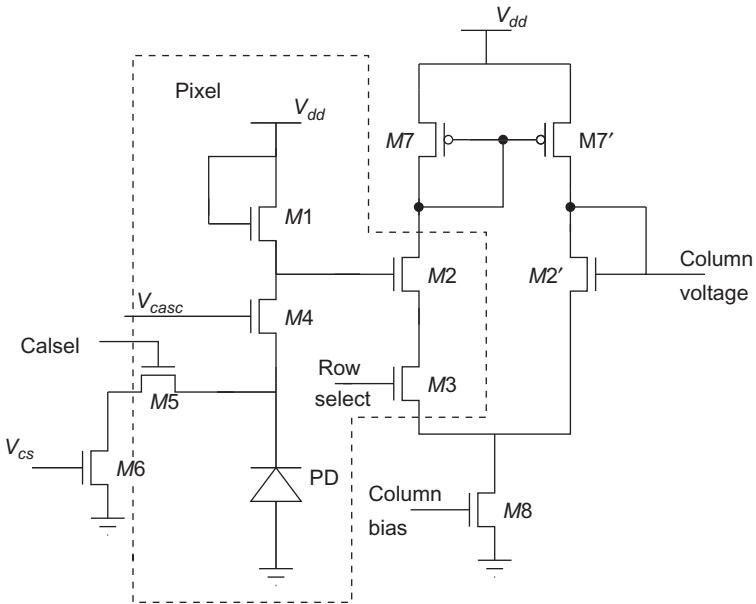
**Fig. 5.17** A typical CMOS logarithmic pixel using an nMOS in weak inversion.

Several approaches have been suggested to reduce the FPN by utilizing reference values for additional FPN correction though with limited success (Kavadia et al., 2000; Labonne et al., 2006; Ni and Matao, 2001; Lai et al., 2004). 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. Postfabrication correction for nonuniformity using hot carrier degradation (Ricquer and Dierickx, 1995) or compensatory circuits (Loose et al., 2001) has 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. Fig. 5.18 shows an example of a compensatory circuit, where the gate voltage of one of the load transistors is adjusted during the calibration process by using external reference voltages, thereby compensating for variations (Loose et al., 2001).

Most of these techniques only reduce the additive FPN; however, the logarithmic pixel has small gain on account of small subthreshold slope of the load device. This means that even a small multiplicative FPN will reduce the image quality (Joseph and Collins, 2001; Choubey et al., 2006). Therefore, logarithmic pixels often require multiparametric calibration to significantly reduce the FPN. Fig. 5.19 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.



**Fig. 5.18** In-pixel calibration circuit proposed by Loose et al. (2001).

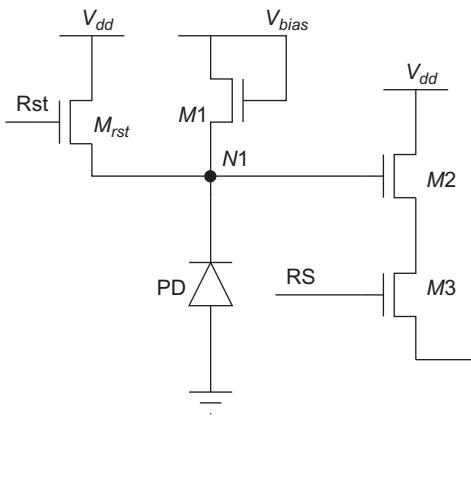


**Fig. 5.19** A logarithmic pixel with one stage of a differential amplifier readout and circuit for electronic calibration (Choubey et al., 2006).

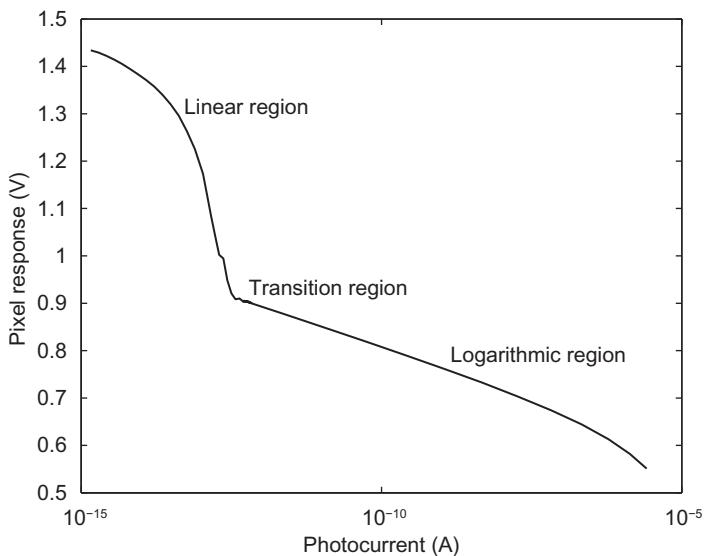
### 5.5.2 Pixels with combined linear and logarithmic response

Logarithmic pixels suffer from limited gain, particularly in low-light regions. Furthermore, transition from high light to low light is problematic on account of the long settling time of these pixels. Circuits have, therefore, been proposed which aim to utilize the fast settling time of linear pixels and the WDR response of logarithmic pixels. A simple approach is to use an APS 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 (Strom et al., 2006). 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 (Fox et al., 2000; Choubey and Collins, 2008) as shown in Fig. 5.20.

The pixel operates as a linear pixel at the start of the frame. It is reset using the  $M_{rst}$  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 APS mode and a linear response is achieved, as shown in Fig. 5.21. However, if



**Fig. 5.20** A pixel with combined linear and logarithmic response.



**Fig. 5.21** Response curve of a combined response pixel showing linear response for low light and logarithmic for high light.

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 (Choubey and Collins, 2007). Operating in two regions, however, introduces higher FPN, particularly in the transition between linear and logarithmic region. A mixture of double sampling as well as calibration can be used to improve the FPN of these sensors.

### 5.5.3 Threshold comparing pixels

It is worth noting that CMOS APSSs lose their ability to record WDR 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 DR 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 (Stoppa et al., 2002; Ikebe and Saito, 2007). 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 Lai et al. (2006). 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 (Andoh et al., 2000). 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 (Stoppa et al., 2002). 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 is 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 (Luo et al., 2006). 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 (Culurciello et al., 2003).

### 5.5.4 Integration time control sensors

Controlling the integration time of pixels can provide another approach to increase their DR. 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 DR of the images captured but allows for a different region of intensities to be captured depending upon the scene (Yadid-Pecht et al., 1991; Kinugusa et al., 1987). An extension to this principle is to locally adapt the integration time for each pixel or group of pixels (Schanz et al., 2000; Yasuda et al., 2003). 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 (Yasuda et al., 2003). 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 WDR signal. Circuits used to achieve these goals are often similar or identical to an APSSs, however,

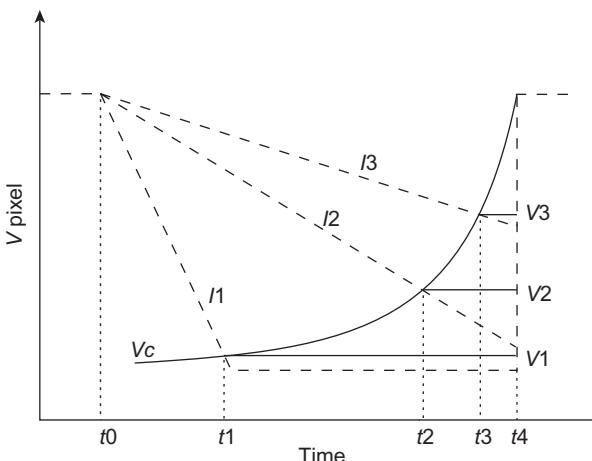
complex external circuitry often at the end of columns, is used to control the reset signal (Schanz et al., 2000). Furthermore, use of local memory and need of comparators often leads to higher noise as well as difficulty in the reconstruction of signals.

### 5.5.5 Threshold comparing as well as integration time control pixels

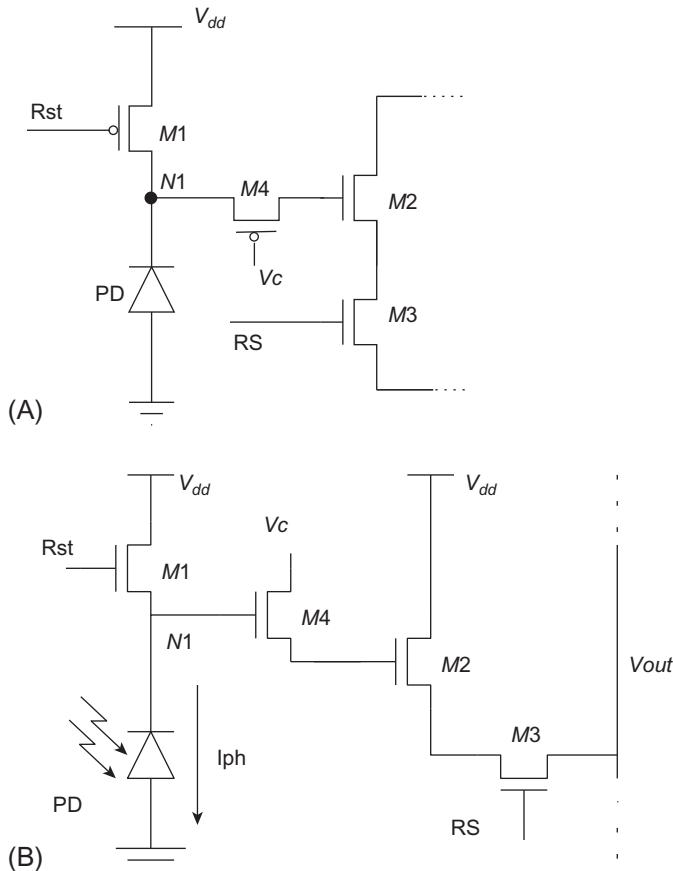
Combining the two approaches of threshold comparison and integration time control provides another approach to obtain WDR information (Cheng et al., 2007; Choubey et al., 2008). Fig. 5.22 shows the signal diagram to reflect the approach. In a linear pixel operation as shown by the dashed lines, high photocurrent,  $I_1$ , leads to saturation. In this approach, however, the integrated signal is compared to a constantly increasing reference signal and the integration is stopped when the two are equal. This means for the high current,  $I_1$ , the integration will stop at lower voltage  $V_1$  than the voltages for lower currents  $I_2$  and  $I_3$ . The pixel output after the integration frame will thus depend on the  $V_c$  signal.

Fig. 5.23A 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 (Choubey et al., 2008). Alternatively, and preferably, an nMOS only pixel is also shown which records the reference value rather than the integrated value (Choubey, 2011b). 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  $V_c$  is lower than  $V_{N1}$  and hence transistor  $M4$  is on. However,



**Fig. 5.22** Signal flow in an integrating pixel with threshold comparison and integration time control.



**Fig. 5.23** Four-transistor integrating pixel capable of producing WDR output of any monotonically increasing transduction function. (A) Uses a pMOS device as switch and (B) uses an nMOS device as switch.

discharging  $V_{N1}$  and monotonically increasing  $V_c$  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  $M2$ .

With a suitable control signal  $V_c$ , 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 (Choubey, 2010). 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 integration time. This further means that the pixel provides inherent global shuttering.

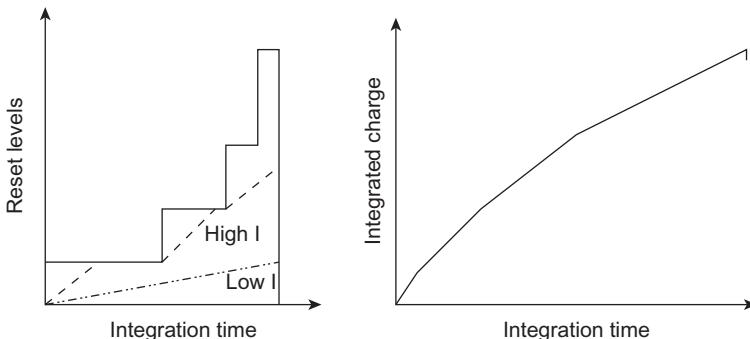
### 5.5.6 Well capacity adjustment

One can also increase the well capacity of the pixel to enhance the amount of charge it can hold and hence its DR. 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 DR of pixels in CCDs and later in CMOS (Decker et al., 1998). The reset device itself can act to modulate the charge storage in the diode (Decker et al., 1998). For example, Fig. 5.24 shows a nonlinear stepping of the reset signal in a typical APS which can be used modulate the charge stored in the pixel and increase the DR.

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 (Guidash, 2001). 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 FD. One can also add additional capacitors to further enhance the ability to integrate spilled charge as well as reduce noise (Akahane et al., 2006). The light to pixel output relationship in such pixels is linear; however, with a sudden change in slope as shown in Fig. 5.25.

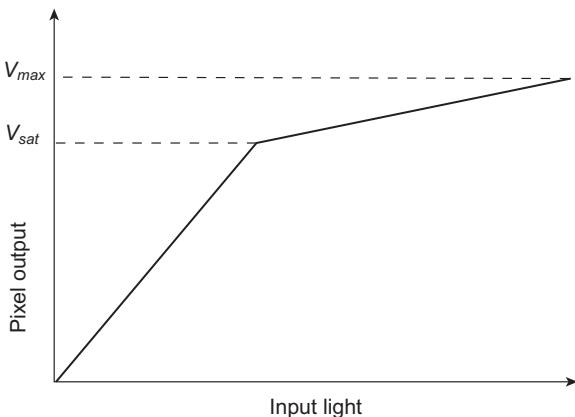
### 5.5.7 Frequency-based WDR image sensors

Mapping a WDR of intensities on to a voltage domain for pixel output increases the requirements from the data conversion stage as 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



**Fig. 5.24** A stepped reset signal and corresponding collected charge used to enhance the DR of the sensor (Decker et al., 1998).

**Fig. 5.25** Pixel response when independently integrating the spilled over charge (Guidash, 2001).



input intensity on to the frequency domain where a large DR can be easily recorded and stored. A number of pixels with such potential have been suggested (Seitz and Raynor, 1997; Wang et al., 2006). However, these approaches often require large circuitry which is incompatible with the requirement of small pixels.

### 5.5.8 Multiple sampling image sensors

Another approach which has been extensively used to increase the DR 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. 5.15. An algorithmic fusion of these two images can then be used to produce a WDR image (Yadid-Pecht and Fossum, 1997; Schrey et al., 2001).

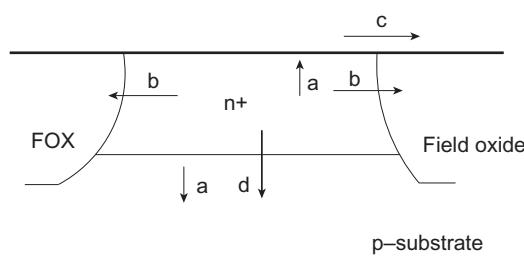
An extension of this approach leading to improved image quality is to capture the same scene with several different integration times, for example, at exponentially increasing times (Yang et al., 2000). In-pixel memories can be used to enhance the performance of multiple sampling techniques. Predictive algorithms have been used to automatically find an optimal integration time (the longest integration time for which the pixel does not saturate) (Acosta-Serafini et al., 2004). This can be achieved by dividing the integration time into slots of different duration 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. In addition, in-pixel memory can also be used to store the reset information (Belenky et al., 2007). Dual and multisampling techniques, however, are unable to capture the full DR in a single frame and often require a large number of frames along with computationally complex postprocessing, thereby making them difficult to use in video and other fast capture applications.

## 5.6 Other high-performance designs

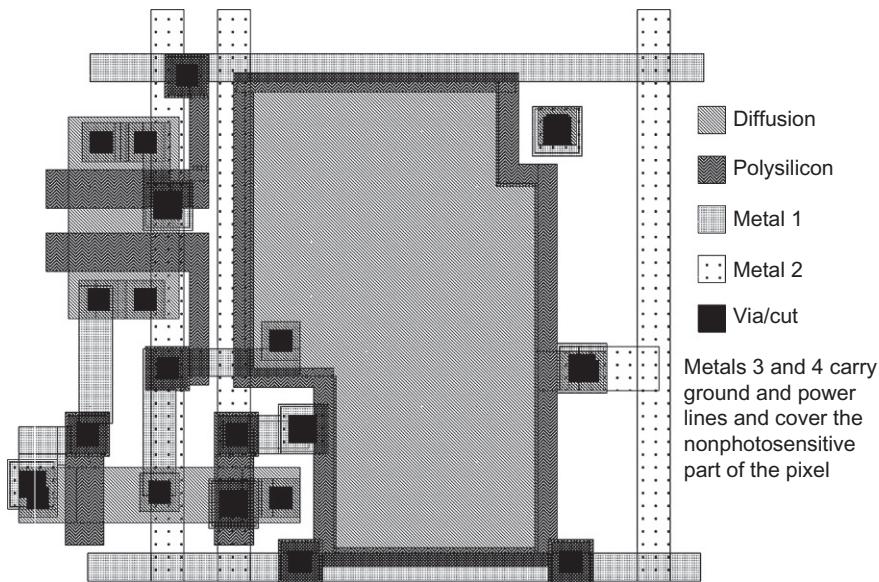
CMOS image sensors as low-cost low-power highly integrable detectors have found applications in a number of diverse fields. This has also prompted development of application-specific image sensors. These include very large format image sensors for radiography, very low-light detectors for astrophysics, 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 (Turchetta et al., 2011). This makes the design of such detectors fairly challenging. Nevertheless, such detectors do not need very small pixel pitch and often pixels with 100  $\mu\text{m}$  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. FPN as well as temporal noise correction need further attention with this multichip 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 (Yamashita et al., 2011).

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 (Choubey and Collins, 2008; Kwon et al., 2004b). This significantly affects the low-light performance of the pixels. Fig. 5.26 shows various sources of leakage in a typical photodiode (Choubey and Collins, 2008). 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 has been proposed to reduce the inherent leakage



**Fig. 5.26** Sources of dark current in a standard n+/p-sub diode in CMOS technology. Thermal drift current (a), lateral (b), and (c) surface leakage currents and diffusion current (d).



**Fig. 5.27** Layout of a logarithmic pixel with a guard ring around the photodiode (Choubey and Collins, 2008).

current in the diode. These include enhanced surface cleaning (Kwon et al., 2004a) and separation of the diode from stressed areas by using pinned or buried devices (Kwon et al., 2004b; Inoue et al., 2003). Furthermore, special barrier layers have also been suggested to reduce the effect on diode leakage (Okita and Suzuki, 2002; Mann, 2003). 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 (Choubey and Collins, 2008; Kopley et al., 2002; Cheng and King, 2002) with a guard ring of polysilicon as shown in Fig. 5.27.

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) (Niclass et al., 2005), 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 subelectron noise floor over hours of measurement.

## 5.7 Conclusion

In this chapter, we have studied approaches to enhance the performance of CMOS image sensors including higher speed, lower noise, lower power, wider DR, and lower dark performance. A number of these have required technology improvement as well as codesign of circuits and technologies. Higher speed has been achieved primarily by

advancing the signal conversion stage. Circuits of CDS as well as the ability to manufacture special diodes has enabled reduction of temporal as well as FPN. A number of circuit approaches have been presented to improve the DR of the image sensor. Low-light performance has primarily been improved by process modifications; however, special circuits such as SPADs 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 nonstandard detectors to enhance spectral sensitivity of image sensors as well as nanophotonic structures to image into the terahertz regime as well.

## References

- Acosta-Serafini, P.M., Masaki, I., Sodini, C.G., 2004. A 1/3"VGA linear wide dynamic range CMOS image sensor implementing a predictive multiple sampling algorithm with overlapping integration intervals. *IEEE J. Solid State Circuits* 39 (9), 1487–1496.
- Akahane, N., Sugawa, S., Adachi, S., Mori, K., Ishiuchi, T., Mizobuchi, K., 2006. A sensitivity and linearity improvement of a 100-db dynamic range CMOS image sensor using a lateral overflow integration capacitor. *IEEE J. Solid State Circuits* 41 (4), 851–858.
- Andoh, F., Shimamoto, H., Fujita, Y., 2000. A digital pixel image sensor for real time readout. *IEEE Trans. Electron Devices* 47 (11), 2123–2127.
- Arai, T., Yasue, T., Kitamura, K., Shimamoto, H., Kosugi, T., Jun, S., Aoyama, S., Hsu, M.-C., Yamashita, Y., Sumi, H., Kawahito, S., 2016. 6.9 A 1.1  $\mu$ m 33 Mpixel 240 fps 3D-stacked CMOS image sensor with 3-stage cyclic-based analog-to-digital converters. In: 2016 IEEE International Solid-State Circuits Conference (ISSCC), vol. 59. IEEE, pp. 126–128.
- Belenky, A., Fish, A., Spivak, A., Yadid-Pecht, O., 2007. Global shutter CMOS image sensor with wide dynamic range. *IEEE Trans. Circuits Syst. II Express Briefs* 54 (12), 1032–1036.
- Bidermann, W., El Gamal, A., Ewedemi, S., Reyneri, J., Tian, H., Wile, D., Yang, D., 2003. A 0.18  $\mu$ m high dynamic range NTSC/PAL imaging system-on-chip with embedded DRAM frame buffer. In: IEEE Int. Solid-State Circuits Conference, 2003. Digest of Technical Papers. ISSCC, vol. 1. IEEE, pp. 212–488.
- Boulenc, P., Robbelein, J., Wu, L., Haspeslagh, L., De Moor, P., Borremans, J., Rosmeulen, M., 2017. High speed TDI embedded CCD in CMOS sensor. In: International Conference on Space Optics—ICSO 2016, Proc. SPIE, vol. 10562.
- Brouk, I., Alameh, K., Nemirovsky, Y., 2007. Design and characterization of CMOS/SOI image sensors. *IEEE Trans. Electron Devices* 54 (3), 468–475.
- Chao, C., Tu, H.Y., Chou, K.Y., Chou, P.S., Hsueh, F.L., 2011. Crosstalk metrics and the characterization of 1.1  $\mu$ m-pixel CIS. Proc. 2011 Int. Image Sensor Workshop, IISS, Hokkaido, Japan, pp. 3–6.
- Chen, T., Catrysse, P., Gamal, A.E., Wandell, B., 2000. How small should pixel size be? Proceedings of SPIE, 3965. SPIE, pp. 451–459. vol.
- Cheng, H.Y., King, Y.C., 2002. An ultra low dark current CMOS image sensor cell using n+ ring reset. *IEEE Electron Device Lett.* 23 (9), 538–540.
- Cheng, H.Y., Choubey, B., Collins, S., 2007. A high dynamic range integrating pixel with an adaptive logarithmic responses. *IEEE Photon. Technol. Lett.* 19 (15), 1169–1171.
- Cheng, Y.W., Tsai, T.H., Chou, C.H., Lee, K.C., Chen, H.C., Hsu, Y.L., 2015. Optical performance study of BSI image sensor with stacked grid structure. In: Technical Digest—International Electron Devices Meeting, IEDM, 2016, February, pp. 30.4.1–30.4.4.

- Cho, K.B., Krymski, A.I., Fossum, E.R., 2003. A 1.5-V 550- $\mu$ W 176  $\times$  144 autonomous CMOS active pixel image sensor. *IEEE Trans. Electron Devices* 50 (1), 96–105.
- Choubey, B., 2010. 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.
- Choubey, B., 2011a. CMOS pixels for wide dynamic range imaging. In: Advances in Electrical Engineering Research. Nova Science Publisher, Hauppauge, New York, pp. 329–375.
- Choubey, B., 2011b. Pixel circuit. GB Patent 20110014898 20110826.
- Choubey, B., 2012. Fixed pattern noise. Encyclopedia of Optical Engineering, Taylor & Francis, Boca Raton.
- Choubey, B., Collins, S., 2007. Models for pixels with wide dynamic range combined linear and logarithmic response. *IEEE Sens. J.* 7 (7), 1066–1072.
- Choubey, B., Collins, S., 2008. Wide dynamic range CMOS pixels with reduced dark current. *Analog Integr. Circ. Sig. Process* 56, 53–60.
- Choubey, B., Otim, S., Aoyama, S., Joseph, D., Collins, S., 2006. An electronic calibration scheme for logarithmic CMOS image sensor. *IEEE J. Sens.* 6 (4), 950–956.
- Choubey, B., Cheng, H.Y., Collins, S., 2008. A wide dynamic range CMOS image sensor with adjustable logarithmic response. In: Proceedings of the SPIE Electronic Imaging, vol. 6816.
- Culurciello, E., Etienne-Cummings, R., Boahen, K.A., 2003. A biomorphic digital image sensor. *IEEE J. Solid State Circuits* 38 (2), 281–294.
- Decker, S., McGrath, D., Brehmer, K., Sodini, C.G., 1998. A 256  $\times$  256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output. *IEEE J. Solid State Circuits* 33 (12), 2081–2091.
- Driggers, R.G., Vollmerhausen, R., Krapels, K., 2001. Target identification performance as a function of temporal and fixed pattern noise. *Opt. Eng.* 40 (3), 443–447.
- Eckardt, A., Reulke, R., Jung, M., Sengebusch, K., 2014. CMOS-TDI detector technology for reconnaissance application. Proceedings of SPIE, vol. 9249, p. 92490X.
- Fish, A., Yadid-Pecht, O., 2008. Low-power smart CMOS image sensors. In: IEEE International Symposium on Circuits and Systems, 2008. ISCAS 2008, pp. 1408–1411.
- Fox, E.C., Nixon, O., 2000. Pinned photodiode five transistor pixel. US Patent no. 6566697.
- Fox, E.C., Hynecek, J., Dykaar, D.R., 2000. Wide dynamic range pixel with combined linear and logarithmic response and increased signal swing. In: Proceedings of SPIE. IS&T/SPIE 12th International Symposium on Electronic Imaging—Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications, 3965. SPIE, pp. 4–10. vol. 3965.
- Gamal, A.E., Fowler, B.A., Min, H., Liu, X., 1998. Modeling and estimation of FPN components in CMOS image sensors. Proceedings of the SPIE, vol. 3301, pp. 168–177.
- Gao, W., Shen, E., Hornsey, R.I., 2005. Low-power realization in main blocks of CMOS APS image sensor. Proceedings of the SPIE, vol. 5969, pp. 507–515.
- Guidash, R., 2001. Variable collection of blooming charge to extend dynamic range. US Patent no. 6307195.
- Hanson, S., Foo, Z., Blaauw, D., Sylvester, D., 2010. A 0.5 V sub-microwatt CMOS image sensor with pulse-width modulation read-out. *IEEE J. Solid State Circuits* 45 (4), 759–767.
- Hardie, R.C., Hayat, M.M., Armstrong, E., Yasuda, B., 2000. Scene-based nonuniformity correction with video sequences and registration. *Appl. Opt.* 39 (8), 1241–1250.
- Holst, G.C., 2007. CMOS/CCD Sensors and Camera Systems. SPIE Press, Bellingham, WA.
- Huang, S., Yamashita, T., Wang, Y., Ong, K.L., 2009. A 2.5 inch, 33 Mpixel, 60 fps CMOS image sensor for UHDTV application. Proc. 2009 Int. Image Sensor Workshop, IISS, Bergen, Norway, pp. 308–311.

- Ignjatovic, Z., Maricic, D., Bocko, M.F., 2012. Low power, high dynamic range CMOS image sensor employing pixel-level oversampling sigma-delta analog-to-digital conversion. *IEEE Sens. J.* 12 (4), 737–746.
- Ikebe, M., Saito, K., 2007. A wide-dynamic-range compression image sensor with negative-feedback resetting. *IEEE Sens. J.* 7 (5), 897–904.
- Inoue, I., Tanaka, N., Yamashita, H., Yagamuchi, T., Ishiwata, 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.
- Jeon, B.-K., Hong, S.-K., Kwon, O.-K., 2017. A low-power 12-bit extended counting ADC without calibration for CMOS image sensors. *IEEE Trans. Circuits Syst. II Express Briefs* 7747 (c), 1.
- Jo, Y.-R., Hong, S.-K., Kwon, O.-K., 2017. A tileable CMOS X-ray line detector using time-delay-integration with pseudomultisampling for large-sized dental X-ray imaging systems. *IEEE Trans. Electron Devices* 64 (1), 211–216.
- Joseph, D., Collins, S., 2001. Modelling, calibration and correction of illumination-dependent fixed pattern noise in logarithmic CMOS image sensor. *IEEE Trans. Instrum. Meas.* 55 (5), 996–1001.
- Joseph, D., Collins, S., 2009. Temperature dependence of fixed pattern noise in logarithmic CMOS image sensors. *IEEE Trans. Instrum. Meas.* 58 (8), 2503–2511.
- Kavadias, S., Dierickx, B., Scheffer, D., Alaerts, A., Uwaerts, D., Bogaerts, J., 2000. A logarithmic response CMOS image sensor with on-chip calibration. *IEEE J. Solid State Circuits* 35 (8), 1146–1152.
- Kim, D., Song, M., 2012. An enhanced dynamic-range CMOS image sensor using a digital logarithmic single-slope ADC. *IEEE Trans. Circuits Syst. II Express Briefs* 59 (10), 653–657.
- Kim, Y.C., Kim, Y.T., Choi, S.H., Kong, H.K., Hwang, S.I., Ko, J.H., Kim, B.S., Asaba, T., Hee, S.H.L., Hahn, J.S., Im, J.H., Oh, T.S., Yi, D.M., Lee, J.M., Yang, W.P., Ahn, J.C., Jung, E.S., Lee, Y.H., 2006. 1/2-inch 7.2 Mpixel CMOS image sensor with 2.25  $\mu\text{m}$  pixels using 4-shared pixel structure for pixel-level summation. In: 2006 IEEE Int. Solid State Circuits Conference—Dig. Tech. Papers, IEEE, pp. 1994–2003.
- Kinugusa, T., Noda, M., Imaide, T., Aizawa, I., Todaka, Y., Ozama, M., 1987. An electronic variable-shutter system in video camera use. *IEEE Trans. Consum. Electron.* CE-33, 249–258.
- Kitamura, K., Watabe, T., Sadanaga, Y., Sawamoto, T., Kosugi, T., Akahori, T., Iida, T., Isobe, K., Watanabe, T., Shimamoto, H., Ohtake, H., Aoyama, S., Kawahito, S., Egami, N., 2011. A 33 Mpixel, 120 fps CMOS image sensor for UDTV application with two-stage column-parallel cyclic ADCs. *Proc. Int. Image Sensor Workshop*, IISS, Hokkaido, Japan, pp. 3–6.
- Kitchen, A., Bouzerdoum, A., Bermak, A., 2004. Time domain analogue to digital conversion in a digital pixel sensor array. In: Second IEEE International Workshop on Electronic Design, Test and Applications, IEEE, pp. 108–112.
- Kiyoyama, K., Lee, K.W., Fukushima, T., Naganuma, H., Kobayashi, H., Tanaka, T., Koyanagi, M., 2011. A very low area ADC for 3-D stacked CMOS image processing system. In: 2011 IEEE International 3D Systems Integration Conference, 3DIC 2011, pp. 3–6.
- Kiyoyama, K., Sato, Y., Hashimoto, H., Lee, K.-W., Fukushima, T., Tanaka, T., Koyanagi, M., 2013. A block-parallel ADC with digital noise cancelling for 3-D stacked CMOS image sensor. In: 2013 IEEE International 3D Systems Integration Conference (3DIC), IEEE, pp. 1–4.
- Kleinfelder, S., El Gamal, A., 2001. A 10000 frames/s CMOS digital pixel sensor. *IEEE J. Solid State Circuits* 36 (12), 2049–2059.
- Kondo, T., Takemoto, Y., Kobayashi, K., Tsukimura, M., Takazawa, N., Kato, H., Suzuki, S., Aoki, J., Saito, H., Gomi, Y., Matsuda, S., Tadaki, Y., 2015. A 3D stacked CMOS image

- sensor with 16 Mpixel global-shutter mode and 2 Mpixel 10000 fps mode using 4 million interconnections. *2015 Symposium on VLSI Circuits (VLSI Circuits)*, IEEE, Kyoto, Japan, pp. C90–C91.
- Kopley, T.E., Vook, D.W., Dungan, T., 2002. Method of manufacturing a structure for reducing leakage currents by providing isolation between adjacent regions of an integrated circuits. US Patent no. 6417074.
- Krymski, A., 2007. A high speed 4 megapixel digital CMOS sensor. *Proc. 2007 Int. Image Sensor Workshop*, IISS, Ogunquit, ME, pp. 78–81.
- Krymski, A.I., Bock, N.E., Van Blerveld, D., Fossum, E.R., 2003. A high-speed, 240-frames/s, 4.1-Mpixel CMOS sensor. *IEEE Trans. Electron Devices* 50 (1), 130–135.
- Kwon, H.I., Kang, I.M., Park, B.G., Lee, J.D., Park, S.S., 2004. The analysis of dark signals in the CMOS APS imagers from the characterization of test structures. *IEEE Trans. Electron Devices* 51 (2), 178–184.
- Kwon, H.I., Kwon, O.J., Shin, H., Park, B.G., Park, S.S., Lee, J.D., 2004. The effects of deuterium annealing on the reduction of dark currents in the CMOS APS. *IEEE Trans. Electron Devices* 51 (8), 1346–1349.
- Labonne, E., Sicard, G., Renaudin, M., Berger, P.D., 2006. A 100 db dynamic range CMOS image sensor with global shutter. In: *13th IEEE International Conference on Electronics, Circuits and Systems*, 2006. ICECS '06, pp. 1133–1136.
- Lai, L.W., Lai, C.H., King, Y.C., 2004. A novel logarithmic response CMOS image sensor with high output voltage swing and in-pixel fixed-pattern noise reduction. *IEEE Sens. J.* 4 (1), 122–126.
- Lai, C.H., King, Y.C., Huang, S.Y., 2006. A 1.2-V 0.25  $\mu\text{m}$  clock output pixel architecture with wide dynamic range and self-offset cancellation. *IEEE Sens. J.* 6 (2), 398–405.
- Le-Thai, H., Chapinal, G., Geurts, T., Gielen, G.G.E., 2017. A 0.18- $\mu\text{m}$  CMOS image sensor with phase-delay-counting and oversampling dual-slope integrating column ADCs achieving 1E-RMS noise at 3.8- $\mu\text{s}$  conversion time. *IEEE J. Solid State Circuits* 1, 1–12.
- Lee, H.J., Donegan, P., Atos, D., Feng, F.H., Fox, E., Ghannoum, R., Guan, J., Maes, W., Marchesan, D., Moser, M., Nixon, O., Ruiter, M., Tsai, T.H., Korthout, L., 2017. Charge-coupled CMOS TDI imager. *Proc. Int. Image Sensor Workshop*.
- Leñero-Bardallo, J.A., Fernández-Berni, J., Rodríguez-Vázquez, Á., 2014. Review of ADCs for imaging. *IS&T/SPIE Electronic Imaging*, p. 90220I.
- Lepage, G., Bogaerts, J., Meynarts, G., 2009. Time-delay-integration architectures in CMOS image sensors. *IEEE Trans. Electron Devices* 56 (11), 2524–2533.
- Levski, D., Choubey, B., 2016. On noise in time-delay integration CMOS image sensors. *Proceedings of the SPIE*, vol. 9891, p. 11.
- Lim, S., Gamal, A.E., 2004. Gain fixed pattern noise correction via optical flow. *IEEE Trans. Circuits Syst. I* 51 (4), 779–786.
- Lin, D.L., Wang, C.C., Wei, C.L., 2010. Quantified temperature effect in a CMOS image sensor. *IEEE Trans. Electron Devices* 57 (2), 422–428.
- Loose, M., Meier, K., Schemmel, J., 2001. A self-calibrating single-chip CMOS camera with logarithmic response. *IEEE J. Solid State Circuits* 36 (4), 586–596.
- Luo, Q., Harris, J.G., Chen, Z.J., 2006. A time-to-first spike CMOS image sensor with coarse temporal sampling. *Analog Integr. Circ. Sig. Process* 47, 303–313.
- Lyon, R.F., Hubel, P.M., 2002. Eyeing the camera: into the next century. In: *The Tenth Color Imaging Conference: Color Science and Engineering Systems, Technologies, Application, IST—The Society for Imaging Science and Technology*, Scottsdale, AZ, pp. 349–355.
- Lyu, T., Yao, S., Nie, K., Xu, J., 2014. A 12-bit high-speed column-parallel two-step single-slope analog-to-digital converter (ADC) for CMOS image sensors. *Sensors (Switzerland)* 14 (11), 21603–21625.

- Mann, R., 2003. Improved semiconductor device for isolating a photodiode to reduce junction leakage and method of formation. International Patent no. WO 03/019675 A1.
- Matsu, S., Bales, T.J., Shoda, M., Osawa, S., Kawamura, K., Andersson, A., Haque, M., Honda, H., Almond, B., Mo, Y., Gleason, J., Chow, T., Takayanagi, I., 2009. 8.9-Megapixel video image sensor with 14-b column-parallel SA-ADC. *IEEE Trans. Electron Devices* 56 (11), 2380–2389.
- McIlrath, L.G., 2001. A low-power low-noise ultrawide-dynamic-range CMOS imager with pixel-parallel A/D conversion. *IEEE J. Solid State Circuits* 36 (5), 846–853.
- Mendis, S.K., Kemeny, S.E., Gee, R.C., Pain, B., Staller, C.O., Kim, Q., Fossum, E.R., 1997. CMOS active pixel image sensors for highly integrated imaging systems. *IEEE J. Solid State Circuits* 32 (2), 187–197.
- Nguyen, A.Q., Shimonomura, K., Dao, V.T.S., Kamakura, Y., Etoh, T.G., 2016. Crosstalk analysis of an image sensor operating at 1 Gfps. *2016 IEEE 6th International Conference on Communications and Electronics, IEEE ICCE 2016*, 176–180.
- Ni, Y., Matao, K., 2001. A CMOS log sensor with on-chip FPN compensation. In: *Proceedings of the European Solid-State Circuits Conference (ESSCIRC)*pp. 128–132.
- Niclass, C., Rochas, A., Besse, P.-A., 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 (9), 1847–1854.
- Nie, K., Yao, S., Xu, J., Gao, J., Xia, Y., 2014. A 128-stage analog accumulator for CMOS TDI image sensor. *IEEE Trans. Circuits Syst. I Reg. Papers* 61 (7), 1952–1961.
- Nie, K., Xu, J., Gao, Z., 2016. A 128-stage CMOS TDI image sensor with on-chip digital accumulator. *IEEE Sens. J.* 16 (5), 1319–1324.
- Okita, A., Suzuki, S., 2002. Photoelectric transducer and imaging device. Japanese Patent publication no. 2002-353430.
- Pelgrom, M.J.M., Duinmaijer, A.C.J., Welbers, A.P.G., 1989. Matching properties of MOS transistors. *IEEE J. Solid State Circuits* 24 (5), 1433–1439.
- Perenzi, M., Gonzo, L., 2010. Solar-powered CMOS image sensor. *Electron. Lett.* 46 (1), 77–79.
- Ricquer, N., Dierickx, B., 1995. Active pixel CMOS sensor with on chip nonuniformity correction. In: *Proceedings of the IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*.
- Roy, F., Mamdy, B., Ahmed, N., Tournier, A., Lu, G.N., 2017. Development of small-sized pixel structures for high-resolution CMOS image sensors. In: *2017 2nd International Conference on Image, Vision and Computing, ICIVC 2017*, pp. 494–500.
- Schanz, M., Nitta, C., Bubmann, A., Hostika, B.J., Wertheimer, R.K., 2000. A high dynamic range CMOS image sensor for automotive applications. *IEEE J. Solid State Circuits* 35 (7), 932–938.
- Schrey, O., Huppertz, J., Filimonovic, G., Bubmann, A., Brovkherde, W., Hosticka, B.J., 2001. 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)*.
- Seitz, P., Raynor, J.M., 1997. A linear array of photodetectors with wide dynamic range and near quantum-noise limit. *Sens. Actuators A* 61 (1–3), 327–330.
- Snoeij, M.F., Theuwissen, A.J.P., Huijsing, J.H., Makinwa, K.A.A., 2007. Power and area efficient column-parallel ADC architectures for CMOS image sensors. *2007 IEEE Sensors*, 523–526.
- Snoeij, M.F., Theuwissen, A.J.P., Makinwa, K.A.A., Huijsing, J.H., 2007. Multiple-ramp column-parallel ADC architectures for CMOS image sensors. *IEEE J. Solid State Circuits* 42 (12), 2968–2977.

- Spickermann, A., Durini, D., Suss, A., Ulfig, W., Brockherde, W., Hosticka, B.J., Schwope, S., Grabmaier, A., 2011. CMOS 3D image sensor based on pulse modulated time-of-flight principle and intrinsic lateral drift-field photodiode pixels. 2011 Proceedings of the ESSCIRC (ESSCIRC), IEEE, Helsinki, pp. 111–114.
- Spivak, A., Teman, A., Belenky, A., Yadid-Pecht, O., Fish, A., 2012. Low-voltage 96 db snapshot CMOS image sensor with 4.5 NW power dissipation per pixel. *Sensors* 12 (8), 10067–10085.
- Stopa, D., Simoni, A., Gonzo, L., Gottardi, M., Betta, G.D., 2002. Novel CMOS image sensor with a 132-db dynamic range. *IEEE J. Solid State Circuits* 37 (12), 1846–1852.
- Strom, G.G., Henderson, R.K., Hurwitz, J.E.D., Renshaw, D., Findlater, K.M., Purcell, M.D., 2006. Extended dynamic range from a combined linear-logarithmic CMOS image sensor. *IEEE J. Solid State Circuits* 41, 2095–2106.
- Suntharalingam, V., Rathman, D.D., Prigozhin, G., Kissel, S., Bautz, M., 2007. Back-illuminated three-dimensionally integrated CMOS image sensors for scientific applications. In: Proceedings of SPIE 6690, Focal Plane Arrays for Space Telescopes III, p. 669009.
- Takahashi, H., Kinoshita, M., Morita, K., Shirai, T., Sato, T., Kimura, T., Yuzurihara, H., Inoue, S., Matsumoto, S., 2004. A 3.9- $\mu$ m pixel pitch VGA format 10-b digital output CMOS image sensor with 1.5 transistor/pixel. *IEEE J. Solid State Circuits* 39 (12), 2417–2425.
- Takahashi, T., Kaji, Y., Tsukuda, Y., Futami, S., Hanzawa, K., Yamauchi, T., Wong, P.W., Brady, F., Holden, P., Ayers, T., Mizuta, K., Ohki, S., Tatani, K., Nagano, T., Wakabayashi, H., Nitta, Y., 2017. A 4.1Mpix 280fps stacked CMOS image sensor with array-parallel ADC architecture for region control. 2017 Symposium on VLSI Circuits, IEEE, pp. C244–C245.
- Takayanagi, I., Nakamura, J., 2012. High-resolution CMOS video image sensors. *Proc. IEEE* 101 (1), 1–13.
- 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.
- Takeshita, H., Sawada, T., Iida, T., Yasutomi, K., Kawahito, S., 2010. High-speed charge transfer pinned-photodiode for a CMOS time-of-flight range image sensor. In: Proceedings of SPIE 7536, Sensors, Cameras, and Systems for Industrial/Scientific Application XI, 7536, SPIE, p. 9. vol.
- Tian, H., Fowler, B., Gamal, A.E., 2001. Analysis of temporal noise in CMOS photodiode active pixel sensor. *IEEE J. Solid State Circuits* 36 (1), 92–101.
- Tochigi, Y., Hanzawa, K., Kato, Y., Kuroda, R., Mutoh, H., Hirose, R., Tominaga, H., Takubo, K., Kondo, Y., Sugawa, S., 2012. A global-shutter CMOS image sensor with read-out speed of 1 Tpixel/s burst and 780 Mpixel/s continuous. In: 2012 IEEE Int. Solid-State Circuits Conference, vol. 43. IEEE, pp. 382–384.
- Tochigi, Y., Hanzawa, K., Kato, Y., Kuroda, R., Mutoh, H., Hirose, R., Tominaga, H., Takubo, K., Kondo, Y., Sugawa, S., 2013. A global-shutter CMOS image sensor with read-out speed of 1-Tpixel/s burst and 780-Mpixel/s continuous. *IEEE J. Solid State Circuits* 48 (1), 329–338.
- Tournier, A., Leverd, F., Favennec, L., Perrot, C., Pinzelli, L., Gatefait, M., Cherault, N., Jeanjean, D., Carrere, J., Hirigoyen, F., Grant, L., Roy, F., 2011. Pixel-to-pixel isolation by deep trench technology: application to CMOS image sensor. Proceedings of 2011 International Image Sensor Workshop, IISW, Hokkaido, Japan, pp. 5–8.
- Toyama, T., Mishina, K., Tsuchiya, H., Ichikawa, T., Iwaki, H., Gendai, Y., Murakami, H., Takamiya, K., Shiroshita, H., Muramatsu, Y., Furusawa, T., 2011. A 17.7 Mpixel 120

- fps CMOS image sensor with 34.8 Gb/s readout. 2011 IEEE Int. Solid State Circuits Conference, IEEE, pp. 420–422.
- Turchetta, R., Guerrini, N., Sedgwick, I., 2011. Large area CMOS image sensors. *J. Instrum.* 6 (1), C01099.
- Wang, T., 2016. A Study on High-Speed Low-Noise Readout Architectures and Column A/D Converters for CMOS Image Sensors (Ph.D. thesis). Shizuoka University.
- Wang, X., Wong, W., Hornsey, R., 2006. A high dynamic range CMOS image sensor with in pixel light-to-frequency conversion. *IEEE Trans. Electron Devices* 53 (12), 2988–2992.
- Wang, X., Bogaerts, J., Vanhorebeek, G., Ruythoren, K., Ceulemans, B., Lepage, G., Willems, P., Meynants, G., 2010. A 2.2M CMOS image sensor for high-speed machine vision applications. In: Proceedings of Sensors, Cameras, and Systems for Industrial/Scientific Application XI, vol. 7536, p. 7.
- Willems, P., Meynants, G., Vanhorebeek, G., Ma, C., 2012. High-speed VGA resolution CMOS image sensor with global shutter. In: Widenhorn, R., Nguyen, V., Dupret, A. (Eds.), Proc. SPIE-IS&T Electron. Imaging, 8298, SPIE, p. 7. V vol.
- Yadid-Pecht, O., 1999. Wide dynamic range sensors. *Opt. Eng.* 38 (10), 1650–1660.
- Yadid-Pecht, O., Fossum, E.R., 1997. Image sensor with ultra-high-linear-dynamic range utilizing dual output CMOS active pixel sensors. *IEEE Trans. Electron Devices* 44 (10), 1721–1723.
- Yadid-Pecht, O., Ginosar, R., Shacham-Diamand, Y., 1991. A random access photodiode array for intelligent image capture. *IEEE Trans. Electron Devices* 38 (8), 1772–1781.
- Yamashita, Y., Takahashi, H., Kikuchi, S., Ota, K., Fujita, M., Hirayama, S., Kanou, T., Hashimoto, S., Momma, G., Inoue, S., 2011. 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.
- Yang, D.X.D., Gammal, A.E., Fowler, B., Tian, H., 2000. A  $640 \times 512$  CMOS image sensor with ultrawide dynamic range floating point pixel level ADC. *IEEE J. Solid State Circuits* 34 (12), 1821–1999.
- Yasuda, T., Hamamoto, T., Aizawa, K., 2003. Adaptive-integration-time image sensor with real-time reconstruction function. *IEEE Trans. Electron Devices* 50 (1), 111–120.
- Yasutomi, K., Itoh, S., Kawahito, S., 2011. A two-stage charge transfer active pixel CMOS image sensor with low-noise global shuttering and a dual-shuttering mode. *IEEE Trans. Electron Devices* 58 (3), 740–747.
- Yu, H., Qian, X., Guo, M., Chen, S., 2016. An antivibration time-delay integration CMOS image sensor with online deblurring algorithm. *IEEE Trans. Circuits Syst. Video Technol.* 26 (8), 1544–1554.
- Zhang, H., Yu, N., Lyu, N., Wang, F., 2016. Linear CMOS image sensor for bio-microfluidic imaging system. In: 2016 13th IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT), IEEE, pp. 1321–1323.
- 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 (10), 1759–1763.

# Smart cameras on a chip: Using complementary metal-oxide-semiconductor (CMOS) image sensors to create smart vision chips<sup>☆</sup>

6

D. Ginhac

Université de Bourgogne, Dijon, France

## 6.1 Introduction

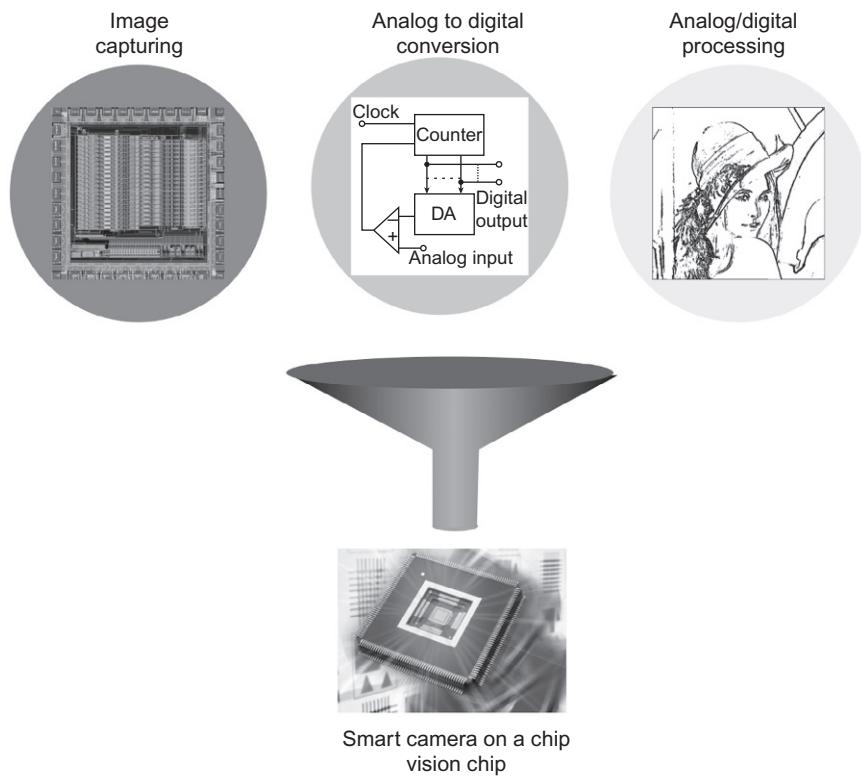
Today, digital smart cameras are rapidly becoming ubiquitous, due to reduced costs and the increasing demands of multimedia applications. Improvements in the growing digital imaging world continue to be made with two main image sensor technologies: charge coupled devices (CCDs) and complementary metal-oxide-semiconductor (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 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-Dessouki 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 (PEs) to one or more columns, or at pixel level by integrating a specific processing unit in each pixel (El El Gamal et al., 1999; El Gamal and Eltoukhy,

<sup>☆</sup> This chapter is a reprint of the chapter originally published in the first edition of “High Performance Silicon Imaging: Fundamentals and Applications of CMOS and CCD Sensors.”

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 (A/D) 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 PEs to one or more columns [typically analog to digital converters (ADCs)], 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



**Fig. 6.1** Conceptual view of a smart camera on a chip.

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 (CNNs) 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 PE 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.

Finally, we set out recent trends on smart vision chips. From a technological point of view, three-dimensional (3D) integrated imagers, based on 3D stacking technology, become an emerging solution to design powerful imaging systems because the sensor, the analog-to-digital converters (ADCs), 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 ADCs 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 (TOF) 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 A/D 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, A/D 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 PEs and
- to minimize the data transfer between cameras and the outside world by only outputting extracted feature information (Zhang et al., 2011b).

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  $320 \times 256$ -pixel CMOS sensor, a 32-bit RISC processor, a neural co-processor, a 10-bit ADC and I/O on to a  $6 \times 6$  mm<sup>2</sup> single chip in a 0.35- $\mu\text{m}$  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 micro-processors in each pixel. Processing units can be realized at chip level (system-on-chip approach), at column level by dedicating PEs to one or more columns (typically ADCs), 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 PE 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, PE area, and fill-factor. This implies various points of view (Ginhac et al., 2008):

- For a fixed fill-factor and a given PE 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 PE area, the photodiode area and the fill-factor increase as technology scales since the area occupied by the pixel transistors in each PE 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 PE 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  $\mu\text{m}$  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

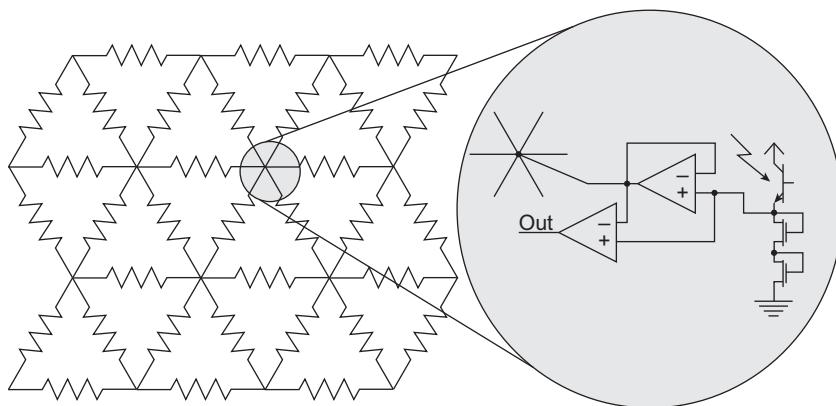
From an historical point of view, the pioneering works have concentrated efforts on spatial and spatio-temporal image processing vision chips in the field of machine vision applications. These vision chips can be viewed as the first smart sensors and came to light during the 1990s under the well-known 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 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, that is, 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. [Delbrück \(1993\)](#) describes a two-dimensional (2D) 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 (1D) velocity sensor that uses the gradient method for spatially resolved velocity computation. [Etienne-Cummings et al. \(1999\)](#) implement a retina for measuring 2D visual motion with two 1D detectors. The pixels are built around a general-purpose analog neural computer and a silicon retina. Motion is extracted in 2D 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



**Fig. 6.2** Architecture of Mead's artificial retina.

From Moini, A., 2000. Vision Chips. Boston; London: Kluwer Academic.

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.

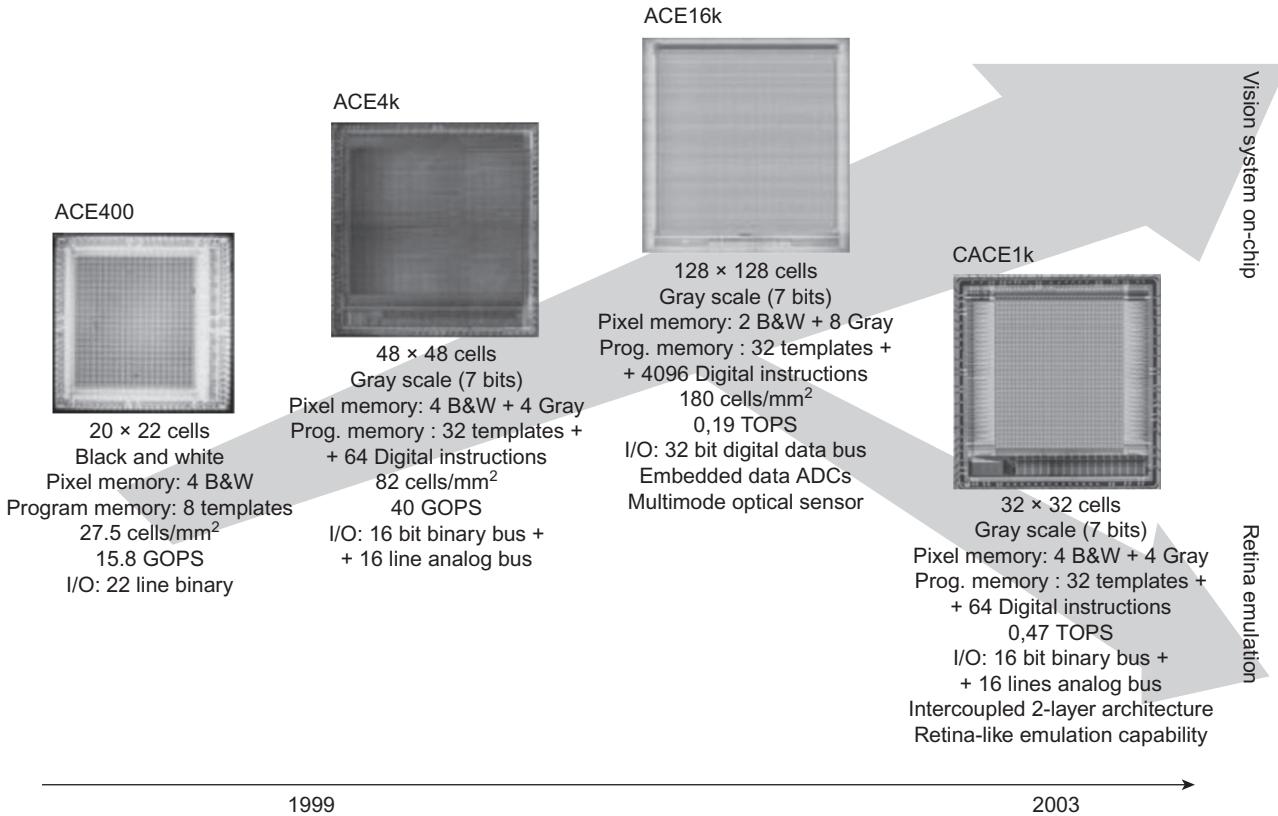
## 6.4 From special-purpose chips to smart computational chips

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 PE, known as computational chip.

For this purpose, the processing function can be based on the paradigm of CNNs, 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  $64 \times 64$  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  $32 \times 32$  grid corresponding to approximately 500,000 transistors in a standard  $0.5 \mu\text{m}$  CMOS technology. An enhanced  $128 \times 128$  version was also described in Rodriguez-Vazquez et al. (2004). The chip designed in a  $0.35 \mu\text{m}$  standard CMOS technology, contains about 3.75 million transistors and exhibits peak computing figure of 330 GOPS. Each PE in the array contains a reconfigurable computing kernel capable of calculating linear convolutions on  $3 \times 3$  neighborhoods in less than  $1.5 \mu\text{s}$ , Boolean combinations in less than 200 ns, arithmetic operations in about  $5 \mu\text{s}$ , and CNN-like temporal evolutions with a time constant of about  $0.5 \mu\text{s}$ . 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



**Fig. 6.3** Historical roadmap of CNN-based vision chips.

From Rodriguez-Vazquez, A., Linan-Cembranon, G., Carranza, L., Roca-Moreno, E., Carmona-Galan, R., Jimenez-Garrido, F., Dominguez-Castro, R. and Meana, S., 2004. Ace16 k: the third generation of mixed-signal SIMD-CNN ACE chips toward VSOCs. IEEE Trans. Circuits Syst. 51(5), 851–863.

vision chips, the pixel size is often over  $100 \times 100 \mu\text{m}$ . 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.

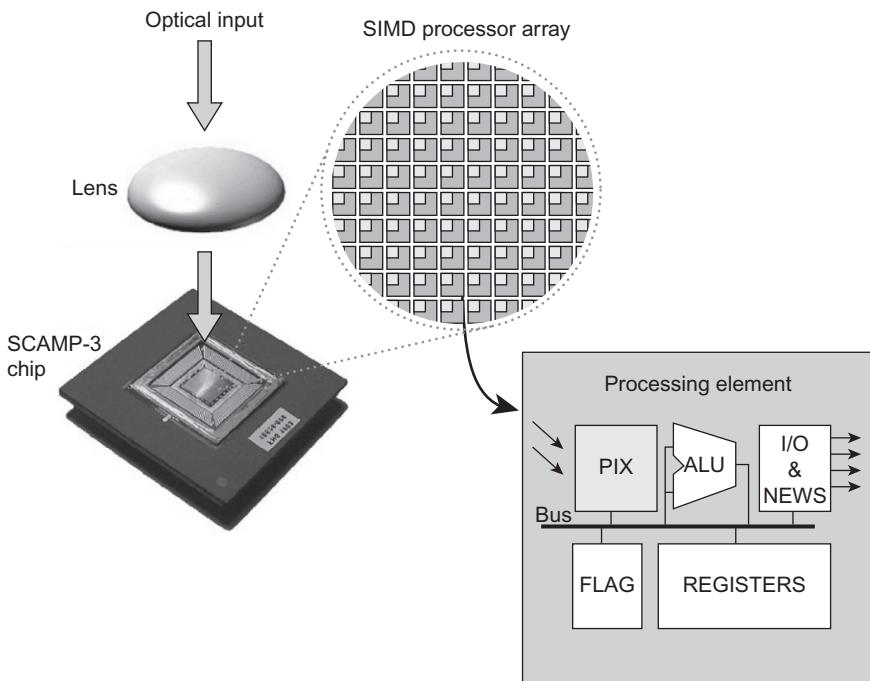
Other architectures in this category are the SCAMP family—SIMD current-mode analog matrix processor (Dudek, 2005; Dudek and Carey, 2006) of software-programmable SIMD (single instruction, multiple data) processor arrays implementing a variety of low-level image processing tasks. A SIMD processor array is built around multiple PEs 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  $0.35 \mu\text{m}$  CMOS technology contains a  $128 \times 128$  processor array and achieves cell density of 410 processors/ $\text{mm}^2$  (a single cell measures under  $50 \times 50 \mu\text{m}^2$ ).

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  $0.35 \mu\text{m}$  CMOS process.

The ASPA family (Lopich and Dudek, 2008) includes vision chips embedding fine-grain processor arrays based on novel control schemes, 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  $160 \times 80$  processor array fabricated in a  $180 \text{ nm}$  CMOS technology with a chip area of  $50 \text{ mm}^2$ .

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  $128 \times 128$  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.

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.



**Fig. 6.4** Overview of the SCAMP architecture.

From Dudek, P., Barr, D., Lopich, A., 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.

## 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 PE 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 PEs 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

**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)	ACE16 k (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 mm <sup>2</sup>	16 mm <sup>2</sup>	145 mm <sup>2</sup>
Pixel pitch	49.35 × 49.35 µm	45.6 × 45 µm	60 × 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 mm <sup>2</sup>	49 mm <sup>2</sup>	76 mm <sup>2</sup>
Pixel pitch	100 × 117 µm	67.4 × 67.4 µm	60 × 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

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 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 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 G-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

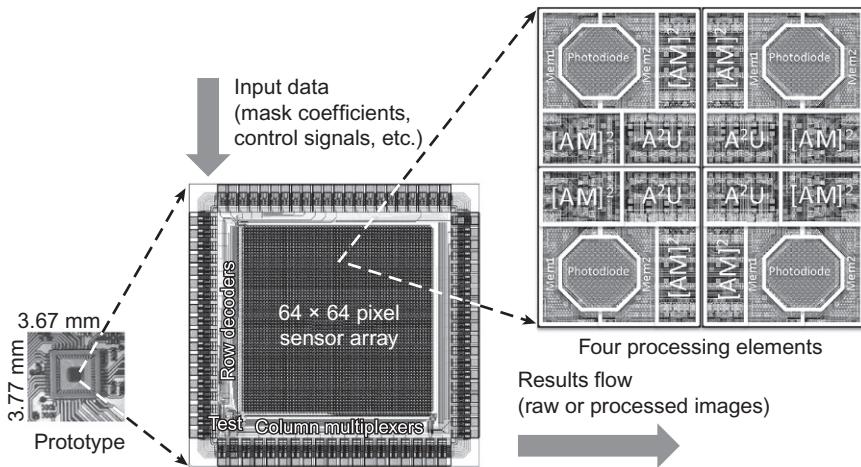
the performance of existing video applications such as motion vector estimation (Handoko et al., 2000; Lim and El Gamal, 2001; Liu and El Gamal, 2001a), multiple capture with dynamic range (Yang et al., 1999; Yadid-Pecht and Belenky, 2001; Stoppa et al., 2002), motion capture (Liu and El Gamal, 2001b), 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 PEs. 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 PE 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

- to capture images at a very high framerate
- 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
- 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 2D array of  $64 \times 64$  identical 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



**Fig. 6.5** Overview of the high-speed vision chip with an analog processor-per-pixel array. From Ginhac, D., Dubois, J., Paindavoine, M., Heyrman, B., 2008. A SIMD programmable vision chip with high speed focal plane image processing, EURASIP J. Embed. Syst. 2008, 961315.

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,
- 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  $2 \times 2$  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  $35 \times 35 \mu\text{m}$  ( $1225 \mu\text{m}^2$ ). The active area of the photodiode is  $300 \mu\text{m}^2$ , 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 2000–5000 frames per second.

Based on the same principle, a large number of equivalent pixel-level image sensors have been designed during the past 10 years, taking advantage of pixel-level PEs 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 A/D 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., 2011a), 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  $\mu\text{m}$  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, ADCs and digital programmable PEs 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. 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 PEs 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 3D 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.

## References

- Acosta-Serafini, P., Ichiro, M., Sodini, C., 2004. A  $\frac{1}{3}$  VGA linear wide dynamic range CMOS image sensor implementing a predictive multiple sampling algorithm with overlapping integration intervals. *IEEE J. Solid State Circuits* 39 (9), 1487–1496.
- Albani, L., Chiesa, P., Covi, D., Pedegani, G., Sartori, A., Vatteroni, M., 2002. Visoc: a smart camera soc. In: Proceedings of the 28th European Solid-State Circuits Conference, ESSCIRC 2002, pp. 367–370.
- Aw, C.H., Wooley, B., 1996. A  $128 \times 128$ -pixel standard-CMOS image sensor with electronic shutter. *IEEE J. Solid State Circuits* 31 (12), 1922–1930.
- Bigas, M., Cabruja, E., Forest, J., Salvi, J., 2006. Review of CMOS image sensors. *Microelectron. J.* 37 (5), 433–451.
- Brajovic, V., 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, pp. 1638–1643.
- Brea, V.M., Vilariño, D.L., Paasio, A., Cabello, D., 2004. Design of the processing core of a mixed-signal CMOS DTCNN chip for pixel-level snakes. *IEEE Trans. Circuits Syst. I* 51 (5), 997–1013.
- Cembrano, G., Rodriguez-Vazquez, A., Galan, R., Jimenez-Garrido, F., Espejo, S., Domniguez-Castro, R., 2004. A 1000 fps  $128 \times 128$  vision processor with 8-bit digitized I/O. *IEEE J. Solid State Circuits* 39 (7), 1044–1055.
- Chi, Y., Mallik, U., Clapp, M.A., Choi, E., Cauwenberghs, G., 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., Yang, L., 1988a. Cellular neural networks: theory. *IEEE Trans. Circuits Systems* 35 (10), 1257–1272.
- Chua, L., Yang, L., 1988b. Cellular neural networks: applications. *IEEE Trans. Circuits Systems* 35 (10), 1273–1290.
- Czuni, L., 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, pp. 51–56.
- Decker, S., McGrath, D., Brehmer, K., Sodini, C., 1998. A  $256 \times 256$  CMOS imaging array with wide dynamic range pixels and column-parallel digital output. *IEEE J. Solid State Circuits* 33 (12), 2081–2091.
- Delbruck, T., 1993. Silicon retina with correlation-based, velocity-tuned pixels. *IEEE Trans. Neural Netw.* 4 (3), 529–541.
- Deutschmann, R., 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, pp. 649–652.
- Dickinson, A., Mendis, S., Inglis, D., Azadet, K., 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., Heyrman, B., 2008. A 10 000 fps CMOS sensor with massively parallel image processing. *IEEE J. Solid State Circuits* 43 (3), 706–717.
- Dudek, P., 2005. Implementation of SIMD vision chip with  $128 \times 128$  array of analogue processing elements. In: IEEE International Symposium on Circuits and Systems, ISCAS. vol. 6, pp. 5806–5809.
- Dudek, P., 2006. An asynchronous cellular logic network for trigger-wave image processing on fine-grain massively parallel arrays. *IEEE Trans. Circuits Syst. Express Briefs* 53 (5), 354–358.
- Dudek, P., Carey, S., 2006. General-purpose  $128 \times 128$  SIMD processor array with integrated image sensor. *Electron. Lett.* 42 (12), 678–679.
- Dudek, P., Hicks, P.J., 2001. An analogue SIMD focal-plane processor array. In: Proceedings of the International Symposium on Circuits and Systems, ISCAS. vol. 4, pp. 490–493.
- Dudek, P., Hicks, P.J., 2005. A general-purpose processor-per-pixel analog SIMD vision chip. *IEEE Trans. Circuits Syst.* 52 (1), 13–20.
- El Gamal, A., Eltoukhy, H., 2005. CMOS image sensors. *IEEE Circuits Devices Mag.* 21 (3), 6–20.
- El Gamal, A., Yang, D., Fowler, B., 1999. Pixel level processing—why, what and how? In: Proceedings of the SPIE Electronic Imaging '99 Conference. vol. 3650, pp. 2–13.
- El-Desouki, M., Jamal Deen, M., Fang, Q., Liu, L., Tse, F., Armstrong, D., 2009. CMOS image sensors for high speed applications. *Sensors* 9 (1), 430–444.
- Espejo, S., Carmona, R., Dominguez-Castro, R., Roriguez-Vazquez, A., 1996. A  $0.8 \mu\text{m}$  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, pp. 280–283.
- Espejo, S., Dominguez-Castro, R., Linan, G., Roriguez-Vazquez, A., 1998. A  $64 \times 64$  CNN universal chip with analog and digital I/O. In: IEEE International Conference on Electronics, Circuits and Systems. vol. 1, pp. 203–206.
- Etienne-Cummings, R., Kalayjian, Z., Cai, D., 2001. A programmable focal-plane MIMD image processor chip. *IEEE J. Solid State Circuits* 36 (1), 64–73.
- Etienne-Cummings, R., Van Der Spiegel, J., Mueller, P., 1999. Hardware implementation of a visual-motion pixel using oriented spatiotemporal neural filters. *IEEE Trans. Circuits Syst. II* 46 (9), 1121–1136.
- Etienne-Cummings, R., Van Der Spiegel, J., Mueller, P., Zhang, M., 2000. A foveated silicon retina for two-dimensional tracking. *IEEE Trans. Circuits Syst. II* 47 (6), 504–517.

- Fontaine, R., 2011. A review of the 1.4  $\mu\text{m}$  pixel generation. In: International Image Sensor Workshop, Hokkaido, Japan.
- Fossum, E., 1993. Active pixel sensors: are CCDs dinosaurs? In: International Society for Optical Engineering (SPIE). 1900, pp. 2–14.
- Fossum, E., 1997. CMOS image sensor: electronic camera on a chip. *IEEE Trans. Electron Devices* 44 (10), 1689–1698.
- Galan, R., Jimenez-Garrido, F., Dominguez-Castro, R., Espejo, S., Roska, T., Rekeczky, C., Petras, I., Rodriguez-Vazquez, A., 2003. A bio-inspired two-layer mixed-signal flexible programmable chip for early vision. *IEEE Trans. Neural Netw.* 14 (5), 1313–1336.
- Gao, Q., Yadid-Pecht, O., 2012. A low-power block-based CMOS image sensor with dual vdd. *IEEE Sensors J.* 12 (4), 747–755.
- Ginhac, D., Dubois, J., Heyrman, B., Paindavoine, M., 2010. A high-speed programmable focal-plane SIMD vision chip. *Analog Integr. Circ. Sig. Process* 65, 389–398.
- Ginhac, D., Dubois, J., Paindavoine, M., Heyrman, B., 2008. A SIMD programmable vision chip with high speed focal plane image processing. *EURASIP J. Embed. Syst.* 2008, 961315.
- Handoko, D., Takokoro, K., Kumahara, M., 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., Syslverter, D., 2010. A 0.5 v sub-microwatt CMOS image sensor with pulse-width modulation read-out. *IEEE J. Solid State Circuits* 45 (4), 759–767.
- Hong, C., 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., Fossum, E.R., 1997. Multiresolution image sensor. *IEEE Trans. Circuits Syst. Video Technol.* 7 (4), 575–583.
- Kim, D., Cho, J., Lim, S., Lee, D., Han, G., 2008. A 5000 S/s single-chip smart eye-tracking sensor. In: IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers, pp. 46–594.
- Kitchen, A., Bermak, A., Bouzerdoum, A., 2005. A digital pixel sensor array with programmable dynamic range. *IEEE Trans. Electron Devices* 52 (12), 2591–2601.
- Kleinfelder, S., Lim, S., Liu, X., El Gamal, A., 2001. A 10 000 frames/s CMOS digital pixel sensor. *IEEE J. Solid State Circuits* 36 (12), 2049–2059.
- Komuro, T., Iwashita, A., Ishikawa, M., 2009. A QVGA-size pixel-parallel image processor for 1000-fps vision. *IEEE Micro* 29 (6), 58–67.
- Komuro, T., Kagami, S., Ishikawa, M., 2004. A dynamically reconfigurable SIMD processor for a vision chip. *IEEE J. Solid State Circuits* 39 (1), 265–268.
- Kozlowski, L., Rossi, G., Blanquart, L., Marchesini, R., Huang, Y., Chow, G., Richardson, J., Standley, D., 2005. Pixel noise suppression via SoC management of target reset in a  $1920 \times 1080$  CMOS image sensor. *IEEE J. Solid State Circuits* 40 (12), 2766–2776.
- Krymski, A., Tu, N., 2003. A 9-v/lux-s 5000-frames/s  $512 \times 512$  CMOS sensor. *IEEE Trans. Electron Devices* 50 (1), 136–143.
- Krymski, A., Van Blerkom, D., Andersson, A., Bock, N., Mansoorian, B., Fossum, E., 1999. A high speed, 500 frames/s,  $1024 \times 1024$  CMOS active pixel sensor. In: 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., Chung, C.H., 2011. A  $192 \times 108$  pixel ToF-3D image sensor with single-tap concentric-gate demodulation pixels in 0.13  $\mu\text{m}$  technology. In: IEEE International Electron Devices Meeting (IEDM), pp. 8.7.1–8.7.4.
- Leon-Salas, W.D., Balkir, S., Sayood, K., Schemm, N., 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., El Gamal, A., 2001. Integrating image capture and processing—beyond single chip digital camera. In: Proceedings of the SPIE Electronic Imaging ‘2001 Conference, San Jose, CA. vol. 4306.
- Lin, Q.-Y., Miao, W., Zhang, W.-C., Fu, Q.-Y., 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. vol. 35. *Photonics Spectra*, pp. 154–158.
- Liu, X., 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, San Jose, CA. vol. 4306.
- Liu, X., 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, pp. 1841–1844.
- Loinaz, M., Singh, K., Blanksby, A., Inglis, D., Azadet, K., Ackland, B., 1998. A 200 mv 3.3 v CMOS color camera IC producing  $352 \times 288$  24-b video at 30 frames/s. *IEEE J. Solid State Circuits* 33 (12), 2092–2103.
- Lopich, A., Dudek, P., 2008. ASPA: focal plane digital processor array with asynchronous processing capabilities. In: IEEE International Symposium on Circuits and Systems, ISCAS 2008, pp. 1592–1595.
- Lopich, A., Dudek, P., 2011. Asynchronous cellular logic network as a co-processor for a general purpose massively parallel array. *Int. J. Circuit Theory Appl.* 39 (9), 963–972.
- Martin, D., Lee, H.-S., Masaki, I., 1998. A mixed-signal array processor with early vision applications. *IEEE J. Solid State Circuits* 33 (3), 497–502.
- Massari, N., Gottardi, M., 2007. A 100 dB dynamic-range CMOS vision sensor with programmable image processing and global feature extraction. *IEEE J. Solid State Circuits* 42 (3), 647–657.
- Mead, C., Mahowald, M., 1988. A silicon model of early visual processing. *Neural Netw.* 1 (1), 91–97.
- Miao, W., Lin, Q., Zhang, W., 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., Ginhac, D., 2008. Nonlinear systems for image processing. In: Hawkes, P.W. (Ed.), Advances in Imaging and Electron Physics. In: vol. 152. Elsevier, pp. 79–151.
- Motoyoshi, M., Koyanagi, M., 2009. 3D-LSI technology for image sensor. *J. Instrum.* 4(03) P03009.
- Murari, K., Etienne-Cummings, R., Thakor, N., Cauwenberghs, G., 2009. Which photodiode to use: a comparison of CMOS-compatible structures. *IEEE Sensors J.* 9 (7), 752–760.
- Nixon, R.H., Kemeny, S.E., Staller, C.O., Fossum, E.R., 1995.  $128 \times 128$  CMOS photodiode-type active pixel sensor with on-chip timing, control, and signal chain electronics. In: Blouke, M.M. (Ed.), Proceedings of the SPIE Charge-Coupled Devices and Solid State Optical Sensors V. vol. 2415, pp. 117–123.
- Ohta, J., 2008. Smart CMOS Images Sensors and Applications. CRC Press, Boca Raton.
- Paillet, F., Mercier, D., Bernard, T., 1998. Making the most of 15 k-lambda-2 silicon area for a digital retina. In: Proceedings of SPIE. vol. 3410, pp. 158–167.
- Paillet, F., Mercier, D., Bernard, T., 1999. Second generation programmable artificial retina. In: Proceedings of the Twelfth Annual IEEE International ASIC/SOC Conference, pp. 304–309.
- Rekeczky, C., Tahy, A., Végh, Z., Roska, T., 1999. CNN-based spatio-temporal nonlinear filtering and endocardial boundary detection in echocardiography. *Int. J. Circuit Theory Appl.* 27 (1), 171–207.

- Rinner, B., Wolf, W., 2008. An introduction to distributed smart cameras. *Proc. 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., Meana, S., 2004. Ace16 k: the third generation of mixed-signal SIMD-CNN ACE chips toward VSOCs. *IEEE Trans. Circuits Syst.* 51 (5), 851–863.
- Roska, T., 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, pp. 120–123.
- Sakakibara, M., Kawahito, S., Handoko, D., Nakamura, N., Higashi, M., Mabuchi, K., Sumi, H., 2005. A high-sensitivity CMOS image sensor with gain-adaptive column amplifiers. *IEEE J. Solid State Circuits* 40 (5), 1147–1156.
- Schrey, O., Huppertz, J., Filimonovic, G., Bussmann, A., Brockherde, W., Hosticka, B., 2002. A 1 k × 1 k high dynamic range CMOS image sensor with on-chip programmable region-of-interest readout. *IEEE J. Solid State Circuits* 37 (7), 911–915.
- Seitz, P., 2000. Solid-state image sensing. In: Jähne, B., Haußecker, H., Geißler, P. (Eds.), *Handbook of Computer Vision and Applications*. In: vol. 1. Academic Press, San Diego, pp. 165–222.
- Seitz, P., Theuwissen, A. (Eds.), 2011. *Single Photon Imaging*. In: *Springer Series in Optical Sciences*, Springer-Verlag, Heidelberg.
- Shi, Y., 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., Murray, A., Anderson, S., Denyer, P., 1998. A single-chip 306 × 244-pixel CMOS NTSC video camera. In: *ISSCC Digest of Technical Papers*, San Fransisco, CA, pp. 170–171.
- Stevanovic, N., Hillebrand, M., Hosticka, B., Teuner, A., 2000. A CMOS image sensor for high-speed imaging. In: *IEEE International Solid-State Circuits Conference, Digest of Technical Papers*, pp. 104–105.
- Stopa, D., Somoni, A., Gonzo, L., Gottardi, M., Dalla Betta, G.F., 2002. Novel CMOS image sensor with a 132-dB dynamic range. *IEEE J. 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*, pp. 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, P.A., Mansoorian, B., Shaver, D., 2009. A 4-side tileable back illuminated 3D-integrated Mpixel CMOS image sensor. In: *IEEE International Solid-State Circuits Conference, Digest of Technical Papers, ISSCC 2009*, pp. 38–39 (39a).
- Wu, C.-Y., 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 J.* 4 (4), 510–518.
- Wuu, S., Chien, H., Yaung, D., Tseng, C., Wang, C., Chang, C., Hsiao, 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., 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.
- Xu, C., Zhang, W., Ki, W.-H., Chan, M., 2002. A 1.0-v vdd CMOS active-pixel sensor with complementary pixel architecture and pulsedwidth modulation fabricated with a 0.25-μm cmos process. *IEEE J. Solid State Circuits* 37 (12), 1853–1859.

- Yadid-Pecht, O., 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, pp. 100–101.
- Yadid-Pecht, O., Belenky, A., 2003. In-pixel autoexposure CMOS APS. *IEEE J. Solid State Circuits* 38 (8), 1425–1428.
- Yang, D., El Gamal, A., Fowler, B., Tian, H., 1999. A  $640 \times 512$  CMOS image sensor with ultra wide dynamix range floating-point pixel-level ADC. *IEEE J. Solid State Circuits* 34, 1821–1834.
- Yeh, S., Lin, J., Hsieh, C., Yeh, K., Li, C., 2011. A new CMOS image sensor readout structure for 3D integrated imagers. In: IEEE Custom Integrated Circuits Conference (CICC), pp. 1–4.
- Yoon, K., Kim, C., Lee, B., Lee, D., 2002. Single-chip CMOS image sensor for mobile applications. *IEEE J. Solid State Circuits* 37, 1839.
- Zhang, M., 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., Wu, N.J., 2011b. A programmable vision chip based on multiple levels of parallel processors. *IEEE J. Solid State Circuits* 46 (9), 2132–2147.
- Zhang, X., Leomant, S., Lau, K.L., Bermak, A., 2011a. A compact digital pixel sensor (DPS) using 2T-DRAM. *J. Low Power Electron. Appl.* 1 (1), 77–96.

## Further reading

- Dudek, P., Barr, D., Lopich, A., Carey, S., 2006. Demonstration of real-time image processing on the SCAMP-3 vision system. In: 10th International Workshop on Cellular Neural Networks and Their Applications, CNNA ‘06.

This page intentionally left blank

# **Part Two**

# **Applications**

This page intentionally left blank

# CMOS image sensor technology advances for mobile devices

7

*Robert J. Gove*  
Synaptics Inc., San Jose, CA, United States

## 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 mobile devices became the fastest growing product category of semiconductor wafers in history, ramping to nearly a million silicon 8" wafers or the equivalent capacity of five wafer fabrication facilities (fabs) over the period of 2003–2005. At this point, in 2017, this mobile image sensor market has doubled several times to a level of four million 12" equivalent wafers shipped per year.

This section describes the enabling technologies that helped position image capture and sharing as the second most used feature (after voice and text) in this “three-billion-unit-per-year” market. Later applications of CMOS mobile imaging, including notebooks, tablets, and handheld gaming devices, drove new capabilities like ultrathin/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-performance and high-resolution camera for general photography or videography needs. However, the secondary, user-facing video camera found on most smartphones uses CMOS image sensors similar to those used in notebooks, tablets, and gaming devices, with the purpose of capturing self-portrait photographs or selfies for sharing with friends on apps like Snapchat, Facebook, Instagram or Twitter, or video conferencing like Apple’s FaceTime. We will explore both sensor types.

Globally, nearly 1.5 billion smartphones shipped in 2017, all of which include one or more cameras. While many features and capabilities differentiate the large number of smartphone products in the market, clearly most smartphone purchasers continue to value the product’s picture and video-taking ability far above most other capabilities. In addition to the smartphone category, non-smartphones shipped an additional 500 million phones, many of those also have cameras, for a total of nearly 2 billion phones in 2017. Also interesting, 99% of all cameras sold are embedded in mobile products, with compact cameras, DSLRs, mirror-less, and others only at 1% of the market.

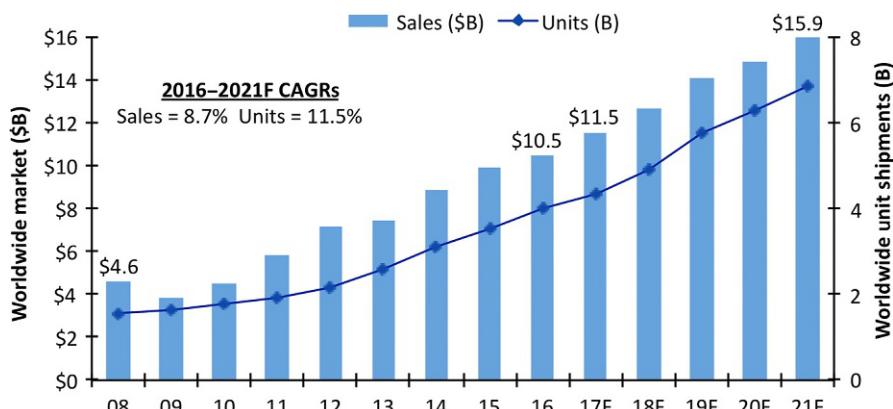
All total, almost 7 billion cell phones are in use today and it’s been estimated that 1.2 trillion photographs were taken in 2017, mostly using these mobile devices. In addition, about 160 million tablets ship per year, most with two mobile cameras.

On average, smartphones have 2.3 cameras per unit today, growing to over 3 and beyond in a few years. Clearly today, mobile interconnected phones and devices with cameras dominate the market for consumer photography. Some interesting camera phone market trends include: (1) significant performance and feature differentiation for the rear-facing (RF) scene camera and front-facing (FF) selfie cameras, (2) minimization of the phone's front display bezel driving to smaller cameras, and (3) addition of advanced three-dimensional (3D) sensing depth cameras, enabling a variety of new interesting use cases for mobile devices. So, in this chapter, we evaluate the emerging technologies for mobile photography and applications based on the image sensing.

### 7.1.1 Evolution of camera phones with CMOS image sensors

Camera phones emerged when semiconductor, optics, and packaging technologies advanced to enable creation of miniature cameras. Simultaneously, mobile phones proliferated to a point where billions of users keep 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 had 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 smartphone segment, where camera performance and features lead other segments, has grown dramatically and remains the leading segment for imaging technology innovation and shipment volume. Fig. 7.1 shows the growth of the CMOS image sensor market. With the mobile phone recently representing 70% of the total sensor sales, the total projected shipments for mobile phones in 2017 should be 3 billion units. Emerging deployment of dual cameras and additional 3D cameras will lift the mobile sensor market further in the coming years.



**Fig. 7.1** CMOS image sensor market growth, with mobile products representing 70% of the market for 3 billion shipments projected in 2017.

Source: IC Insights, June 2017.

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, imaging 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 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 (HDR) 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, multicore 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 innovative imaging applications. Fig. 7.2 highlights many such applications. Examples include:

- face detection for immediate cataloging or tagging of photos (Klik by [Face.com](#)) and face authentication using 3D cameras (Apple's iPhoneX),
- 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),
- 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),
- 3D face scanning using a 3D depth camera (such as Bellus3D) to create 360 degree, realistic face or body models in seconds, followed by online shopping to achieve a perfect clothing fit, and
- guiding a grocery shopper to quickly locate their favorite items in a grocery store using augmented reality (AR) applications, such as Apple's ARkit API, where the user follows

Camera-enabled smartphone



Mobile hand gesture detection



#### Picture taking

1. Voice actuated shutter
2. Stabilization (LumaCam app)
3. Slow shutter for motion blur (Slow shutter app)
4. High-Dynamic range and tone mapping ([Instagram.com](#))
5. Beam picture anywhere using QR code (Scalado® PhotoBeamer™)

#### Video/picture editing

1. Automatic face detect for cataloging photos
2. ClearView™ 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 ([Metairo.com](#))
4. Camera-based gesture control ([Eyesight.com](#))
5. Information enhanced picture searching ([Pixulu.com](#))
6. Augmented Reality navigation ([AccessAir.com](#), [Wikitude.com](#))

#### Multishot picture fusion

1. Panoramic image stitching (Cloudburst AutostitchPanorama)
2. Stitching multishots 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 ([ViewdI.com](#))
2. AR game with virtual liquid flow by user orientation ([ReFlow](#))

**Fig. 7.2** Mobile applications for camera-enabled smartphones.

Top photo by [istockphoto.com](#) file#000019780418 and bottom photo by eyesight at [eyesight-tech.com](#), used with permission.

interactive, location-aware waypoints on their smartphone's display ([https://www.youtube.com/watch?time\\_continue=21&v=GpvuSfOV7wo](https://www.youtube.com/watch?time_continue=21&v=GpvuSfOV7wo)).

AR 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 real and synthetic images, followed by selective overlay of useful data, such as interactive visual directions, or waypoint to waypoint, through a subway station, or, in the example above, directions to your desired item in a grocery store. 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 (the SocialCamera app from [Viewdl.com](http://Viewdl.com) tagged your photos and uploaded them to Facebook on the fly). In summary, innovative imaging-centric applications have and will continue to emerge because the mobile devices offer easy image and imaging information capture followed by abundant processing, 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 coupled 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 12 megapixel image sensors, increasing to 24 megapixels in some cameras phones.

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 thousands 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 cameras available, yet be smaller and easier to use, without regard to how the picture was taken. However, 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 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 backside 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. Fig. 7.3 shows how CMOS has overcome CCD technology, particularly in pixel size, critical to mobile applications. 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. Fig. 7.3 shows the

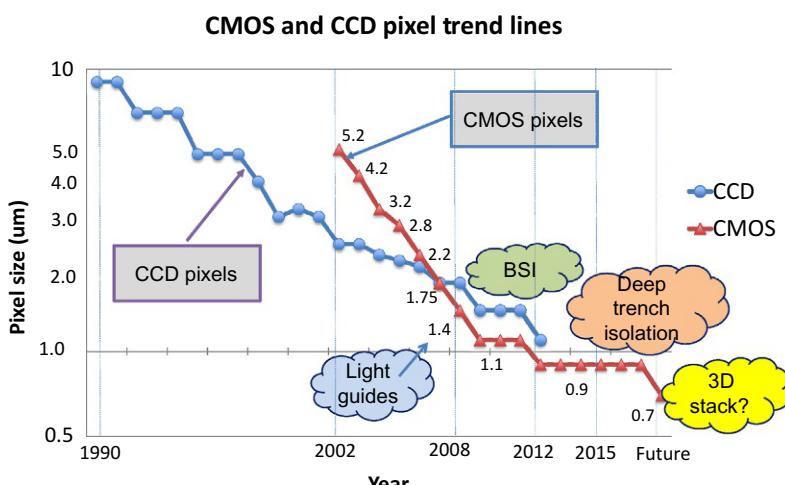


Fig. 7.3 CMOS imaging Moore’s Law trend.

industry progression of shrinking pixel sizes for mobile devices. The smallest CMOS pixels used in smartphone products today are “recently improved” 0.9- $\mu\text{m}$  sized pixels. As well, Fig. 7.3 highlights the slowing of this pixel size advancement trend. Conventional pixel architectures moved to the use of special light guides (LGs) per pixel, to BSI technology, to wafer stacking 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, 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:

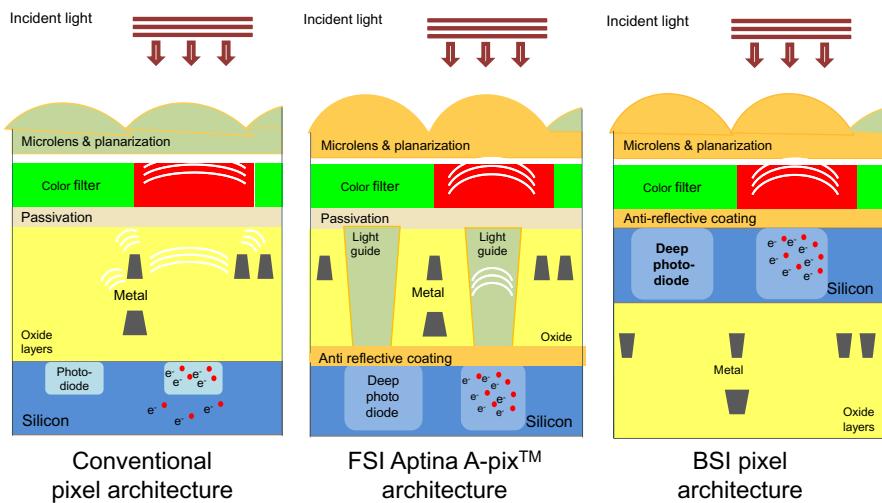
1. the underlying optimization of the fab process to create successful operation (photon collection, storage, and data transfer) of the pixel,
2. the architecture of the pixel, and
3. the analog circuits for readout of the pixels.

For mobile applications, the limitations were plentiful 15 years ago. Initial mobile sensor pixels were 5.6  $\mu\text{m}$  on a side, producing a video graphics array (VGA) image

Categories for Mobile CMOS Imaging Core Technologies				
Process	Pixel	Sensor	Image Computing and System	Optics and Packaging
<ul style="list-style-type: none"> <li>Dark current</li> <li>Hot Pixels</li> <li>Defects</li> <li>Dzero</li> <li>Die size</li> <li>Light guides</li> <li>BSI</li> <li>Stacking</li> </ul>	<ul style="list-style-type: none"> <li>2.2<math>\mu</math>m, 1.75<math>\mu</math>m, 1.4<math>\mu</math>m, 1.1<math>\mu</math>m, 0.9<math>\mu</math>m pixels</li> <li>FSI &amp; BSI</li> <li>Low light sensitivity</li> <li>Dark current</li> <li>High-dynamic range (HDR)</li> <li>Light guide</li> <li>Color-filter-array (CFA)</li> <li>Micro-lenses</li> </ul>	<ul style="list-style-type: none"> <li>Row or column noise</li> <li>Defect replace</li> <li>ADC speed/power/watt</li> <li>High-dynamic range (HDR)</li> </ul>	<ul style="list-style-type: none"> <li>Color processing</li> <li>Shading correction</li> <li>Fast auto focus</li> <li>Image stabilization</li> <li>Zero shutter lag</li> <li>User tracking</li> <li>Co-processor ASIC</li> <li>2-chip stacking</li> <li>Software support</li> <li>Fast interface (HiSPi, MIPI, ...)</li> </ul>	<ul style="list-style-type: none"> <li>High chief-ray angle (CRA) tolerance</li> <li>Fast voice coil auto focus</li> <li>Image stabilization support</li> <li>Low-cost manufacturing</li> <li>Re-flowable module</li> <li>Wafer-level module</li> <li>Chip-Scale package</li> </ul>

Technology constraints: image capture performance, cost, size, and feature requirements

**Fig. 7.4** Categories of camera phone core technology advancements.



**Fig. 7.5** Mobile pixel architectures and technology advancement.

with a  $1/4''$  optical format array. By far, the  $1/4''$  optical format had dominated the image sensor markets due to an effective compromise of performance, cost, and size. The industry evolved to the following optical formats for mobile sensors:

- (1) a *performance-leading* node originally at  $1/3''$  and now larger formats like  $1/2.3''$ ,
- (2) a *mainstream* node at  $1/4''$ , and
- (3) a *value* node at  $1/5''$  or smaller.

Some technology areas which impeded success in this early period included color performance, crosstalk between pixels, uniformity across the sensor, color shading, and eclipse, where objectionable distortion patterns are readout of a sensor with sun in the field-of-view where electrons spill from the photodiode to the floating diffusion, causing a reset to the incorrect level due to the added electrons (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 1.12  $\mu\text{m}$  pixels, and just recently 0.9  $\mu\text{m}$ , yet 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 beyond 0.9  $\mu\text{m}$ .

### 7.2.2 *The CMOS camera semiconductor process*

We have discussed the advantages of “drafting” CMOS process technology, specifically producing image sensors by leveraging industry standards and processes already running in wafer fabrication 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 relate to achieving;

- lower dark current and
- the minimization of hot pixels.

Dark current is produced by interface states at the substrate-oxide boundary and impurities in the silicon substrate which reduce the energy required by electrons to cross the silicon band gap. The electrons produced as dark current are indistinguishable from signal electrons and therefore, contribute to the noise characteristics of the sensor (Janesick, 2001). Dark current creates a fundamental noise signal, which builds or accumulates overtime within the pixel’s photodiode. 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. Essentially, a hot pixel acts like a pixel with a high dark current rate so that it saturates with extremely short integration times. 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 dark current 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 backside 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 to 40 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 VGA resolution CMOS sensors with larger pixels have been fabricated with nearly 100% yield per wafer, whereas larger sensors may only achieve 80% yield in some cases. Clearly, attention to the process optimization can greatly influence the resultant yield and cost.

### ***7.2.2.1 Color filter arrays***

After silicon front-end fabrication, CMOS image sensors are taken to another wafer manufacturing line that can include spin-coat/bake for creating a 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.

### ***7.2.2.2 New semiconductor processes enable LG and BSI pixel architecture advancements***

Two technology developments enabled the evolution of high-performance 1.4  $\mu\text{m}$  pixels:

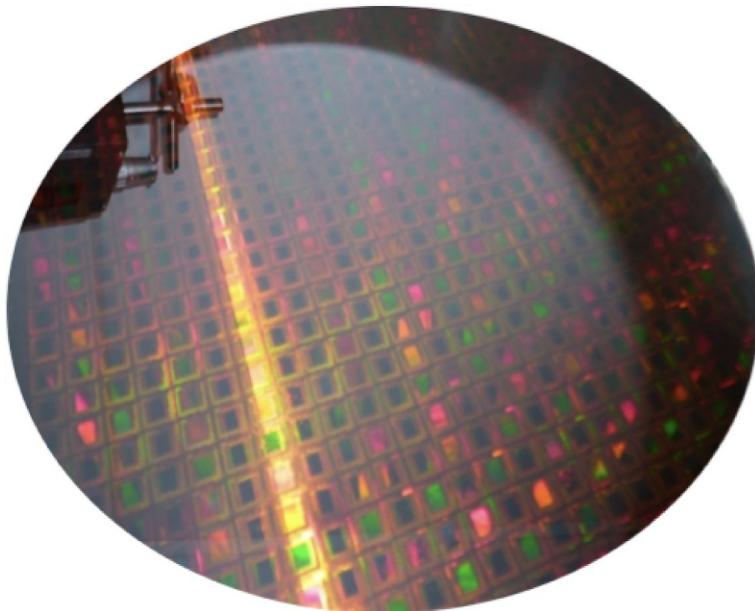
1. the LG process technology illustrated in Fig. 7.5 as A-Pix, and as found in nature in Fig. 7.6 (Franze et al., 2012), and
2. wafer-level BSI process technology (Iwabuchi et al., 2006).



**Fig. 7.6** Light guides, as found in nature, optical fibers in the retina (Muller cells, used with permission).

BSI is necessary for acceptable performance at pixels sizes of  $1.1\text{ }\mu\text{m}$  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 QE 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 ultrathin 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



**Fig. 7.7** Photograph of a thinned BSI wafer, only several microns thick.

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 is typically thinned to 100  $\mu\text{m}$  in thickness. As a result, the BSI wafer is attached to a thin carrier wafer, introducing stretching and warping, and ultimately causing problems as CFA and micro-lenses 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. Fig. 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 possible future product depicted in Fig. 7.8 as 28 megapixel, 0.9  $\mu\text{m}$  pixel, 1/2.5" format. However, today's popular 1/2.3" format and 0.9  $\mu\text{m}$  pixels corresponds to 24 megapixel sensor and 1.12  $\mu\text{m}$  pixels corresponds to 16.4 megapixel sensor. 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 LGs, 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 [Hyncek \(1979\)](#) and the pinned photodiode by [Teranishi et al. \(1982\)](#) used in CCDs for low

Image sensor area diagrams  
(pixel and total die)

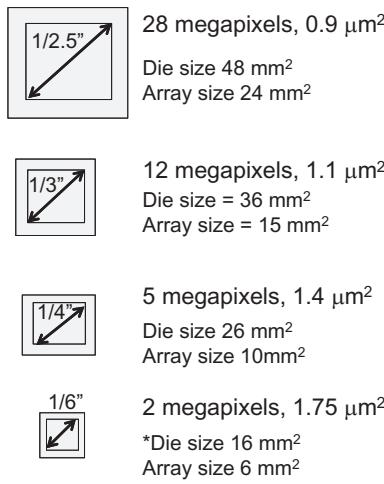


Fig. 7.8 Mobile image sensor array size, die size, and resolution.

Table 7.1 Relation of pixel size and resolution to various optical formats (note, other optical formats are in use in the market as well).

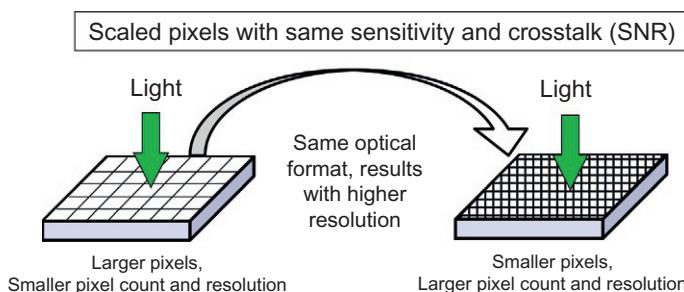
Number of pixels (resolution)	Optical format					
	1/2.5"	1/3"	1/4"	1/8"	1/6"	1/12"
1.3 Mpix				2.2 $\mu\text{m}^2$	2.2 $\mu\text{m}^2$	1.4 $\mu\text{m}^2$
2 Mpix				1.75 $\mu\text{m}^2$	1.75 $\mu\text{m}^2$	1.1 $\mu\text{m}^2$
3 Mpix				1.4 $\mu\text{m}^2$	1.4 $\mu\text{m}^2$	0.9 $\mu\text{m}^2$
5 Mpix	2.2 $\mu\text{m}^2$	1.75 $\mu\text{m}^2$	1.4 $\mu\text{m}^2$	1.1 $\mu\text{m}^2$		
8 Mpix	1.75 $\mu\text{m}^2$	1.4 $\mu\text{m}^2$	1.1 $\mu\text{m}^2$	0.9 $\mu\text{m}^2$		
12 Mpix	1.4 $\mu\text{m}^2$	1.1 $\mu\text{m}^2$				
18 Mpix	1.1 $\mu\text{m}^2$	0.9 $\mu\text{m}^2$				
28 Mpix	0.9 $\mu\text{m}^2$					
Pixel size ( $\mu\text{m}^2$ )						

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 of the active pixel sensor (APS) technology by [Lee et al. \(1995\)](#), combining a pinned photodiode

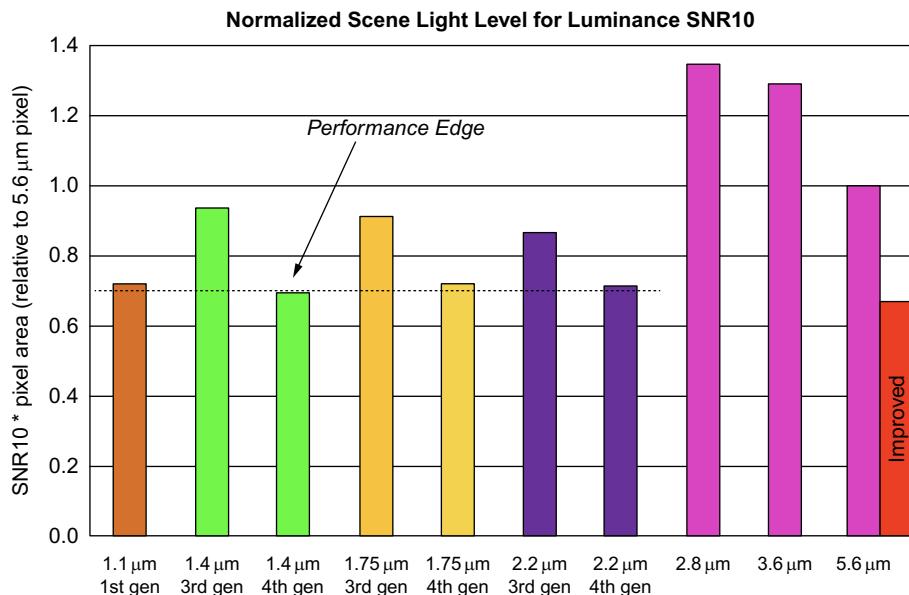
and a transfer gate achieving nondestructive readout and CDS leading to sufficient signal-to-noise ratios (SNRs) for good performance image capture with underlying CMOS process technology. The APS sensor permitted charge transfer to an in-pixel amplifier using CDS, with integrated circuits (ICs) for fixed-pattern noise removal (Guidash et al., 1997; 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  $\frac{1}{3}$ " and  $\frac{1}{4}$ " **optical** formats, sensor manufacturers must produce sensors that achieve a “scaled” performance. Fig. 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 of an inch square, the sensor always achieves the same SNR. Fig. 7.10 shows an example of the evolution of pixel SNR-10 values overtime. Essentially, the largest pixel 5.6  $\mu\text{m}$  was developed first, establishing the relative performance target of 1.0, then the next smallest, and so on. The graph shows the smaller pixels, which were developed later, such as the 1.4  $\mu\text{m}$  fourth-generation pixels, perform better relative to the initial 5.6  $\mu\text{m}$  pixel. Specifically, less luminance per area or scene illumination is needed to achieve the same SNR. As well, the improvements from the development of the smaller pixel are usually reapplied to the larger pixels to create subsequent generations of each pixel, overtime (note the 5.6  $\mu\text{m}$  pixel to the far right in the graph is comparable to the later 1.1  $\mu\text{m}$  pixels). Fig. 7.11 shows that for a particular optical format ( $\frac{1}{3}$ "), the SNR-10 improves overtime as solutions emerge to the challenge of scaled performance with ever smaller pixels.

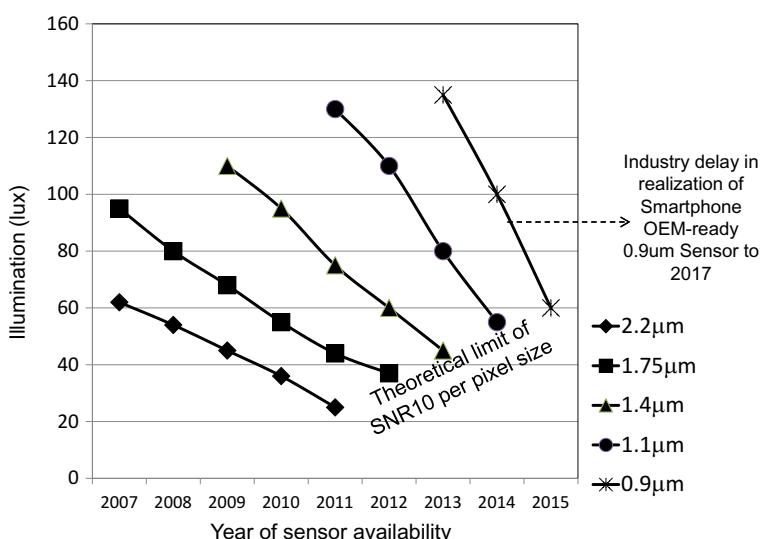
The industry challenge remains achieving the next pixel reduction in a timely manner. Historically, pixel advances were achieved every 18 months, typical of Moore’s Law of CMOS transistor density doubling every 18 months. This Moore’s Law of CMOS imaging created a well predictable industry, in terms of performance and product delivery. However, the trend has slowed recently as the limitation of CMOS and optics increase the challenges. “More-than-Moore” technologies have been developed



**Fig. 7.9** Preserving SNR performance with an arbitrary area as pixels shrink.



**Fig. 7.10** Example scaled pixel SNR performance per pixel generation.



**Fig. 7.11** Industry SNR-10 performance trends for a  $\frac{1}{3}$ " sensor.

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 recombinations possible. At least 4000 electron capacity is needed for mobile. Mobile sensor designers strive to achieve between 6000 and 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 4 T pixel, to as low as 1.5 transistors per pixel.

Shading or nonuniformity 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°). Attention to the symmetry of the pixel and the tuning of the optical pathway either by use of special materials or optical LGs 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**

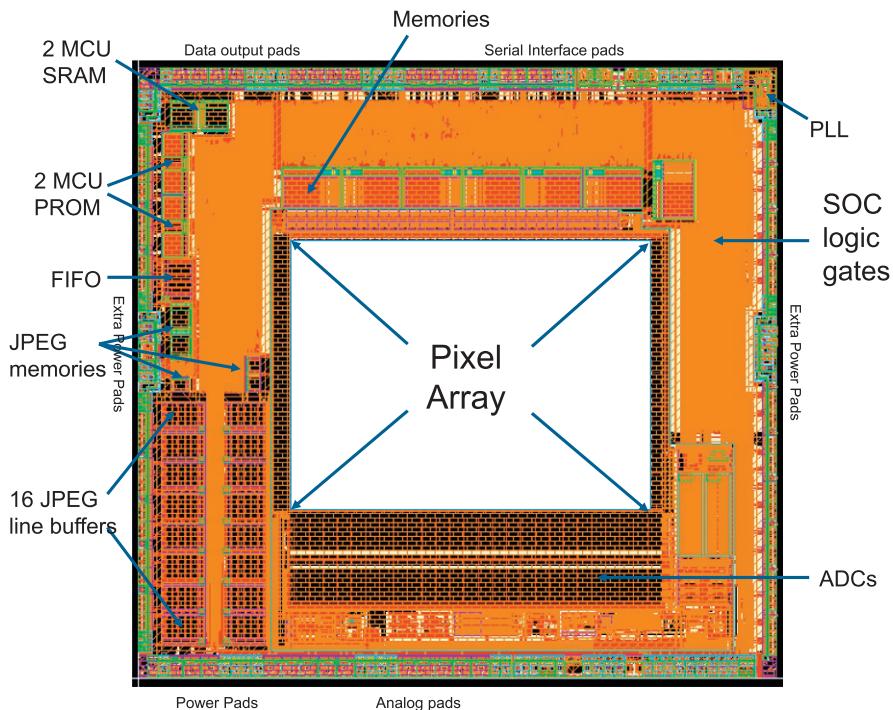
Two primary functions of the CMOS sensor are readout of the pixel array and conversion of the analog pixel signals to an output digital data stream (sometimes compressed). The image sensor has limitations of speed, power, and size. Typical CMOS sensors for mobile consume between 500 and 900 mW. The choice 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 frames per second (fps).

Mobile sensor ADCs are usually designed with 12–14 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 fps (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/s per lane. The MIPI alliance serial interface can realize up to 1 Gbits/s per lane, with 1.5 Gbits/s in development. In practice, a mobile sensor may use multiple lanes for multiples of those data rates. However, as sensors increased from 12 to 22 megapixels and at 30 fps or even 60 fps, serial data rates and number of lanes increased dramatically.

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](#). [Fig. 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



**Fig. 7.12** Die layer geometry showing physical layout for a typical CMOS image sensor SOC.

(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 filtering. In other cases, a full camera image signal processor (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 subsampled 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 1/30th a second exposure interval. A camera using ISO 400 at the same 1/30th of a second exposure time would perform better, with much less visual noise. Larger, more sensitive pixels, such as those with a 1/2.5" sensor would create a higher ISO camera than a 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 reflow, 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 5% 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 1/4", 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 AF 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 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 HDR in isolation from the sensor, such as with the mobile phone’s 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. [Fig. 7.13](#) illustrates the key technologies developed or in development to create superior image and video capture experiences in camera phones. This section will include an update of these technologies, from pixels to cameras. In addition, [Fig. 7.14](#) also shows some important sensor-embedded technologies that we will discuss in this section. HDR, global shutter (GS), 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 Mobile silicon imaging: Pixel advances

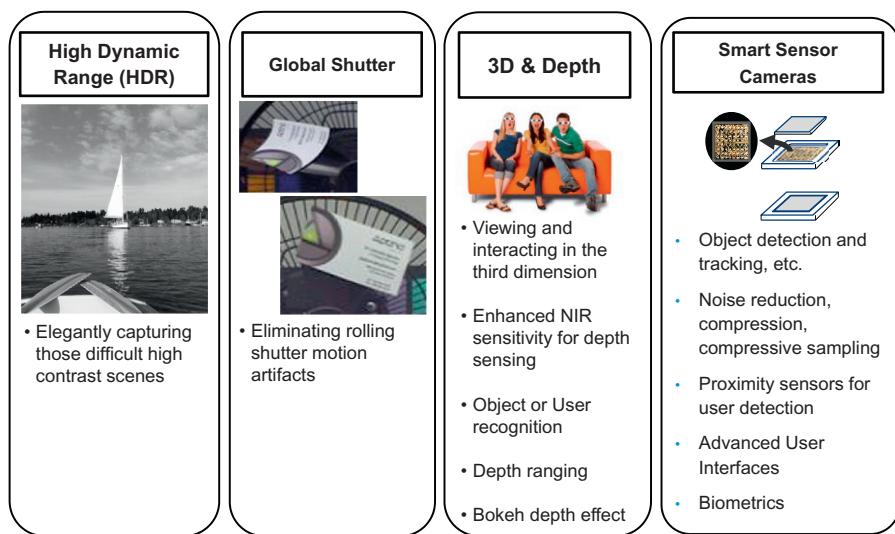
#### 7.3.1.1 Smaller pixels, but not too small

The smallest pixel in production for smartphones now is the 0.9- $\mu\text{m}$  pitch pixel. However, the best performing smartphones usually don’t feature such a small pixel in the RF, scene-photographic camera (RF) due to poorer performance of that small of a pixel. Rather a new trend is to promote even higher resolution and smaller pixels for the selfie camera (FF), due to its popularity. In one case, the Vivo V7+ smartphone



Quality images in all lighting conditions	Fast, responsive camera operation	Small devices to capture everywhere	Seamless video capture	Power reduction
<ul style="list-style-type: none"> <li>Low light sensitivity enhancements</li> <li>HDR capabilities</li> <li>Advanced image processing</li> <li>Image stabilization for longer exposure</li> <li>Dual camera</li> </ul>	<ul style="list-style-type: none"> <li>Context mode switching</li> <li>Fast PDAF algorithms</li> <li>Reduced startup latency</li> <li>User presence and tracking</li> </ul>	<ul style="list-style-type: none"> <li>Pixel shrink for smaller optical formats</li> <li>WLM technologies</li> <li>High CRA tolerance</li> </ul>	<ul style="list-style-type: none"> <li>HD &amp; UDTV video modes</li> <li>Image stabilization</li> <li>Fast interface</li> </ul>	<ul style="list-style-type: none"> <li>Design and process</li> <li>"Always-on"</li> <li>User detection</li> <li>Power down modes</li> <li>Ambient Light Sensing (ALS)</li> </ul>

**Fig. 7.13** Key technologies to achieve superior image and video capture experiences.



**Fig. 7.14** Sensor-embedded technologies.

uses a 24 megapixel selfie FF camera with 0.9  $\mu\text{m}$  pixels, coupled with a lower-resolution scene RF camera at 16 megapixel and bigger pixels—a reversal from earlier smartphones. The technology required to provide good performance at these small dimensions has recently emerged from sensor suppliers, primarily enabled by 3D

stacking [both Samsung's Tetracell (for Vivo) and OmniVision's PureCel-S sensors]. Another enabling technology found within these image sensors is a four-cell binned pixel design that uses a smaller pixel for higher resolution, yet also adds a feature to bin each four adjacent pixels to gain sensitivity for low-lit situations (with lower resolution). This binning approach is ideal for supporting low-lit indoor selfie capture, or selectively, higher-resolution outdoor shots in situations where more light are usually available.

Key enabling technology for 0.9  $\mu\text{m}$  pixels includes the use of deep-trench-isolation (DTI) and shallow-trench-isolation (STI) semiconductor processes to provide isolation between the pixels and thereby less crosstalk between pixels, especially with higher aspect ratio pixels needed for small area pixels. With the aspect ratio of the pixel extending as the pitch reduces, developers use advanced integrated optics with unique micro-lens filters and aperture grid metal isolating the color light paths. In the Vivo case, using a Samsung sensor, the pixel depth is 3.8  $\mu\text{m}$  for a 0.9- $\mu\text{m}$  pitch pixel, defining a "4:1" aspect ratio. Small pixels have extended the substrate depth to achieve acceptable performance.

Counter to the consumer digital camera and DSLR markets where ever smaller pixels and higher resolutions continue to emerge in the market, the trend for makers of top smartphones rather shows a *reduction* of sensor resolution with bigger, more sensitive pixels to achieve improved sensitivity and faster integration times, and new features in the pixel. This strategy includes increasing the camera's optical format (or active area) and improving the lens system to bring more light into the sensing array (decreasing the F-number). Today, the average resolution of the many top-rated world-facing cameras in smartphones is 12 megapixels with pixels between 1.25 and 1.55  $\mu\text{m}$ , and with a 1/2.3" optical format. Certainly, there are exceptions to this trend with others driving even smaller pixels at higher resolutions. On the other hand, selfie cameras have recently improved to feature much higher resolutions such as 12–24 megapixels and small pixels.

The fabrication requirements of achieving high-performing small CMOS pixels, coupled with new unique capabilities, are driving most sensor manufacturers to now use 65  $\mu\text{m}$ , 55  $\mu\text{m}$ , or 40 nm CMOS process nodes for manufacturing of mobile image sensors, which carries additional cost. This is primarily driven by optical and electrical "steerage or control" of photons and then electrons within the sensor and isolation of the photo-generated electrons to retain them within pixels that are ever smaller and taller. Smaller pixels enable design of a higher resolution or lower-cost sensor. The taller pixels enable improvement of spectral response, such as boosting near-infrared (NIR) performance, or special features. Finally, some of the motivation for a shift to denser process nodes for image sensors is due to the availability of certain wafer-level processing equipment, like bonding or other steps, sometimes only available on 12" CMOS fabrication lines with these higher density minimum dimensions. However, as 3D wafer stacking becomes more prevalent, we expect to see somewhat of a reversal in node shrinking as the pixel layer moves to larger process nodes for cost reasons. Some are choosing to use the advantages of 3D stacking though to drive to smaller pixels as well.

The previously mentioned requirement for NIR sensing has emerged for depth imaging applications and in particular the use of “invisible” LED or VCSEL laser illumination so that the smartphone can “see” or detect the user’s presence or identity without annoying flashes of visible light that would distract the user. These invisible flashes of light are used with a depth sensor to measure the distance and shape of objects in the scene. This information is then used to control the behavior of the smartphone. Certain frequencies of NIR (like 940 nm) also carry great advantage to function well in sunlight, which is greatly diminished at that frequency enabling less susceptibility to the bright sun swamping the illumination event. As a result, much semiconductor process development is being deployed to increase the QE of sensors at 940 nm.

### ***7.3.1.2 Special phase detection focus pixels (PDAF)***

While CMOS pixel performance continues to improve with better design and manufacturing processes, new capabilities like in-pixel focus methods have emerged in smartphone solutions. Quickly determining lens focus parameters with special pixels in the CMOS sensor was first designed for consumer-market, mirror-less digital cameras. It enables fast and accurate focus across the image vs old techniques of repetitive cycling of moving the lens, then measuring focus by computing image contrast, then moving the lens again, repeating until the best focus is found. This capability was initially only placed at a few specific locations in the array, but now has extended to all pixels in the array, in some cases. With an in-pixel, phase-detection autofocus (PDAF) approach, in each frame time the focus can be determined continuously and accurately for every pixel ([Śliwiński and Wachel, 2013](#); [Fontaine, 2017](#)).

With PDAF, essentially each phase detection pixel sees a slightly different location, shifted by half a pixel, which is similar to human eyesight parallax. PDAF also better detects fast moving objects moving into or away from the camera. It not only indicates which direction to move the lens, but also it estimates the distance for fast AF. The downside is that phase detect has difficulty in low illumination situations. New methods from Samsung use dual pixels AF system that eliminates the use of auxiliary laser-assisted AF with every pixel performing AF vs other approaches with a similar percentage of select pixels (perhaps 5%) in the array having phase-detection capabilities. Another method involves using computer vision and multiple exposures to estimate depth. Hybrid phase and contrast methods from Sony have enabled video rate 30 ms focus times.

### ***7.3.1.3 Special pixels distance-for-Bokeh photograph blurring***

A photographic effect called Bokeh, where the image naturally blurs in off-focus regions, results from the lens aberrations and aperture shape usually found in large lenses DSLRs. Several methods to emulate this Bokeh depth-effect portrait mode are found in smartphone world cameras today. They detect the distance between subjects in a scene and control a blurring by computing a new image as a combination or modification of focused and defocused images. Some algorithms blur selective

portions of the image usually in different depth planes from the subject. The traditional method of measuring contrast differences in the image and “searching” by moving the lens in and out can be used for a “slow” implementation of Bokeh. A second, much faster method is derived from the use of dual cameras of differing focal lengths or apertures to measure distance in the image. A third Bokeh AF method results when using a sensor with these dual aperture phase detection pixels across the array (PDAF), in which case the Bokeh feature is fast and accurate.

In summary, the marketing of smartphones has changed from *resolution-is-king* or *how many pixels one has* to other measures of imaging value like color rendition, ability to focus, infrared performance, and low-light capability. Conclusively, a balance of factors creates the best user experience (UX) with camera phones. Currently found in some of the best performing camera phones use a Sony IMX378 image sensor with only 12 megapixels, but in a larger 1/2.3" optical format for higher sensitivity. The pixel is 1.55  $\mu\text{m}$  with PDAF, with a total of  $4056 \times 3040$  pixels. As well, the IMX378 utilizes BSI and stacked technology to create a high-performance solution. On the other hand, the marketing of selfie cameras has changed to stress resolution advances.

### 7.3.2 **Mobile silicon imaging: Camera advances**

#### 7.3.2.1 *Dual world camera*

Many smartphones today add a second scene-capture RF camera beside the traditional scene-capture RF camera to enable new features such as zoom, depth measurements, resolution, or low-light sensitivity. The side-by-side placed cameras can vary by focal length with each using a different field-of-view, varied resolution, varied color spectrum (perhaps one camera senses RGB light, while the second senses monochrome light), or other, such as aperture in unique cases. Some smartphone manufacturers feature a smooth blending electronically switching between the two images from two different focal-length lenses achieving a continuous zoom function. In some cases, stereopsis methods are used to compute depth from the two images, which can then be used to establish focus or depth effects.

Dual cameras are designed as either *nonsymmetric*, with one low-resolution and one high-resolution sensor, or *symmetric*, with dual sensors of equal resolution, in configuration. In addition to computation to resolve depth, dual cameras require calibration to dynamically correct for any mechanical misalignment due to mechanical shock of the smartphone, which prevents blurring of the blended images. Sensor fusion algorithms combine imagery from both sensors to create zoom effects or to enhance image quality with combination of color RGB and monochrome sensors. HDR can be implemented with simultaneous multiple exposures or combination of color and monochrome images. Depth computation can be classified into sparse or dense depth mapping. Sparse is necessary for fast applications like AF and object tracking, while dense is necessary for Bokeh or segmentation. This can be implemented in the form of special application-specific integrated circuits (ASICs) or the smartphone AP. The ASIC implementations can operate with lower latency.

### 7.3.2.2 *Performance assessment with “DxO mobile mark”*

To assist consumers and manufacturers assess the relative performance of personal mobile device cameras in the market, DxO ([www.dxomark.com](http://www.dxomark.com)) provides a service to capture and score the performance of all major mobile phones using a database of over 1500 images and 2 hours of videos in lab and outdoor situations. Parameters of noise, resolution or detail, color rendering or color depth, demosaicing artifacts, dynamic range, sharpness, AF, depth effect, low-light ISO, video stabilization, and other parameters are measured for comparison.

As of 2017, the highest DxO Sensor mark for a smartphone camera is 98 for the Google Pixel 2, whereas the highest score for a consumer digital camera is 102 for the Hasselblad X1D-50c. The Hasselblad uses a huge sensor with 50 megapixels at  $5.3 \mu\text{m}$  per pixel. Certainly, the use of high-end, interchangeable lens solutions moves consumer cameras like mirror-less and DSLR cameras into different categories of performance than small, fixed-purpose lenses found in most smartphones. However, Google’s advantage lies in computational imaging rather than lens differences—note most other smartphones today use dual cameras to achieve what Google does with one. Other sections describe the algorithmic methods used to achieve this performance advantage.

## 7.3.3 *Mobile silicon imaging: Semiconductor processes advances*

### 7.3.3.1 *3D stacking methods and attributes*

3D stacking manufacturing technology, where multiple wafers (and associated die) of differing design and process are stacked together with high-interconnect density, enables the development of particularly unique imaging solutions for mobile applications. Mostly 3D stacking enables improved camera performance, size, potential power consumption and tightly coupled computation for higher speed or more fruitful algorithms. Areas improved can be noise, dynamic range, speed, and functionality.

A three-level 3D stacking sensor approach, where three silicon wafers or layers are bonded and interconnected together (the pixel, processor, and memory layers), allows for new types of mobile imaging products. The benefits of a tight coupling of sensing and computation enable advanced imaging capabilities not possible without stacking.

While perhaps even more significant than BSI, 3D stacking enables designers to boost performance, functionality, or integration. Three important categories enabled by 3D stacking are: (1) new generations of pixel arrays, (2) new methods for digitizing, I/O, and image computation, and (3) new camera capabilities or features.

#### *3D stacking category 1: New generations of pixel arrays*

- (1) (*Unique split pixel architectures*) Once wafers can be interconnected at sufficient density to enable a connection at the pixel pitch or group-of-pixels pitch, then inherent processing advantages can be realized to enable a larger, reduced-noise pixel “active” area for a given pixel pitch. Moving some of the transistors or digital processing from the pixel layer to a different digital layer can enable a higher “fill-factor” enabling more photo-generated

- electrons to collect in the pixel. Many combinations are possible to extend the sensitivity, dynamic range, or resolution per area.
- (2) (*Increased pixel-pixel connectivity*) Each pixel can utilize more logic transistors, more analog transistors, and more pixel-to-pixel connections via a tight interconnection to the pixel circuits on adjacent layers.
  - (3) (*Smaller pixels*) Each pixel can shrink with stacking as the buried electrodes of the photodetector and signal processing can be located on the second layer. However, the interconnect method directly limits the ability to benefit from this for very small pixels.
  - (4) (*Split process layers*) With two, three, or more layers bonded together via 3D stacking, each layer's semiconductor process can be optimized for performance or more efficiently achieve some function. For example, in a non-stacked CMOS and especially BSI sensors, CMOS pixel layers can have constrained metal, other materials, thermal steps, or other doping profiles to minimize generation of dark current or other parameters, whereas a process optimized for logic density is not necessarily optimized for dark current. Stacking isolates the process difference to different wafer layers. Stacking then enables enhanced performance by minimizing process constraints.
  - (5) (*Spectral enhancements*) With unique processes per layer possible, QE can be enhanced for different spectrum and hyper-spectral solutions can result.

### *3D stacking category 2: New methods for digitizing, I/O, and image computation*

- (1) (*Unique Pixel Access Architectures*) With 3D stacking, pixels can be more-easily *locally* processed in analog, such as binning, multiple storage nodes or multiple sense transistors. This partitioning can be tuned to match any application's needs.
- (2) (*ADC and I/O*) The complexity and variation of the ADC can vary greatly with stacking. Readout and I/O structures can be more flexible to the application.
- (3) (*8K/33 megapixel sensors at 240 fps*) 3D Stacking has been utilized to optimize the ADC into a middle layer for high-speed and low-noise readout necessary for 8K image sensors (Kawahito, 2016). Note, the top layer has an array of 1.1  $\mu\text{m}$  pixels, and the bottom layer has the logic/readout layer. These methods may be applied to mobile sensors for exceptional video and photographic performance as well in the future.
- (4) (*Processing and computation*) With more *local* computational area available to designers when using 3D stacking methods, sensor designers can integrate functions like enhanced noise-reducing filters, feature recognition and optical flow, and even convolutional neural network (CNN) processing into a single sensor chip. Data can be classified and complete decisions formed by this additional computation embedded in a 3D stacked sensor.

### *3D stacking category 3: New camera capabilities or features*

- (1) (*Small camera size*) Integrated camera solutions with processing, memory and sensing dies combined form the smallest x-y area solution possible, with the processor and memory underneath the sensor with only 100s of microns in height added rather than the addition of a separate packaged die, which becomes compelling for camera phone OEMs. For the camera phone, this approach can enable simpler placement of the camera with fewer constraints of physical size of the camera module.
- (2) (*Super high-speed sensors*) With more signal processing circuits per pixel or group of pixels enabled by adding more circuits onto other stacked layers, architectures can be devised to achieve extremely high speeds. Sony's use of memory as a third layer enables 960 fps readout for super slow-motion playback, and predictive capture modes so key photographic shots are less likely to be missed (Haruta et al., 2017).

These advantages should drive wide-scale adoption of 3D stacking for mobile image sensors. Several manufacturing processing methods have been introduced and are in use to achieve 3D stacking IC fabrication, as follows:

- (1) *TSV*: Initial approaches to 3D stacking used TSVs and oxide-oxide bonding for chip-to-chip interconnect, mostly with two-layer bonding. The stacking density limits uses to only row or column driver interconnection to the second layer. Note, TSV is proven, in volume, and used to manufacture BSI with interconnect in the periphery and direct oxide bond elsewhere. Moving the sensor's logic to a die under the sensor reduces the total die area vs standard CMOS sensors and BSI.
- (2) *Hybrid bonding*: Newer approaches to 3D stacking use hybrid oxide and metal interface bonding across the die to achieve more flexible interconnect spacings (or pitch). With hybrid bonding, stacking can be utilized both inside the pixel array and outside the array in the periphery. Improvements in gate-oxide noise levels and light guiding has been implemented with hybrid bonding. With thicker silicon, deeper and narrower backside DTI, SNR-10 for a 1.0-μm pixel has improved from 90 to 80 lux ([Venezia et al., 2017](#)). This as well enables even further die size reduction or performance enhancement. This hybrid bonding has been demonstrated by TSMC/OmniVision with 3.7 μm pitch interconnects and bond size at 1.8 μm diameter.
- (3) *Three-level stacking*: Xperi's Direct Bond Interconnect (DBI) creates a bump-less bonding technology and was implemented by Sony adding copper-to-copper bonding and their Exmor-RS BSI process for up to three-layer bonding. Others have followed recently.

The first three-layer stacked CMOS image sensor die in volume production, the IMX400, was designed into Sony's Xperia XZs phone. The device has the following three layers (with metal layers shown in parenthesis: Al = aluminum, Cu = copper, W = tungsten).

- (top layer) 90 nm BSI imaging array (1 Al, 5 Cu),
- (mid layer) 30 nm 1 Gbit DRAM and row drivers for the imaging array (3 Al, 1 W), and
- (bottom layer) 40-nm ISP logic and the ADCs (6 Cu, 1 Al).

This IMX400 chip has an optical format of 1/2.3" at 19.3 megapixels. Super slow motion is enabled by this stacking architecture, with 960 fps storage of video into an integrated DRAM layer ([Haruta et al., 2017](#)).

Sony has also used stacking to achieve a 1-μm pixel pitch. They introduced a back-illuminated CIS (BI-CIS) with hybrid bonding of two substrates with Cu—Cu metal bonding and an interlayer dielectric (ILD) oxide bonding. The interconnect pitch between the CIS and ISP layers is 3 μm with over 3 M interconnects. The device is a stacked BI-CIS with 22.5 megapixels 1/2.6" with 1 μm pixels and an ISP ([Kagawa et al., 2017](#)).

Also to enhance small pixel performance, OmniVision and TSMC features a three-layer stack based on 55-nm logic wafers enabling the design of a high-performance 24 megapixels sensor with a compact 1/2.5" optical format and 0.9 μm pixels—PureCel—S ([Venezia et al., 2017](#)).

The challenge for broad-scale adoption of 3D stacking includes cost, yield, and manufacturing capacity. Note, with three wafers stacked, the impact of yield is multiplicative (yield of die on layer-1, times yield of die on layer-2, times yield of die on layer-3), which can greatly impact final cost of the solution. Manufacturing costs vary from \$100 to \$300 per wafer before yield consideration,

with TSV having the highest cost and lowest density ( $\sim 10 \mu\text{m}$  interconnect limit vs  $\sim 2 \mu\text{m}$  for DBI).

### 7.3.4 Mobile silicon imaging: 3D camera advances

#### 7.3.4.1 Machine perception and 3D depth imaging

The “Next Generation of Imaging” for personal mobile devices will center on usage of 3D depth imaging for a wide variety of exciting new UXs and features. Similar to Microsoft’s X-box Kinect solution with embedded depth imaging capabilities, mobile devices will increasingly include 3D depth sensing devices. Uses include advanced user interfaces, biometrics, ecommerce, and AR applications. While photographic and video quality will remain a differentiator in product performance, using information sensed by embedded 3D depth sensors will further enhance the media capture quality, it will as well enable new experiences for these personal mobile devices. These new experiences will include:

- (1) *Intelligent photography*: With depth information augmenting the capture of a two-dimensional (2D) image of a scene or subject, computational imaging can create improvements to the photo or video in the form of better focus, better contrast or lighting controls, and with other intelligent photography methods.
- (2) *Augmented Reality*: AR applications where the synthetic objects are rendered to appear at specific locations, casting the perception that the object is actually located at that 3D position in space.
- (3) *Biometrics*: Biometrics for user authentication and ecommerce such as Apple’s iPhone-X Face ID which uses structured light methods to measure the shape of the human face and time-of-flight (ToF) to determine user presence.
- (4) *Avatars*: Creation of simple animojis for chatting or interactive games.
- (5) *e-Commerce*: Mapping candidate items for purchase digitally to the user’s shape so accurate fit and style choices can be made during online shopping.
- (6) *Advanced user interfaces*: Computer vision will also enable other unique *user interfaces*, such as augmenting touch with optical gesture controls. With depth information and low-power computer vision, the mobile device will be able to perform *persistent identification* of the user, simplifying the UX so that authentication is performed less often.
- (7) *Holoportation*: Creation of 3D holograms of objects that will be rendered by AR/virtual reality (VR) applications enabling “holoportation.” A special (mobile) camera with 3D capture can be used to place a person or object in physical space anywhere in the world as a hologram.

Some competing technologies to enable this compact, low-power, low-cost 3D sensing, and machine vision computation following. Key requirements are to minimize depth ambiguity, motion blur, ambient light interaction, and multipath issues.

- *Time-of-Flight (ToF) solutions*: feature lock-in pixels where the ToF sensor and a phase modulated VCSEL illuminator create a complete 3D depth image at video rates. ToF pixels are special versions of an NIR pixel with dual storage nodes, and some special timing. Some implementations of ToF sensors consumed higher power and were susceptible to high levels of background illumination (the reflected signal from the illuminator became overwhelmed by sunlight when operating outdoors). However, these early problems have been overcome.

- *Structured light (SL) stereopsis solutions:* use a sensor where NIR GS pixel images a reflected scene illuminated by a fixed pattern VCSEL. SL has high computational complexity and added latency. Implementations have difficulty with edge disparity, such as a person's hairline.
- *Stereo vision solutions:* use two sensors with GS NIR pixels and heavy computation of the depth disparity. Implementations can have difficulty with latency and in applications that require depth without texture in the scene, which is prevalent in indoor applications.
- *Unique methods and hybrid solutions:* Samsung has introduced a MEMS, laser-scanned RGBZ sensor ([Wang et al., 2017](#)), enabling design of RGBZ solutions with accurate time-stamped solutions of 2D RGB images and 3D IR depth images. In addition, Bellus3D ([www.bellus3d.com](#)) has created combined stereopsis with color and NIR cameras, and structured light for higher performance, especially for their 360° face scanning application.

#### **7.3.4.2 Improved NIR Sensitivity for 3D cameras**

The key technology to enable personal mobile 3D imaging for machine awareness is NIR sensitivity, especially at 940 nm wavelength (and beyond). Both invisible operation to humans and a gap in the sunlight spectrum provides unique advantages for these applications. Beyond standard CMOS image sensor processes, the following new semiconductor process methods will enable a boost of QE at 940 nm in a range of 30%–50%, necessary for mobile 3D sensing products. As well, some benefit can be realized in the visible spectrum from these methods.

- (1) *Silicon-on-Insulator (SOI) Layer:* Today's highest-wafer-volume NIR sensor in production, used in Apple's iPhoneX FaceID solution, boosts the front-side imager's NIR sensitivity using an SOI layer, DTI, and with a thick 6.1 μm substrate. The added SOI layer coupled with deep trenches (filled with dielectric) at the pixel boundary enable collection of more longer-wavelength photons recombined deeper in the pixel—enhancing the QE for NIR. Enhanced QE results by trapping light above the buried oxide (BOX) and reducing crosstalk between pixels. With the BOX, metal contaminants are less likely to migrate into the pixels ([Yoshimi et al., 2008](#)). In contrast, most image sensors today are manufactured using back-thinned, epi-doped silicon substrates, where the thickness of the epi-doping determines the operating spectrum.
- (2) *Nyxel Pixel:* This new pixel architecture uses thick substrates and DTI to achieve up to 5 × boost QE to 40% for 940 nm for a 2.8-μm pixel. Key to improved QE sensitivity, the surface scatters the light in the pixel to lengthen the path it travels, thereby increasing the number of electrons captured from longer wavelength photons ([Wilson, 2017](#)).
- (3) *Pyramid surface light-diffraction structures:* New process flows have been developed to forego the large capital investment of high-energy ion implanters to create deep NIR photodiodes. The pyramid-surface process creates 400 nm dimensioned pyramid structures at the pixel surface and DTI along the pixel boundaries. Similar to the Nyxel, longer light pathways are critical. Initial results show 30% QE at 850 nm. Sony demonstrated a 1.2-μm pixel and SmartSens a 2-μm pixel with this type of structure.
- (4) *Quantum dot photodetector (0.7–2 μm):* Quantum dots are implemented as a unique monolithic layer or film of pixel material enabling a lower cost and greatly extended IR performance. The QE at a specific frequency can be “tuned” by design of specific QD materials. Utilizing 130 nm process, 40% QE was achieved at 940 nm. The advantage of QD remains the ability to extend sensitivity to longer wavelengths, beyond 1.1 μm for even greater advantage in sunlit situations. Crosstalk between the pixels can be one of the limitations of current quantum dot solutions ([Malinowski et al., 2017](#)—InVisage, [Barrow et al., 2017](#)—IMEC).

- (5) *Hybrid CCD-on-a-CMOS*: Fully depleted, back-illuminated CCD pixels have been used for some time to extend the sensitivity to NIR. A newer approach forms a hybrid of these fully depleted CCD pixels (Holland, 2006) on a CMOS wafer to extend the QE to 70% at 905-nm NIR wavelength (Popp et al., 2013, Espros, TSMC). This approach can as well be used to create multispectral sensors, perhaps a new trend in mobile sensors.

### 7.3.5 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 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% gain in dynamic range results, which leads to infrequent use of this capture mode.

#### 7.3.5.1 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 12–14 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.

### 7.3.5.2 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 HDRs. 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 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.

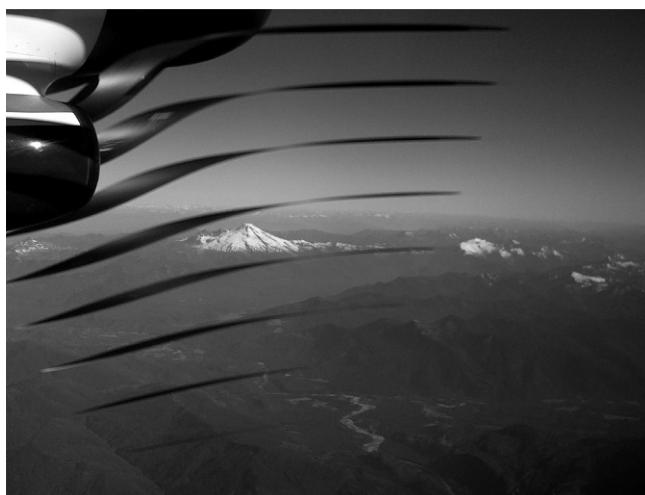
Applications which demand very large dynamic range can be addressed with sensors using more and larger variation in exposure control. For example, Aptina’s three-row multi-exposure sensor offers extended dynamic ranges with two knee points to approximate an exponential capture transfer curve. The sensor 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.

### 7.3.6 *Limitations of rolling shutter and GS 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.



**Fig. 7.15** Captured photograph using CMOS rolling shutter sensor showing interesting distortion of actual image captured of a rotating aircraft propeller.

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 integrates 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 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 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 GS 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 interline transfer methods solve the problem, but at a severe area cost (between 50% and 100% of the integrated sensing area).

The CMOS GS 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, GS reduces the sensitivity by about 6%–10% 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 GS pixels could include using a faster and more sensitive sensor. Mobile CMOS image sensors began at 10–15 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 GS until stacking emerges to enable more transistors per pixel.

### **7.3.7 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 Kinect-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. From Google/Lenovo’s project Tango to iPhone-X’s FaceID, 3D cameras embedded in smartphones have entered the market and is expected to increase rapidly in volume.

**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; projector complexity
Depth of focus	Capture multiple images at narrow field of view with varying focus, then extract depth	Accuracy; motion artifacts
Time-of-flight	(1) 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  (2) Phase modulation type: measure phase delay per pixel	(1) Power necessary to generate the pulses; sensitivity of pixels required to capture “pulsed” light drives either large, costly die size or low resolution  (2) Sensitivity and storage of phases limits pixel shrinkage, but advantage in sunlight, simpler illumination, camera size, etc.

### 7.3.7.1 Proximity sensors

Smartphones use proximity sensors to determine the nearness and attention of a user. In general, proximity devices prevent useless power consumption and accidental touch actuations when one holds the phone to the ear. Higher levels of proximity detection also activate features like biometrics and change the user interface if the owner is present and has turned his/her attention to usage of the mobile device.

Recently proximity sensors have moved from *simple* ambient illumination detection coupled with measurement of a variety of other changes in the device (like position changes via GPS, gyroscope movement, accelerometer, etc.), to recently in smartphones, a new fully automatic assessment of the number and location of people around the phone's display. Some ambient light sensors detect the eye's reaction to the display illumination with a photodiode and NIR filter. Many smartphones today use ToF NIR sensors to directly measure the existence and location of the user(s). Single-photon avalanche diode (SPAD) architectures have been selected for these ambient sensors from ST Microelectronics.

## 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, 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
- metadata creation.

Metadata 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 RF and FF sensors

While 99% of mobile phones manufactured today feature at least one camera. Smartphones typically use two sensors, (1) a RF camera, sometimes called a world-facing camera, located at the rear of the phone and facing the distant subjects or scene when the user looks at the device's display and (2) a FF camera, sometimes called a selfie camera, located at the front side of the phone, facing the user and usually situated within an arm's length from the user. 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 RF camera has an image and video performance mandate driving high resolution and larger lenses, while the FF has the challenge of even smaller size since it is secondary and constrained by size to fit beside today's ever growing displays, yet with image and video performance historically limited to enable user identification and tracking or video conferencing. However, now the FF selfie camera application is driving flagship phones to enhance image quality and even include optical stabilization for even better selfies on some flagship phones.

### 7.4.2 RF scene or world 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 12–24 megapixels,
- 100 lux illumination level at SNR of 10 ([Alakarhu, 2007](#)),
- QE greater than 90%,
- pixel full well capacity of 6000 electrons, and
- operation at 15–30 fps.

These performance requirements are for a  $\frac{1}{3}$ " optical format camera module that is typically 8.5 by 8.5 mm in area and 5.5 mm in height. DSC level performance is attained in many flagship smartphone products featuring larger optical formats with 1/2.5", 1/2.3", or 1/1.8" sensors and optics. At the extreme of resolution and size in camera phones, some time ago Nokia Corporation introduced a 41 megapixel camera phone product (containing  $7728 \times 5368$  pixels), called the Pure View 808, where the camera used oversampling algorithms to combine pixels from a very large 1/1.2" optical format sensor to form a very high quality 5 or 8 megapixel 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 selfie 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 ultrasmall 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 WLO, 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 TSV technology for component interconnect from the bottom of the sensor (Han et al., 2010). This simplifies the camera construction as the optical elements only reside on the top of the sensor and the bond pads then reside on the bottom of the sensor. Sometimes even more important, the area of the silicon sensor can reduce if one uses a redistribution interconnect layer and locates the bond pads within the active area of the sensor (under the sensor).

Lenses for WLM were a key 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.

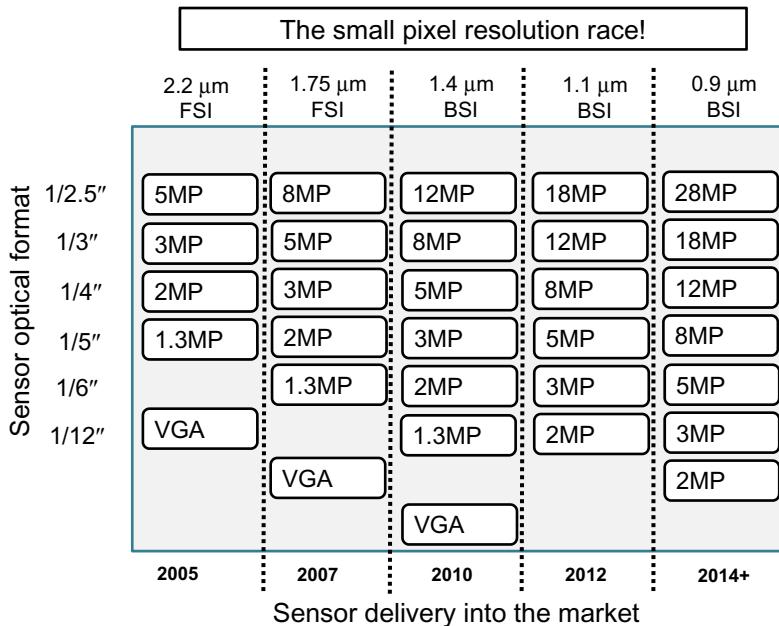
For the camera phone designer, the WLM camera is simple for the integrator at every level. As new applications emerge for the FF camera, image and video quality become more important, such as for video conferencing. However, the limitations of cost and yield for wafer-level methods typically restrict the area to small sizes (1/6"–1/15"). As an example, the wafer-level TSV and optics layers add hundreds of dollars of cost to each combined wafer. A standard 200 mm CMOS image sensor wafer (with 28,000 die candidates per mm<sup>2</sup>) may cost up to \$1000 per wafer. This translates into \$0.57 for a 16 mm<sup>2</sup> 1/6" sensor, yet a much larger cost of \$0.92 for a 26 mm<sup>2</sup> 1/4" sensor, ignoring typical yield decreases for the bigger die. However, when WLM costs are added to the total cost, say up to \$2000, the small die is then \$1.14 and the larger die is \$1.85, with \$0.92 in equivalent packaging costs. As die sizes increase beyond these small die sizes, WLM packaging costs exceed conventional packaging costs.

#### 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, 1.75, and 1.4  $\mu\text{m}$ , 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.

Fig. 7.16 shows each year's product portfolio enabled by new generations of pixels (2.2, 1.75  $\mu\text{m}$ , 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 5 megapixel, 1/3" sensor, produced over 7 years ago, still ships in high volume in nonmobile markets.

In the recent generations of pixel advancements, specifically at the 0.9- $\mu\text{m}$  node, the industry has delayed the generation cycle to employ more advanced semiconductor and optical technologies. This delay is shown in Figs. 7.16 and 7.17. Essentially,

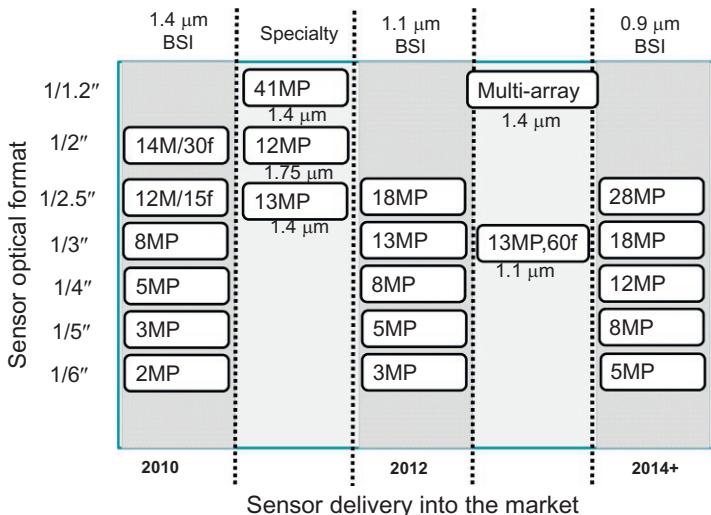


**Fig. 7.16** Mobile sensor product evolution and industry trend.

delay results from the innovation and process investment needed to sustain “scaled” performance beyond conventional CMOS technology. Fig. 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



**Fig. 7.17** Mobile sensor specialty product differentiation grows, as pixel race slows.

or memory within a camera module. Each mobile image sensor pipeline of processing contains:

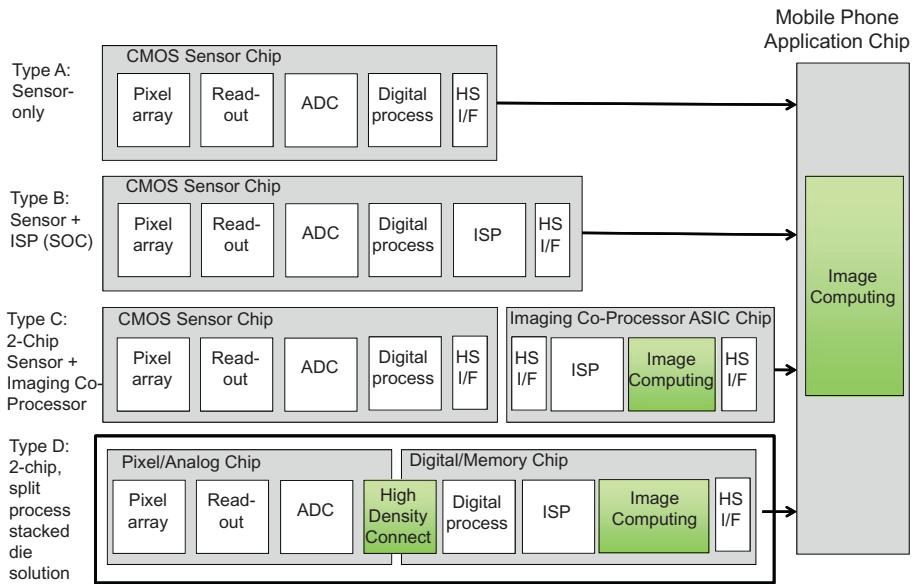
- the pixel array,
- analog processing,
- digital processing, and
- high-speed interface to the end-using system, as follows and shown in [Fig. 7.18](#).

#### 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 system processor (ISP), 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

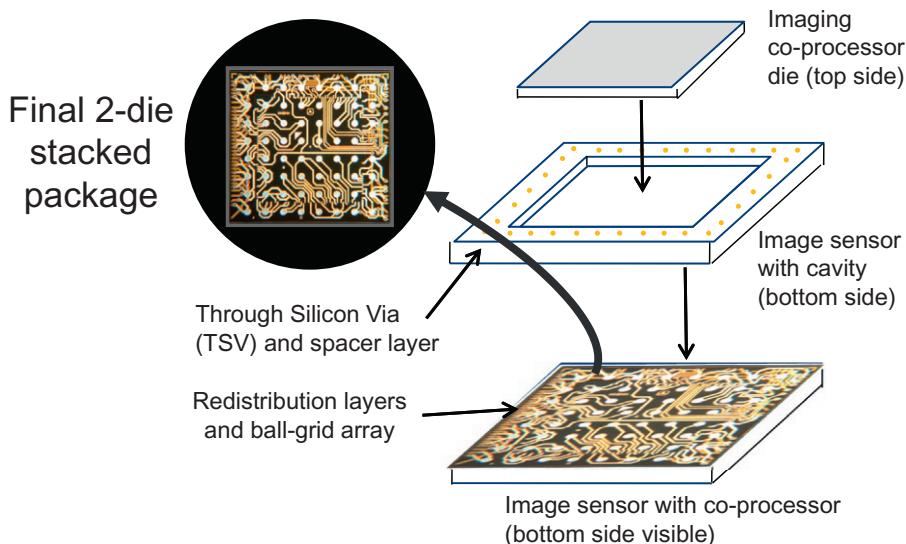


**Fig. 7.18** Four CMOS image sensor solution architectures.

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 500K 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 high-quality ISP, HDR, face tracking, sensor fusion, and computer vision algorithms, all of which can offer competitive advantage to the market. Aptina and others have developed an imaging coprocessor that offers the computational power to perform these functions in a small ASIC chip. The coprocessor 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 with the multicore GFLOPS capability of application processors. As well, the capability of stacking together three separate layers, such as



**Fig. 7.19** Die stacking enables sensor and coprocessor integration, creating a smart camera.

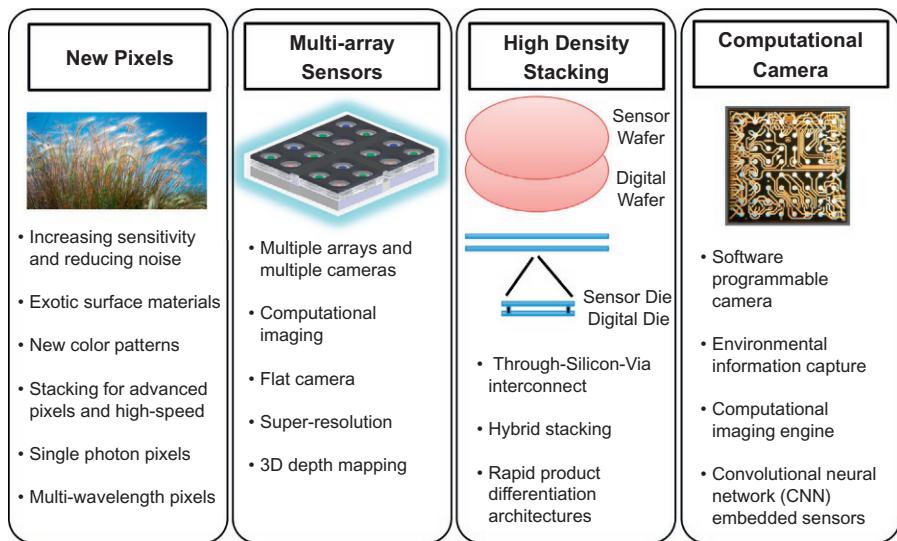
the sensor, processor, and memory wafers/dies may emerge. Ultimately, with die/wafer stacking and power reduction, sensor-embedded image processing for camera phones solution may transition to “frame-based” solutions, supporting full-frame memory-based algorithms for image processing or computer vision.

#### 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 fps, 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 6 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. [Fig. 7.20](#) shows particular categories of innovation where we expect camera phone core technology advancements.



**Fig. 7.20** Emerging mobile imaging technologies.

### 7.5.1 *Pixels: New exotic materials, new color patterns, and new nonplanar, 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). To date, several mobile image sensor manufacturers have been developing such 0.7 μm pixels. Methods in development to achieve smaller pixels include the use of exotic materials for high sensitivity pixels, such as organic films (Ihama et al., 2010) or perhaps quantum films (Greenemeier, 2010). As well, others in the industry developed new color pattern architectures, so-called non-Bayer patterns, for increased sensitivity, by introducing more white light into the sensor with full white (W) pixels in addition to the RGB pixels (Sony Corporation, 2010), as shown in Fig. 7.21. Note that a typical Bayer RGBG pattern diminishes luminance by over 50%.

Finally, 3D wafer stacking methods pixels that leverage vertical structures within the pixels (also called 3D structure pixels), vs today's planar pixels, remain in active development to perhaps enable shrinking the pixels for mobile devices to 0.7-μm pixel pitches. 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 (illustrated Fig. 7.3).

#### 7.5.1.1 *Non-Bayer subsides*

While non-Bayer color filter architectures were found in mobile smartphones a few years ago, they have been replaced again by standard Bayer solutions in most cases. The inclusion of white or clear pixels boosted sensitivity. However, both for computational

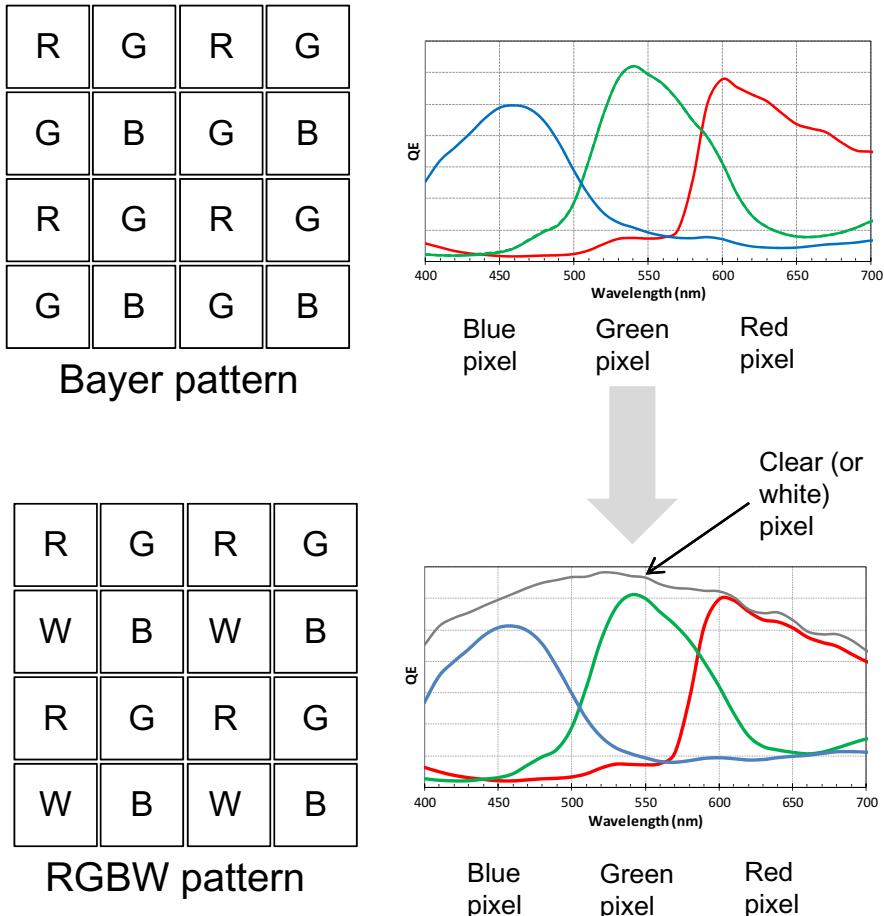


Fig. 7.21 Higher sensitivity color matrix pattern.

complexity or uniqueness reasons, coupled with lower performance at some corner cases, have resulted in standard Bayer solutions regaining prominence. Some exceptions are the TetraCell and 4-Cell approaches for high-resolution selfie FF cameras.

### 7.5.2 New computational imaging multi-array sensors

We expect the camera phone industry to 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, multi-array, coded aperture, diffuser aperture, 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 is 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 refocus, motion blur removal, multispectral measurements, and others. [Table 7.3](#) highlights some of those advantages.

[Fig. 7.22](#) shows the concept of a multi-array sensor and its replacement of a conventional RF, 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

**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

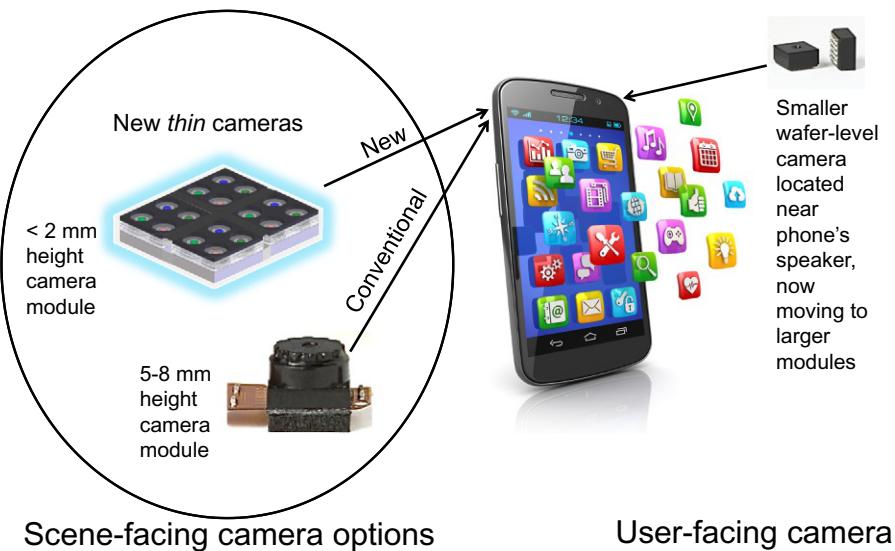


Fig. 7.22 Multi-array sensor and other flat optics for mobile phones.

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 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 this explosive growth of available compute cycles are enablement of new smarter cameras and a broader class of useful applications.

Mostly this computational power enables new methods of imaging for mobile products. Traditional mobile imaging 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. 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 ClearView. 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 multispectral and HDR 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 New mobile imaging applications**

Leveraging the computational model of the camera phone, countless applications will emerge like new user interfaces that leverage face recognition for camera control, imaging and tracking for gesture control, and AR. As well, in some cases,

the smartphone is enhanced with a physical plug-in camera module, such as at the top portion of the device, creates 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.

#### 7.5.4.1 *Smartphone spectrometer sensing for analysis of materials*

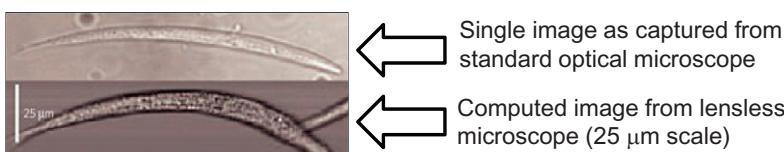
Increasing the measured spectrum of illumination for CMOS imagers and coupling such a sensor with transmissive gratings enables a smartphone spectrometer ([McGonigle et al., 2018](#)). Imaging beyond the human visible spectrum with a smartphone enables interesting new applications to understand the underlying condition of consumables, and others. For example, these new type of sensors can not only image in the visible but also broadly into IR and UV ranges, dedicating specific portions of the sensor to an array of different light wavelengths. This can be used to find objects and detect the underlying materials, helping users understand the underlying chemical similarity of different samples. Food can be sensed to determine if meat, fish, dairy, or other is spoiled or not, or the identity of a sample. These devices will find their way into smartphones in the future.

#### 7.5.4.2 *Smartphone optical fingerprint sensing*

Optical fingerprint sensors have been shown to support OLED infinity or bezel-less displays offering a unique look and feel for the newest smartphones. In this case, the usual capacitive fingerprint sensor normally placed under the home button is replaced with a unique CMOS image sensor located underneath the display that images the user's fingerprint. Synaptics initially introduced this technology with a Vivo smartphone ([Maxham, 2018](#)). The advantage is fast and accurate fingerprint authentication without consuming any space on the front side of a smartphone. We expect this trend to continue.

#### 7.5.4.3 *Smartphone opto-fluidic sensors for biomedical applications*

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 technology, we can magnify particles about 60 times.



**Fig. 7.23** Computed image from sequence of images with an opto-fluidic microscope camera.

## 7.6 Vision for the future of mobile silicon imaging

This subchapter will look into the future for personal mobile imaging devices, including smartphones. This includes advanced imaging in the form of image sensor arrays, new types of image sensors with embedded computation, and new forms of computational imaging. Possible new types of image sensors and computational imaging solutions employing deep learning and other concepts or directions are introduced and suggested for adoption by makers of mobile imaging devices.

Mobile silicon imaging will transition through several *mega-application-phases*, all of which will drive unique embedded silicon to perform image sensing and computation:

(*Application Phase-1) Photography and video*—Most of today's mobile imaging products are designed for the photographic/video application and continue to advance with better pixels, optics, and processing to enhance the captured image quality. However, these sensors are not always ideal for computer or machine vision applications.

(*Application Phase-2) Embedded 3D depth sensing and computational imaging*—In order to more quickly and accurately sense and understand the user and his/her surroundings and enable computer vision applications in a smartphone, today's newest smartphones embed advanced depth imaging methods like dual cameras with stereo processing, 3D structured light (SL) or ToF cameras, coupled with computational imaging. These sensors and embedded computation enable: (1) rapid authentication of the user's identity, (2) multi-person ranging to determine the presence of a user, and (3) unique levels of photographic quality similar to the use of a DSLR. Ultimately new classes of vision applications become practical in the popular smartphone form factor.

(*Application Phase-3) Embedded 3D interactive imaging*—As the performance (depth accuracy), speed (near 100 fps), power consumption (operating for hours), and field-of-view (eventually 360°) of these 3D sensors advance, future smartphones will allow full immersive applications where the user is interactive in many ways including AR/VR with full 3D mapping and localization of the situation around the device using combinations of dual cameras, 3D cameras and beyond to create exceptional new services.

Note these new application phases are additive to the original photography and video phase in that they augment and enhance photography and video in new ways with embedded 3D technology. For example, not only will focus be determined by measurement of subjects and location, the device will collect a deeper level of understanding of the surrounding content and intent of the user to create new experiences, including intelligent photography.

### 7.6.1.1 Computational imaging becomes as important as the image sensor

Much of the photographic experience and performance of new mobile smartphones can be due to *computational imaging*, rather than *silicon imaging* performance. This is true in the case of the highest DxO-Mobile score today, the Google Pixel2. In that

case, machine learning processing is featured for photo exposure and portrait-mode focus. Google likewise implemented machine learning in a new processing chip called the Pixel Visual Core. Their imaging strategy is to use a fast optical sensing system (sensor and lens) so they can capture multiple frames quickly for subsequent computation to result in performance better than one long exposure (Levoy, 2015). Imaging methods developed include: short burst frame imaging followed by accumulation, focal sweep, burst averaging to reduce noise in shadows with preservation of local contrast, and tracking of features in the image to adapt processing to preserve edges (critical at low light).

This fast *burst-exposure imaging*, where multiple exposures are analyzed to form the best photo in low light realizes exceptional low light camera performance. Using these fast frame samples of the same scene, image blur is reduced computationally and the motion of the subjects in the scene is better determined than with one longer frame capture. With a stacked DRAM and processor attached to the sensor, even faster *burst-exposure imaging* followed by a computational imaging model becomes possible.

Traditionally, the best smartphones employ optical image stabilization (OIS) and longer exposures to boost performance in low light where longer exposures would blur the image without OIS. Optically, larger pixels can help stabilize the images vs small pixels, helping to drive the trend to larger pixels and optical formats for smartphones today. The phone's gyroscope can also be sampled and analyzed at 200 times per second to help this stabilization process. Motion sensors, the gyro and the accelerometer are all interconnected for successful performance optimization. Fast *burst-exposure imaging* followed by computational imaging will simplify much of this in the future.

This trend to include computational imaging to improve the imaging experience will continue to include more aspects of imaging. This will include the camera having a better understanding of the photographic situation and adapting appropriately.

#### **7.6.1.2 The smartphone becomes “mobile AI” with edge-based machine perception**

The future of mobile imaging will include Internet-edge-located, machine perception sensors. Many recent applications of machine learning to imaging and machine vision were focused on cloud-based solutions, such as cloud-based processing of photo or video libraries. However, embedded vision or machine perception at the edge, with local vision processing within the mobile device, becomes necessary for the future of mobile image capture. For interactive applications, machine perception must occur at the edge, not the cloud.

Led by the introduction of the iPhoneX’s 3D imaging capabilities, new applications beyond face recognition and authentication will emerge, leveraging user presence determination and situation awareness capabilities of these 3D imaging devices. These emerging applications could include personal safety, better user interfaces, collision-free navigation, and security. The benefits of edge-based machine perception include lower latency, full-time usage (not only when the network is available), new levels of privacy, and lower cost.

To support this edge-based machine perception in the consumer Internet of things (IoT) market, Amazon introduced the DeepLens camera. Featuring embedded deep neural networks (DNN), the DeepLens includes off-the-shelf, preloaded and pre-trained learning to recognize objects like cats, dogs, household objects, motion, and OCR. This will enable a wide range of new applications within a large ecosystem for ecommerce and smart homes, and likely migrate to other mobile devices.

The requirements for such machine perception image sensors include combined NIR and visible light sensing at 100 fps for hand, face, and eye tracking as highlighted in Liu (Liu et al., 2017—Oculus Facebook).

### 7.6.1.3 *Image sensors for smartphone AR/VR viewers*

The future of mobile imaging will include screen-less VR viewers, where a head-mounted frame holds one's smartphone close to one's eyes, to create a mobile VR experience where the users drops into an interactive viewing experience. Watching suspense/action movies, socializing with friends, and navigating through real-life 3D models of cities (with 3D Google maps) become exciting. The market is currently the largest application of VR viewers (Samsung Gear/Oculus). This will continue until low-cost, tightly integrated head-up display (HUD) AR/VR viewers dominate the market in a few years.

Most challenges for this market are relative to display resolution and foveated imaging to detect the user's fixation points to optimize the screen's refresh to achieve the maximum refresh possible at those fixation points. However, sensing the user's gaze, hand position or gestures, and others requires unique sensors that will be integrated directly into the smartphone in the future for this application.

### 7.6.1.4 *New mobile UXs: Unique methods to sense users and context awareness for advanced UXs*

The future of mobile imaging will include emerging image sensor technology such as *nearly* full spherical 360° imaging to advanced perceptive user interfaces with optical solutions. While touch and voice are typical user interface technologies in use today, the future will include increased adoption of 2D and 3D imaging to autonomously determine the intent of the user. The user interface will adapt to the user and environment. From swiping and tapping a display and voice queries, mobile device users will move to use more intuitive and natural interfaces. That will include machine perception of what you are doing and how you could, should or are expected to interact with the device.

These new image sensors will sense and process full surrounding 2D and 3D information, and as well as handle subject presence detection and context awareness. Sensors will detect the user's attention and begin to change its use case or application as a function of the situation. For example, if one wishes to operate the device with gloves on or in wet environments, the user interface will change. Also the skill, identity, and emotion of the users will be assessed continuously. These use cases will be

personalized as well. However, in all cases, new unique image sensors will be necessary to enable these dynamic new use cases.

A key aspect of these new user interfaces or experiences will be the need for “always-on and contextually aware,” driving ultralow-power requirement such as less than 1 mW. Contextual awareness ranges from simple motion detection to person location and others. Since the solution must compute at the edge and within this power requirement, as the mobile device may or may not be awake or connected to the Internet.

To design a practical surrounding environment or *situation* sensing solution for a smartphone or other mobile device, we predict that compressive sampling (CS) methods will be implemented in future mobile sensors (Dadkhah et al., 2013). With CS either the resolution, optical sampling (like gaze or other), electrical readout, or other are minimized to reduce power consumption needed for situation sensing, then those parameters can be selectively reinstated as needed, depending on the situation. Many CS camera approaches have been developed to enable utilization of standard sensors. However, this approach will evolve into integrated solutions for small mobile devices, rather than continually adding more and more special-purpose sensors to the platform.

#### 7.6.1.5 Single-photon CMOS sensors

The future of mobile imaging will also include single-photon sensors. Observing ever shrinking pixel sizes with diminished storage capacities, and increasing computational power becoming available within image sensors, Quanta image sensor (QIS) were introduced with specialized binary jot pixels. A jot pixel image shows the presence or the absence of a photoelectron, composing a binary image. Either high resolution or high frame rate can be realized by appropriate design of a sensor with jot pixels. This QIS sensing architecture may emerge for mobile applications. As well, SPAD CMOS sensors with 3D hybrid bonding and BSI offer good performance at the small size necessary for presence detection. This 3 μm pitch, 1 μm SPAD could enable JOT devices in the future (Ma et al., 2017; Al Abbas et al., 2016; Fossum et al., 2016).

#### 7.6.1.6 Multiwavelength sensors

The future of mobile imaging will likewise include multiwavelength sensors. Demonstrations of new sensors providing multispectral features have been realized. We expect future applications of smartphones utilizing these to detect food spoilage and food identity (to name a few) since the sensor sees a spectrum beyond the eye’s ability. One approach is to use a CCD-in-CMOS architecture to extend the spectral response into different bands (San Segundo Bello et al., 2017—IMEC).

#### 7.6.1.7 Flat cameras and coded aperture with computational imaging

The future of mobile imaging will include lens-less sensors. Mobile device manufacturers desire as flat as possible cameras to have more flexibility for placement of the cameras in a device. Eliminating the lens is a pathway to achieve low-profile

solutions. Lens-less technology has ranged from *coded-aperture* solutions to a recent *diffuser* sensor either with covers or masks over the sensor rather than a lens, each of which are followed by extensive computation to extract the 2D or 3D depth image. Using a conventional 2D RGB image sensor and a machine learning algorithm, the *DiffuserCam* synthesizes a four-dimensional (4D) RGBD light field with color and depth for each ray (Srinivasan et al., 2017). It has been demonstrated to reconstruct 100 million voxels, when using a conventional 1.3-megapixel pixel image with the diffuser layer, but without any special scanned illumination. In comparison, most SL or ToF mobile product solutions today create 30,000–40,000 3D pixels, with maybe 100 voxels per ray (for a 1% depth accuracy), or 3–4 million possible voxels.

With lens-less imaging such as FlatCam, a coded aperture mask is placed closely between the image sensor and the scene (Asif et al., 2015). While a FlatCam solution is less than 0.5 mm thick, the speed of the lens is comparable to a conventional mobile glass lens (F/2.5 vs F/1.8) of even more thickness (10–20 mm). A pinhole is the alternative solution but presents much less light to the sensor ( $\sim$ F/22) for a lower sensitivity.

With DiffuserCam, the lens-less mask varies phase or amplitude and is placed a small distance in front of a sensor (Srinivasan et al., 2017). The transparent mask with coherent light creates high-frequency pseudorandom *caustic patterns* for capture by the sensor. The *caustic patterns* vary with the 3D position of the source with both lateral shift and axial shift for scaling. For a mobile application, one requirement of the DiffuserCam is the use of coherent illumination.

A limitation of most masked-based lens-less solutions is the added latency of many frame times necessary to compute the resulting images. While these approaches will not *replace* high-performance mobile cameras, they will find use as additional cameras in mobile products to enable new advanced interactive or immersive applications.

#### 7.6.1.8 Deep learning smart image sensors

The future of mobile imaging will include deep learning smart image sensors. Deep learning with CNNs methods have been proposed to utilize special angle-dependent pixels and diffraction patterns to reduce the complexity of CNN by implementing part of it optically in a CMOS image sensor (Chen et al., 2016).

## 7.7 Conclusion

Abundant innovation and rich technology advancement over a broad spectrum of areas 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, image computation algorithms, 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, ultrahigh-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.

## References

- Al Abbas, T., Dutton, N., Almer, O., Pellegrini, S., Henrion, Y., Henderson, R., 2016. Backside illuminated SPAD image sensor with 7.83 $\mu$ m pitch in 3D-stacked CMOS technology. In: Proc. IEEE International Electron Devices Meeting, San Francisco.
- 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>.
- 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>.
- Asif, S., Ayremlou, A., Sankaranarayanan, A., Veeraraghavan, A., Baraniuk, R., 2015. FlatCam: thin, bare-sensor cameras using coded aperture and computation. In: Proc. 2015 International Conference on Computer Vision (ICCV 2015).
- Barrow, L., Bock, N., Bouvier, A., Clocchiatti, D., Feng, J., Kollil, N., Pattantyus, A., Sharma, V., Shu, T., Mandelli, E., 2017. A quantum film based quad VGA 1.5 $\mu$ m pixel

- image sensor with over 40% QE at 940 nm for actively illuminated applications. In: Proc. 2017 International Image Sensor Workshop—R53 (IISW 2017).
- 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, pp. 41–44.
- Chen, H., Jayasuriya, S., Yang, J., Stephen, J., Sivaramakrishnan, S., Veeraraghavan, A., Molnar, A., 2016. ASP vision: optically computing the first layer of convolution neural networks using angle sensitive pixels. In: Proc. International Conference on Computer Vision and Pattern Recognition (ICCVPR 2016).
- Cossairt O, 2011. Tradeoffs and Limits in Computational Imaging. PhD thesis, Columbia University.
- Dadkhah, M., Deen, M., Shirani, S., 2013. Compressive sensing image sensors-hardware implementation. Sensors (13), 4961–4978.
- Fontaine, R., 2017. A survey of enabling technologies in successful consumer digital imaging products. In: Proc. 2017 International Image Sensor Workshop—R06.
- Fossum, E., Ma, J., Masoodian, S., Anzagira, L., Zizza, R., 2016. The quanta image sensor: every photon counts. Sensors 16 (8), 1260.
- 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 Vertebrate Retina. Available from: <http://www.vision-research.eu/index.php?id=160>.
- Greenemeier, L., 2010. Light improvements: could quantum dots boost the quality of cell phone pix? Sci. Am. <https://www.scientificamerican.com/article/quantum-dots-cell-camera/>.
- Guidash, R.M., et al., 1997. A 0.6 mm CMOS pinned photodiode color imager technology. In: IEDM Technical Digest, pp. 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.
- Haruta, T., Nakajima, T., Hashizume, J., Nomoto, T., et al., 2017. A 1/2.3-inch 20Mpixel 3-layer stacked CMOS image sensor with DRAM. In: Proc. of ISSCC 2017 (4.6).
- Henrysson A, 2007. Bringing Augmented Reality to Mobile Phones. PhD Thesis No. 1145, Linkopings University.
- Holland, S.E., 2006. Fully depleted charge-coupled devices'. In: Proc. of 2006 SNIC Symposium, Stanford, California.
- Hynecek, J., 1979. Virtual phase CCD technology. In: IEDM Tech. Dig, pp. 611–614.
- Hynecek, J., 1988. A new device architecture suitable for high resolution and high-performance image sensors. IEEE Trans. Electron Devices 35 (5), 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, pp. 14–17. Available from: [http://www.fujifilm.com/about/research/report/055/pdf/index\\_ff\\_rd055\\_all\\_en.pdf](http://www.fujifilm.com/about/research/report/055/pdf/index_ff_rd055_all_en.pdf).
- 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. In: ISSCC Dig. Tech. Papers, pp. 302–303.
- Janesick, J., 2001. Scientific Charge Coupled Devices. SPIE—The International Society for Optical Engineering.

- Kagawa, Y., Fuji, N., Aoyagi, K., Kobayashi, Y., Nishu, S., Takeshita, S., Taura, J., Takahasi, H., Nishimura, Y., Tatani, K., Kawamura, M., 2017. An advanced CuCu hybrid bonding for novel stacked CMOS image sensor. In: Proc. 2017 International Image Sensor Workshop (IISW 2017).
- Kawahito, S., 2016. 3D-stacking architecture for low-noise high-speed image sensors. In: Proc. of 3rd International Workshop on Image Sensor and Systems (IWiSS 2016).
- 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. In: ISSCC Dig. Tech. Papers, pp. 56–57.
- 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. In: IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors.
- Levoy, M., 2015. Extreme imaging using cell phones: see in the dark. In: Proc. of ICCV Extreme Imaging Workshop.
- Liu, C., et al., 2017. Sensors for future VR applications. In: Proc. 2017 International Image Sensor Workshop (IISW 2017).
- Ma, J., Masoodian, S., Wang, T.J., Fossum, E., 2017. Experimental comparison of MOSFET and JFET 1.1 m pitch jots in 1Mjot stacked BSI quanta image sensors. In: Proc. 2017 International Image Sensor Workshop—R18 (IISW 2017).
- Mabuchi, M., et al., 2004. CMOS image sensor using a floating diffusion driving buried photodiode. In: ISSCC Dig. Tech. Papers, pp. 112–113.
- Malinowski, P., Georgitzikis, E., Maes, J., Mamun, M., Enzing, O., Fazzica, F., Van Olmen, J., De Moor, P., Heremans, P., Hens, Z., Cheyns, D., 2017. Monolithic near infrared image sensors enabled by quantum dot photodetector. In: Proc. 2017 International Image Sensor Workshop—R54 (IISW 2017).
- 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. In: Symp. VLSI Circuits Dig. Tech. Papers, pp. 138–139.
- Maxham A, 2018. Hands on With Synaptics ‘Clear ID’ in-Display Fingerprint Sensor—CES 2018. Android Headlines. <https://www.androidheadlines.com/2018/01/hands-synaptics-clear-id-display-fingerprint-sensor-ces-2018.html>.
- McGonigle, A., Wilkes, T., Pering, T., Willmott, J., Cook, J., Mims III, F., Parisi, A., 2018. Smartphone spectrometers. Sensors 18, 223.
- Mori, M., et al., 2004. A  $\frac{1}{4}''$  2M pixel CMOS image sensor with 1.75-transistor/pixel. In: ISSCC Dig. Tech. Papers, pp. 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<sup>TM</sup>. In: IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors, pp. 13–14.
- Nayar, S., 2012. Focal sweep photography. In: Presented at the IEEE International Conference on Computational Photography.
- Ng R, 2006. Digital Light Field Photography. PhD Thesis, Stanford University.
- 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. In: ISSCC Dig. Tech. Paperspp. 500–502.
- Nobukazu, T., 2012. Dark Current and White Blemish. Presented at Image Sensors 2012.

- 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. In: Proc. SPIE, Solid State Sensor Arrays and CCD Cameras. vol. 2654, pp. 63–71.
- Popp, M., De Coi, B., Huber, D., Ferrat, P., Ledergerber, M., 2013. High speed, backside illuminated 1024x1 line imager with charge domain frame store in Espros photonic CMOS™ technology. In: Proc. 2013 International Image Sensor Workshop (IISW 2013).
- San Segundo Bello, D., De Bock, M., Bouleuc, P., Vandebriel, R., Wu, L., Van Olmen, J., Malandruccolo, V., Craninckx, J., Haspeslagh, L., Guerrieri, S., Rosmeulen, M., Borremans, J., 2017. A 7-band CCD-in-CMOS multispectral TDI imager. In: Proc. 2017 International Image Sensor Workshop—P18 (IISW 2017).
- Schuler, C., Hirsch, M., Harmeling, S., Scholkopf, B., 2012. Fast removal of non-uniform camera shake and blind removal of optical aberrations. In: Presented at the IEEE International Conference on Computational Photography.
- Śliwiński, P., Wachel, P., 2013. A simple model for on-sensor phase-detection autofocusing algorithm. *J. Comput. Commun.* 1, 11–17.
- 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](http://www.sony.net/Products/SC-HP/cx_news/vol68/pdf/sideview_vol68.pdf).
- Srinivasan, P., Wang, T., Sreelal, A., Ramamoorthi, R., Ng, R., 2017. Learning to synthesize a 4D RGBD light field from a single image. In: Proc. 2017 International Conference on Computer Vision (ICCV 2017).
- 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. In: ISSCC Dig. Tech. Papers, pp. 108–109.
- Sun, L., Hays, J., 2012. Super-resolution from Internet-scale scene matching. In: Proceedings of the IEEE International Conference on Computational Photography.
- Takahashi, H., et al., 2004. A 3.9 mm pixel pitch VGA format 10b digital image sensor with 1.5-transistor/pixel. In: ISSCC Dig. Tech. Papers, pp. 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). In: ISSCC Dig. Tech. Papers, pp. 96–97.
- Teranishi, N., Kohno, A., Ishihara, Y., Oda, E., Arai, K., 1982. No image lag photodiode structure in the interline CCD image sensor. In: IEDM Tech. Dig, pp. 113–116.
- Venezia, V.C., Shih, C., Yang, W., Zang, Y., Lin, Z., Grant, A., Rhodes, H., 2017. 1.0  $\mu\text{m}$  pixel improvements with hybrid bond stacking technology. In: Proc. 2017 International Image Sensor Workshop, R03 (IISW 2017).
- Wakabayashi, H., et al., 2012. 10.3 Mpix back-side illuminated imager. In: IEEE ISSCC Dig. Tech. Papers.
- Wang, Y., Ovsiannikov, I., You, J., Deane, P., Smits, D., Park, Y., Niesten, M., Hwang, S., Chung, C., 2017. A low-power low-cost high-speed 2D/3D camera for virtual reality headsets, mobile devices and automobiles. In: Proc. 2017 International Image Sensor Workshop—R28 (IISW 2017).
- Wernersson, M., 2012. Next generation camera phones—trends and technologies. In: Presented at Image Sensors 2012.
- 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.

- Wilson, A. (Ed.), 2017. Near Infrared Imaging: Process Improvements Boost NIR Quantum Efficiency. *Vision Systems Design Magazine*.
- Yang, W., Kwon, O.-B., Lee, J.-I., Hwang, G.-T., Lee, S.-J., 1999. An integrated  $800 \times 600$  CMOS imaging system. In: ISSCC Dig. Tech. Papers, pp. 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.
- Yoshimi, M., Delpra, D., Cayrefourcq, I., Celler, G., Mazure, C., Aspar, B., 2008. Current status and possibilities of wafer-bonding-based SOI technology in 45nm or below CMOS LSIs. In: Proc. 2008 International Conference on Solid-State and Integrated-Circuit Technology.
- 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, 1759–1763.

## Further reading

- Antipa, N., Kuo, G., Heckel, R., Mildenhall, B., Bostan, E., Ng, R., Waller, W., 2017. Diffusercam: lensless single-exposure 3D imaging. *Optica* 5 (1), 1–9.
- Forward Concepts, 2012. Cellular Hand Tablet Core Chip Trends 2012. Report No. 1210. Available from: <http://www.fwdconcepts.com/celltab/index.htm>.
- Ku T K, 2011. 3D TSV Stacking IC Technologies, Challenges and Opportunities. Available from: [http://sites.amd.com/us/Documents/TFE2011\\_026ITR.pdf](http://sites.amd.com/us/Documents/TFE2011_026ITR.pdf).
- Takemoto, Y., Kato, H., Aoki, J., Kondo, T., Saito, H., Gomi, Y., Matsuda, S., Tadaki, Y., 2017. Optical characterization of multi-storyed photodiode CMOS image sensor with 3D stacking technology. In: Proc. 2017 International Image Sensor Workshop—P42 (IISW 2017).

# Complementary metal-oxide-semiconductor (CMOS) image sensors for automotive applications<sup>☆</sup>

C. De Locht, H. Van Den Broeck  
Melexis Technologies NV, Tessenderlo, Belgium

## 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:

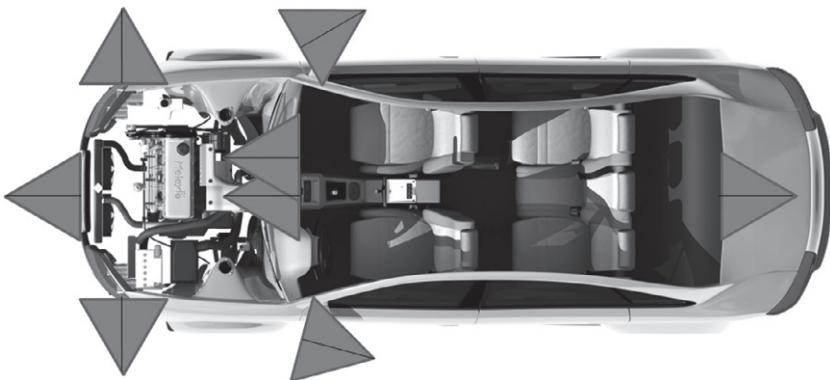
- rear view
- surround view
- night vision and
- road crossing monitoring.

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.

Fig. 8.1 shows some common camera locations for the different applications.

<sup>☆</sup>This chapter is a reprint of the chapter originally published in the first edition of “High Performance Silicon Imaging: Fundamentals and Applications of CMOS and CCD Sensors.”



**Fig. 8.1** Common camera locations for the different applications.

## 8.2 Vision systems

### 8.2.1 Rearview systems

Rearview 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, that is,  $640 \times 480$  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 United States, 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 (FOV) 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 FOV of around 180 degrees. 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 800–1000 µm. The human eye is not sensitive for these wavelengths, so upcoming traffic is not blinded, while CMOS cameras are sensitive in the NIR wavelength wave. With night vision systems with active light, the driver can see approximately three times further ([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](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)).

Low light sensitivity, high dynamic range and extended responsivity in the NIR band



**Fig. 8.2** Example of an automotive night vision image.

are key camera specifications for this application. Automotive night vision systems can also be based on far infrared (FIR) cameras, also called thermal 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 (<http://www.clarion.com/xe/en/tech/technology/camera/index.html>).

## **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.

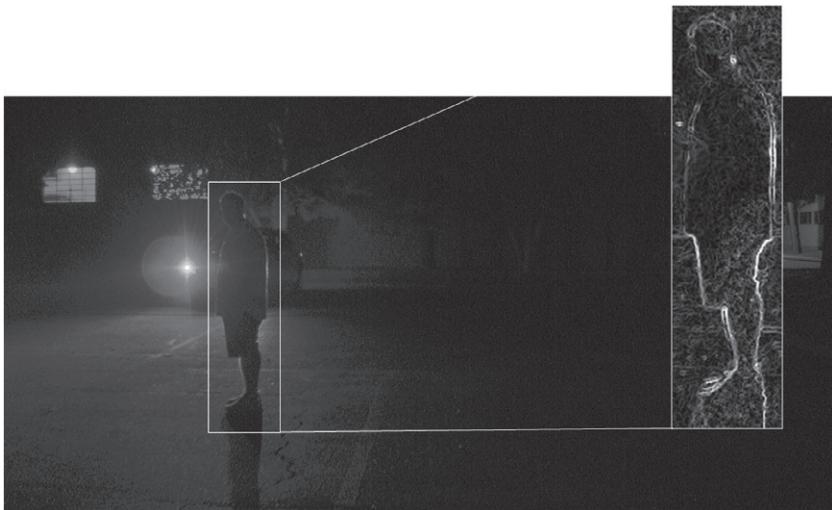
Common comfort and safety ADAS functions include:

- 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 FOV 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 (<http://delphi.com/manufacturers/auto/safety/active/racam/>).

#### **8.3.1.1 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 (<http://www.euroncap.com/home.aspx>) attributes safety points to cars that have city-AEB and urban-AEB. City-AEB looks 6–8 m ahead, detects vehicles and large obstacles to avoid low speed impacts up to 20 km/h to avoid whiplash.



**Fig. 8.3** Example of a pedestrian detection application.

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 (<http://www.clarion.com/xe/en/tech/technology/camera/index.html>). 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 FOV in the range of 40–60 degrees.

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 FOV. This application requires a typical lens FOV between 170 and 190 degrees. 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 BSD, the mounting space for a camera is extremely limited. At the time of writing, common optical formats for automotive cameras range from  $\frac{1}{4}$ " to  $\frac{1}{3}$ ". Combining this optical format with a resolution range between VGA and 1.2 Mpix leads to pixel sizes between 3.7 and 6  $\mu\text{m}$ . A pixel size of 3.7  $\mu\text{m}$  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).

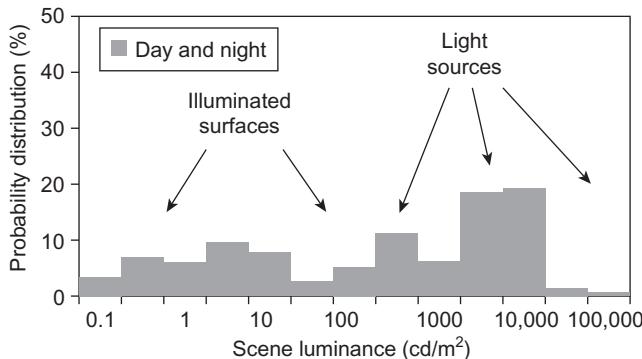
As an objective measure of low light performance, the European Machine Vision Association (EMVA1288) (<http://www.emva.org/cms/index.php>) recommends the use of SNR1 (signal-to-noise ratio) and SNR10. SNR1 provides the light irradiance level so that on the output of the image sensor the signal level 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<sup>2</sup> 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.

### 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. Fig. 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



**Fig. 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.



**Fig. 8.5** Measured dynamic range of a night scene.

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 range in every single frame. Examples of automotive scenes that require a high dynamic range camera include:

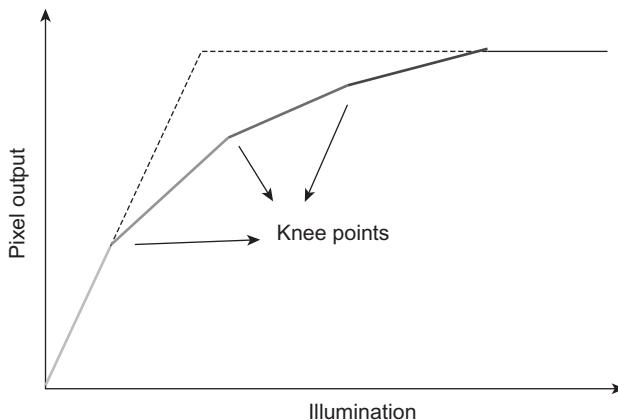
- driving in and out tunnels, parking garages, etc.
- driving towards the sun
- dusk
- driving at night with lights of upcoming traffic shining into the camera lens.

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 light range of the scene. If this condition is not met, the detection algorithms will not be able to function properly for specific gray tone values.

#### 8.4.3.1 *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



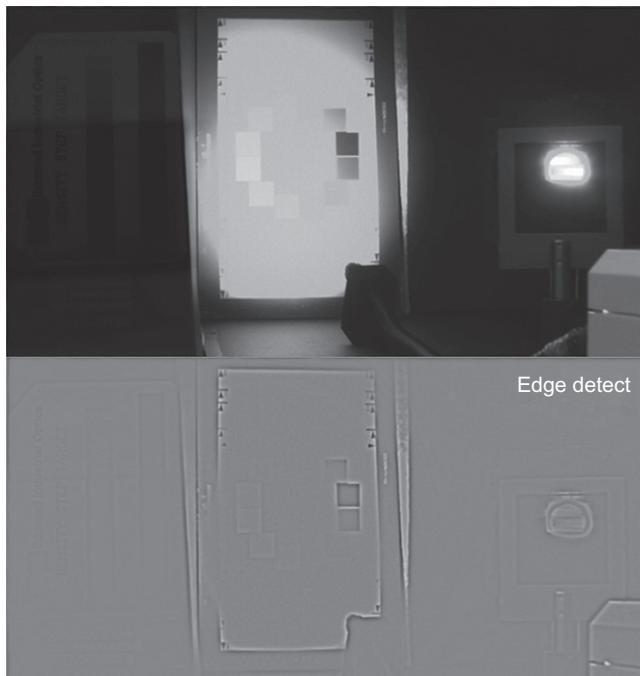
**Fig. 8.6** Example of a piece-wise linear response curve with three kneepoints.

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, that is, 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 (<http://www.melexis.com/Assets/Autobrite-Whitepaper-5815.aspx>).

#### 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  $-40^{\circ}\text{C}$  and as high as  $105^{\circ}\text{C}/115^{\circ}\text{C}$  (<http://www.melexis.com/Optical-Sensors/Optical-Sensing/MLX75411-696.aspx>).

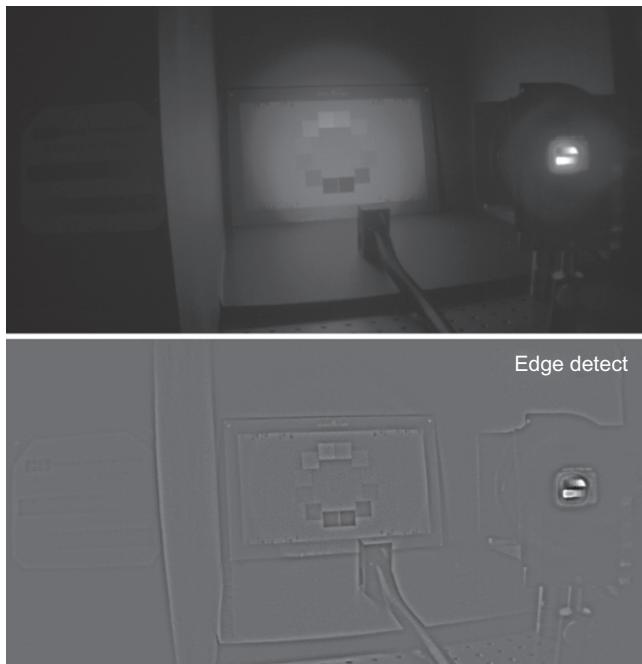


**Fig. 8.7** Picture taken with an HDR imager with two kneepoints: there is information loss in the mid tones when the dynamic range of the scene is larger than 80 dB: the edge detect algorithm does not detect some of the mid tones in the middle of the image.

#### 8.4.5 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 (<http://www.iso.org/iso/catalogue-detail?csnumber=43464>) 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.



**Fig. 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.

#### 8.4.6 Image sensor interface

A common digital interface for automotive image sensors is the low-voltage TTL parallel output interface (e.g., 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, e.g.) 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) (<http://www.melexis.com/Optical-Sensors/Optical-Datacommunications/MLX75605-671.aspx>) 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 (<http://www.open-sig.org/>) 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) (<http://www.ieee802.org/3/RTPGE/>) 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. TSR and 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 NIR light is because NIR light may interfere with color quality. Recent advances in post-processing algorithms enable conserving color quality (e.g., 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. FOV is application dependent and can range from 30 degree for night vision systems to more than 180 degrees for surround and rear view cameras.

Special care needs to be taken for HDR functions. Under HDR conditions, a bright light source (e.g., 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 (<http://www.aecouncil.com/AECDocuments.html>). 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 rearview and blind spot mirrors: ISO16505 “Road vehicles—ergonomic and performance aspects of camera-monitor systems” (<http://www.iso.org/iso/home.html>) 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.

# CMOS and CCD image sensors for space applications

9

P. Jerram<sup>a</sup>, K. Stefanov<sup>b</sup>

<sup>a</sup>Teledyne e2v, Chelmsford, United Kingdom, <sup>b</sup>The Open University, Milton Keynes, United Kingdom

## 9.1 Introduction

Space imaging has a number of unique challenges and typically requires image sensors that are highly optimized. In general it is very costly to launch a camera into space and so the electro-optical performance of the sensor is critical. Because of this CCDs have dominated the space market long after they have ceased to be used for nearly all terrestrial applications. The largest projects currently in development in the world such as Plato (ESA PLATO, n.d.) and Euclid (Endicott et al., 2012) still use CCDs partly due to very high uniformity obtained from a CCD, and partly due to the conservative nature of the space market where heritage is one of the most critical requirements.

As CMOS imaging technology continues to improve this situation is changing and some instruments are now using CMOS technology such as Bepi-Colombo (Flaminii et al., 2010), MTG VisDA (Jerram and Morris, 2016), and JUICE (Janus) (Soman et al., 2016). This is because, for some applications, the advantages of CMOS technology for space now mainly outweigh the remaining disadvantages. The main advantages of CMOS technology are a much simpler interface, low power consumption, low noise even at high frame rates, and critically they are not susceptible to charge transfer efficiency (CTE) degradation from proton irradiation. Backthinned CMOS sensors can now be manufactured with quantum efficiency (QE) comparable to CCDs except at the longest wavelengths, and response uniformity is continually improving. The main remaining advantages of CCDs are still better QE at the longest wavelengths, higher dynamic range, and better uniformity which are particularly critical for space science and hyperspectral applications.

In this chapter we have discussed the particular requirements for imaging from space and how that has driven sensor designs that can be significantly different from those used for terrestrial applications.

## 9.2 Imaging modes in space applications

Space imaging falls into two significantly different categories: instruments that look down at the Earth or at other planets, and those that look out into space. In this section we will review different imaging modes and the types of sensors that are used for those modes.

### 9.2.1 Snapshot or staring mode operation

This is the imaging mode that corresponds to that used for normal terrestrial camera operation, where an image is taken of a stationary or near stationary object with an exposure time that is determined by the light level and by the motion in the scene.

A large majority of astronomical systems use staring arrays because the relative motion of the objects being imaged is low and there is often a requirement to image very faint objects. Typically astronomical imaging systems such as Hubble, Kepler, and Euclid used very large format arrays of sensors that are backthinned to give a maximum sensitivity, with long integration times of up to several minutes to give extremely clear and detailed images (Fig. 9.1). The requirement for very high dynamic range and uniformity means that these missions all currently use CCD technology despite the fact that CTE reduction after radiation means considerable effort is required to correct the images.

Planetary observation can also use staring arrays if the relative motion of the scene is low, which is generally the case if the satellite is at a significant distance from the object to be imaged. Hence Earth observation (EO) satellites that operate from geostationary orbit such as GOCI ([Choi et al., 2012](#)) also operate in this way.



**Fig. 9.1** Hubble image of the Crab Nebula.  
Image courtesy of NASA and STScI.

## 9.2.2 Scanning imagers

There are a number of different types of scanning imaging systems that are used for EO and planetary observation applications when the relative motion of the detector and the scene is high.

### 9.2.2.1 Linear sensors

A linear sensor is essentially the same as the sensor that is used in a photocopier, with the sensor mounted on a satellite moving over the surface of the Earth at a high speed. The sensor typically consists of at least four separate lines with a different color filter over each line. Each line of the sensor is read once for a given line on the ground. Thus for a given required ground sampling distance (GSD) the time to read each line of the sensor is dependent on the satellite's velocity, pixel size, and sampling rate. A linear sensor can typically achieve a GSD of around 2.5 m.

However, the integration time for a linear sensor is very short and reduces as the GSD reduces. The pixel size also will need to be reduced as the GSD is reduced; this means that the amount of light falling in each pixel becomes extremely small for a GSD of <2 m and it is not possible to obtain sufficient signal-to-noise ratio to produce good-quality images at high line rates.

### 9.2.2.2 Time delay integration imaging

For systems requiring the highest resolution multiple lines must be added together using time delay integration (TDI) operation to improve the signal-to-noise ratio to a point where good quality images can be obtained. The preferred means to achieve TDI operation is to move the signal in electrons down the image sensor at a rate that corresponds to the motion of the satellite over the Earth. This mode of operation is referred to as charge domain TDI. Operating in this way means that the signal from many lines can be added together with no additional noise. The alternative is to read-out an entire array for every line imaged on the ground and then add the lines together off-chip. This is usually called digital TDI. Digital TDI has two main disadvantages: firstly the data rate increases with the number of rows to be read, and secondly the read noise increases as the square root of the number of rows to be read and so is only practicable for a small number of rows.

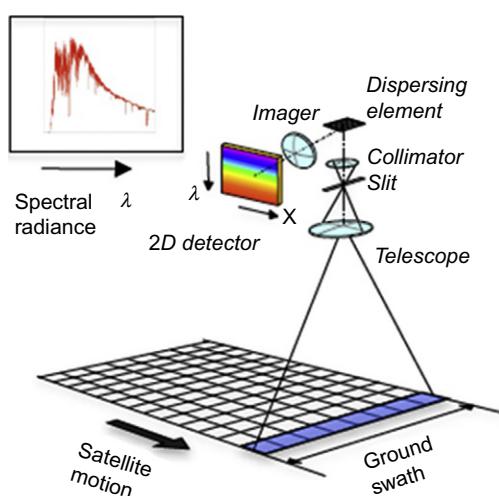
### 9.2.2.3 Hyperspectral imaging

In hyperspectral systems a single line on the ground is imaged as for a linear imager but the spectrum from the line is dispersed across an area imager as shown in Fig. 9.2. This produces an image that contains spatial information across the imager in the row direction and spectral information in the column direction. Because the entire sensor must be readout once per line on the ground the GSD is large, typically hundreds of meters, however, extremely good quality spectral information is produced. Hyperspectral sensors are used for scientific EO applications such as pollution, agriculture, and ocean monitoring.

### 9.3 Important additional requirements for image sensors in space

Space is a hostile environment and electronic instruments onboard spacecraft are subjected to damaging radiation, temperature extremes, and also shock and vibration during launch. Image sensors are particularly vulnerable because they have to look out to observe their object and cannot be provided with the same level of protection as systems deep inside a satellite. The most important requirements for space-grade image sensors, in addition to those asked of the typical image sensor, derive from their challenging operating conditions and are listed below.

- Radiation hardness—image sensors operating in space are not protected by the Earth’s atmosphere, and encounter high-energy protons, electrons, and galactic cosmic rays (GCRs). Satellites employ shielding as a first line of defense against radiation, but many of these particles are so energetic that it is impractical to shield against them. This radiation bombardment causes both gradual deterioration of sensor performance and instantaneous single event effects (SEE). The long-term effects are caused by ionization damage in the insulators (typically oxides and nitrides) and their interfaces to the active silicon, and also by displacement damage (DD) to the silicon crystal lattice. Both can lead to an undesirable increase in the dark current, noise, and charge transfer inefficiency (CTI). Instantaneous effects include single event upsets (SEU) which can corrupt memory elements in CMOS sensors or interfere with the normal operation of their logic circuitry. While these can be considered “soft” transient errors, heavier particles and energetic protons can cause single event latchups (SEL) in CMOS circuits, which can damage an image sensor permanently. The radiation environment in certain places in the solar system, for example, around Jupiter, can be much worse than that in low Earth orbit (LEO) and require particularly high radiation tolerance. Qualifying the radiation hardness of an image sensor is usually a time-consuming job and requires great level of expertise in device physics and operation.



**Fig. 9.2** Hyperspectral imaging system schematic.

- High detection efficiency—due to the expense of lifting optics into orbit satellite designers are aiming to reduce mass and therefore the size of the optical systems, and as a result the optical performance is rarely as good as that in terrestrial systems. An image sensor with excellent detection efficiency offers a way to compensate the shortcomings of the optics, being usually a small, low mass part of the system. This is why image sensors for space applications frequently push the limits of technology for QE, and often feature backside illumination, thick sensitive silicon, and advanced antireflection coatings.
- Power dissipation—power onboard spacecraft is rarely abundant and all instruments are under strict power constraints. Image sensors for space are normally designed to have low power dissipation so that they do not place undue burden on the spacecraft.
- System considerations—an image sensor does not operate independently but is supported by optics, readout electronics, power supplies, radiation shielding, thermal management, data-processing software, and communications. The complexity of these support systems could easily outweigh the complexity of most image sensors. Features that can relax the requirements for the support systems and lead to reduced mass and cost are highly desirable, for example, on-chip signal digitization and processing, wide operating temperature range, and low power dissipation. The CMOS image sensors (CISs) can integrate many system elements on chip, such as ADCs and image processing.
- Packaging—many sensors have to deal with environmental extremes including temperature cycling, shock, and vibration, which places stringent requirements on the construction of the package and the materials used in it. Similar coefficients of thermal expansion to silicon and low outgassing are particularly important.
- Reliability—space qualification is a time-consuming and expensive procedure that could take years, especially regarding radiation damage effects. Stringent control of manufacturing methods, materials, and test procedures are required.
- Long-term availability—many space missions, especially in science, take decades from conception till launch. The timescale of commercial technology development is usually shorter, and this creates concerns regarding the long-term availability of the preferred manufacturing processes (particularly for CMOS due to the evolving consumer market), and of semiconductor vendors remaining in the business.
- Heritage—the space industry is very conservative and risk-averse due to the huge costs of building and launching spacecraft, and the impracticalities of repairs and replacements in orbit. Image sensors and cameras with problem-free flight heritage are highly regarded and often preferred to sometimes more capable alternatives, which however have not been flight-proven. The technology readiness level (TRL) ([ISO 16290, 2013](#)) is widely used in the industry to inform satellite designers about the heritage and the level of qualification of a sensor.

## 9.4 Performance of CMOS and CCD image sensors for space

### 9.4.1 Electro-optical performance

For most space applications electro-optical performance is critical. The most important parameters are typically QE, noise, dynamic range (DR), dark signal (or leakage current), uniformity, and performance repeatability.

#### 9.4.1.1 Quantum efficiency

In general nearly every photon that falls on a sensor should produce a signal that can be detected. This means that the sensor noise must be minimized and critically the QE must be as close to 100% as possible. In a normal CCD or CMOS array there is structure in front of the photosensitive area which reflects some of the light and reduces the detection efficiency. In the case of a CCD this mainly consists of polysilicon electrodes whereas for a CMOS array there are up to six layers of metal on the top of the array. This can partly be overcome by micro-lenses which can produce detection efficiencies nearly 80% at visible wavelengths by focusing the light onto the photo diode within the pixel, however, backthinning can offer very high QE from deep ultra-violet (UV) to near-infrared (NIR) wavelengths.

For backthinning a sensor is attached to a silicon support wafer and then the back of the active wafer is removed so that illumination can be applied directly to the active epitaxial layer. Critical aspects of the backthinning process are as follows:

*Surface passivation:* The back surface of the silicon must be passivated to avoid electrons recombining and hence losing signal charge, and to avoid the generation of high dark signal. There are several different techniques available for surface passivation, but the one that has been used for the large majority of space projects is boron implant followed by an anneal with a short duration, high-power laser pulse to melt the surface layer as this gives very good stability. For high-volume terrestrial application atomic layer deposition (ALD) can be used which is a low-cost process, but relies on charging within the ALD layer and therefore has concerns over the long-term stability.

*Antireflection (AR) coating:* A bare silicon surface will reflect approximately 36% of the visible light incident on the back surface. To optimize the QE, an AR coating is required, matched to the wavelengths of interest. For wide-band detection a multilayer AR coating is used. Examples of the QE achievable with typical AR coatings are shown in Fig. 9.3. For the detection of deep UV wavelengths below 100 nm and down into the soft X-ray range the AR-coating material will either not be effective or will absorb the incoming signal and so no coating is used.

*Active silicon thickness:* Typically space imagers need to image more than the visible light spectrum. For effective detection at longer wavelengths the active silicon must be as thick as possible. Fig. 9.4 shows the impact of using thicker silicon on the longer wavelength QE. For operation at the maximum thicknesses a voltage must be applied to the back surface of the sensor to give a field throughout the silicon to ensure that charge is effectively gathered in the charge collection region at the front of the device. This is referred to as HiRho (from high resistance) technology. For CCDs HiRho technology is used when the active silicon thickness is required to be  $>40\text{ }\mu\text{m}$ . For CISs the very much lower voltages mean that HiRho technology is required when the silicon thickness is greater than around  $10\text{ }\mu\text{m}$ .

Silicon detectors are also used for soft X-ray detections missions such as XMM, Chandra, and ATHENA. Silicon also becomes more transparent for X-rays as the energy increases, so for X-ray detection missions of higher energies even thicker silicon is needed.

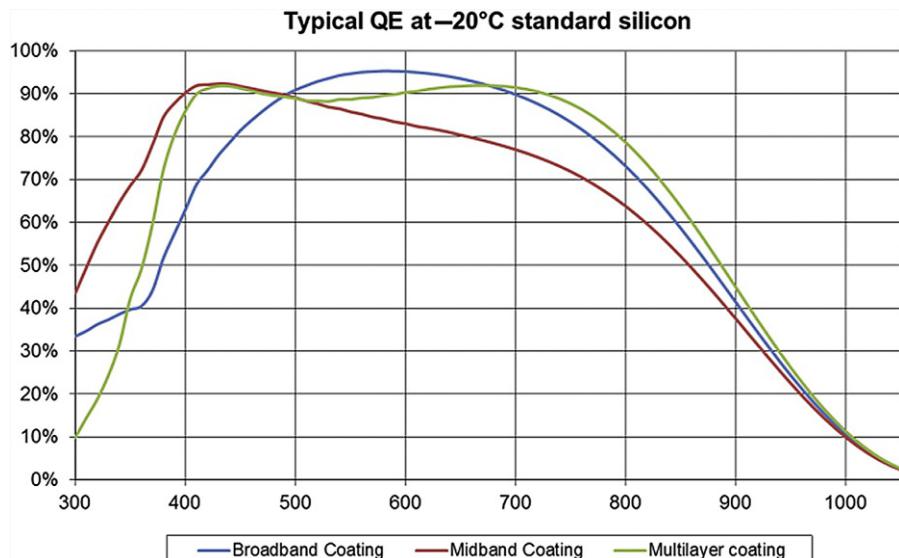


Fig. 9.3 Typical backthinned sensor QE at  $-100^{\circ}\text{C}$  with different AR coatings.

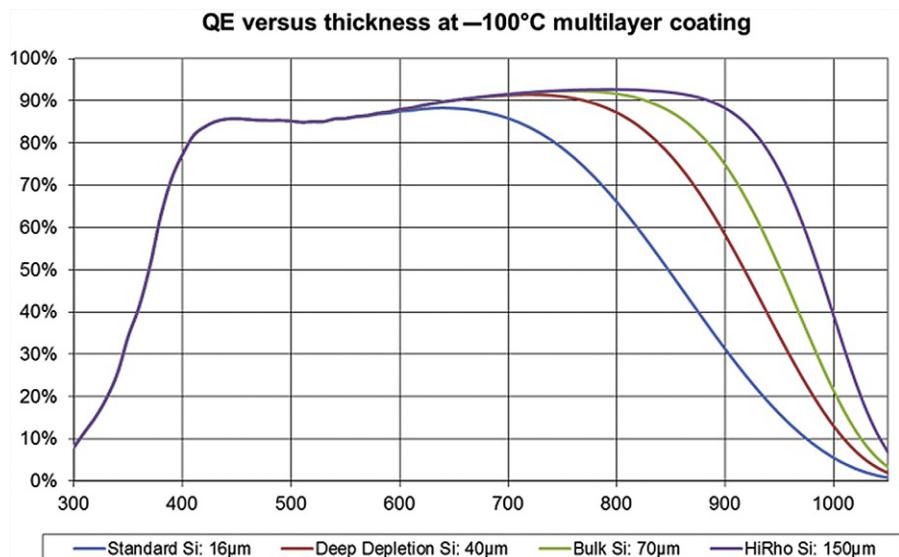


Fig. 9.4 Impact of silicon thickness on NIR QE.

#### 9.4.1.2 Noise

Astronomical applications typically require the detection of low signal levels and hence the noise floor must be as low as possible, typically in the range 2–10 electrons depending on the mission. Although CCDs can only achieve this at low speeds as they are readout through a single amplifier or a small number of amplifiers this is generally not a problem for astronomy where the frame rates are low. The CMOS sensors can achieve noise levels close to one electron even at high frame rates due to the typical column-parallel readout. Higher speed applications will typically operate at higher peak signals where the signal-to-noise ratio is shot noise limited rather than dependent on read noise, and hence the readout noise is less critical.

#### 9.4.1.3 Dynamic range

Dynamic range is the ratio of the noise floor to the peak signal that can be usefully read from a sensor. If a sensor saturates then image data is lost, so sensors must have as high a dynamic range as possible. This is particularly true for astronomical applications, but the signal range can also be very high for some EO applications particularly hyperspectral imaging.

Note that dynamic range is different than signal-to-noise ratio which is the ratio of the signal to the noise at the given signal level. For high signal applications this will generally be dominated by the shot noise on the signal.

#### 9.4.1.4 Dark signal

Sensors are generally operated so that the dark signal generated during integration does not contribute significantly to the total noise. As dark signal is thermally generated with a rate of change exponentially dependent on the activation energy of the source, detectors are frequently operated at a reduced temperature. This means that EO detectors are typically operated in the range  $-40^{\circ}\text{C}$  to  $+20^{\circ}\text{C}$  whereas astronomical detectors which operate at small signals and often need long integrations are operated in the range  $-120^{\circ}\text{C}$  to  $-60^{\circ}\text{C}$ .

The CMOS devices have significantly lower dark signal at the start of life and so it might be assumed that they will be able to operate at higher temperatures, however cooling is still required to a similar level to remove the dark signal increase generated by the space radiation, and especially dark signal spikes from proton radiation (see [Section 9.4.2](#) for details).

#### 9.4.1.5 Measurement repeatability

Precise, long-term and stable calibration is also critical to many space applications. For example, small differences in response from a hyperspectral instrument are used to determine pollution or different types of agriculture. The instrument can only be calibrated on the ground and so there must be no changes in orbit, for example, to detection efficiency or linearity. The main concern is generally the impact of radiation on detector performance. Hence all sensors for space applications will undergo a

qualification or a lot acceptance test (LAT) program to determine the long-term stability and reliability of the sensors. These acceptance programs will usually follow the ESA 9020 guidelines or the equivalent NASA guidelines.

## 9.4.2 **Radiation hardness**

### 9.4.2.1 *Radiation environment in space*

Space radiation is much more intense and different in nature compared to the radiation background on Earth. One of the main differences between commercial and space-qualified image sensors is the ability to operate in radiation environments.

Protons, electrons, and low-energy heavy ions trapped in Earth's magnetic field continuously circulate in two radiation belts around the planet (Barth et al., 2003). Protons have energies reaching hundreds of MeV and constitute the largest and most harmful fraction of trapped particles, occupying the first radiation belt at distances below 3 times Earth's radius. Electrons can have energies up to 10 MeV and extend to 10 times Earth's radius in the second radiation belt. Solar activity can dramatically change the particle population in the radiation belts, with solar flares and coronal mass ejections in particular able to increase their intensity by several orders of magnitude.

The GCRs are the second major radiation component, especially important for higher orbits and deep space. They originate outside the solar system and consist of 87% protons, 12% alpha particles, and 1% heavy ions (Bourdarie and Xapsos, 2008). Cosmic rays can have very large energies exceeding  $10^{14}$  MeV, making shielding difficult and expensive. Transiting protons and heavy ions originating from the Sun have energies in the GeV range and add to the GCR background, especially during solar eruptions.

The LEOs at low inclination, where the majority of satellites operate, are protected by the Earth's magnetic field from both radiation belts and the majority of GCR. The exceptions are the South Atlantic anomaly (SAA) (Bourdarie and Xapsos, 2008), where the proton radiation belt comes within 200 km from the surface, and polar orbits. The SAA is the most significant source of irradiation for satellites in LEO.

Elsewhere in the solar system the radiation environment could be much more hostile than in the Earth's orbit. For example, the magnetic field around Jupiter being 20 times stronger than Earth's, the trapped protons and electrons in Jupiter's radiation belts have much higher energy and intensity (Johnston, 2011). For a number of spacecraft exploring the Jovian system (e.g., Galileo, Juno, JUICE) this necessitates that image sensors are characterized to radiation levels an order of magnitude higher than those encountered in Earth orbits.

When interacting with the image sensor and the materials of the spacecraft, highly energetic particles create a cascade of secondary electrons, protons, neutrons, gammas, and X-rays through ionization and nuclear reactions, generating large numbers of additional particles. Charged particles can cause ionizing damage and SEE (Stassinopoulos and Raymond, 1988) in semiconductors leading to permanent and transient effects. Furthermore, charged and neutral particles with high enough energy

can displace atoms in the crystal lattice of the semiconductor and create defects which affect many aspects of device's operation.

#### 9.4.2.2 Ionizing damage

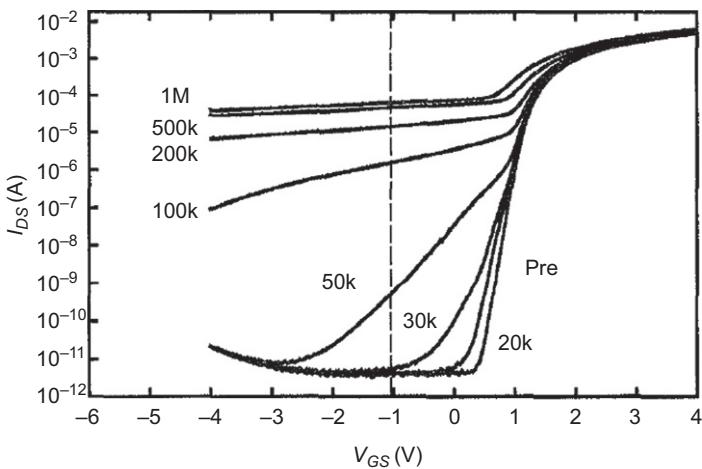
Ionizing radiation of sufficient energy can generate electron-hole pairs both in silicon and in the oxide layers used in the construction of a sensor. While the former is fundamental to image sensor operation, the latter can produce trapped oxide charge and defects at the silicon-oxide interfaces.

The total ionizing dose (TID) is used as a measure of the ionization damage effects. The TID is given as the total energy absorbed from ionization in a particular material in units of gray ( $1 \text{ Gy} = 1 \text{ J/kg}$ ), and also in rad, where  $100 \text{ rad} = 1 \text{ Gy}$ . Most sensors operating in LEO will receive TID below 10 krad over their lifetime and this is the usual dose required for space characterization. For spacecraft traversing the radiation belts or operating in deep space and in hostile places in the solar system the required TID tolerance could be above 50 krad.

Silicon dioxide ( $\text{SiO}_2$ ) is used as gate dielectric and for transistor isolation in the form of shallow trench isolation (STI) in modern CISs. Trapped positive charge in the oxides and at the Si- $\text{SiO}_2$  interface can cause undesirable threshold shift in MOS transistors and increased drain-source leakage current due to the formation of a parasitic conduction channel. The threshold shift is proportional to the square of the oxide thickness (Ma and Dressendorfer, 1989), but for thin gate oxides ( $< 10 \text{ nm}$ ) used in today's CMOS circuits the trapped charge dissipates via tunneling and the voltage shift is usually negligible. The STI represents much bigger danger due to its thickness, up to 300 nm in most cases, and can create large voltage shifts (Johnston et al., 2010). This can cause severe leakage currents in typical n-type MOSFETs (metal-oxide-semiconductor field-effect transistor) (Shaneyfelt et al., 1998; Fig. 9.5) at high TID, however, enclosed geometry transistors are very effective in preventing parasitic conduction from occurring (Tan et al., 2012).

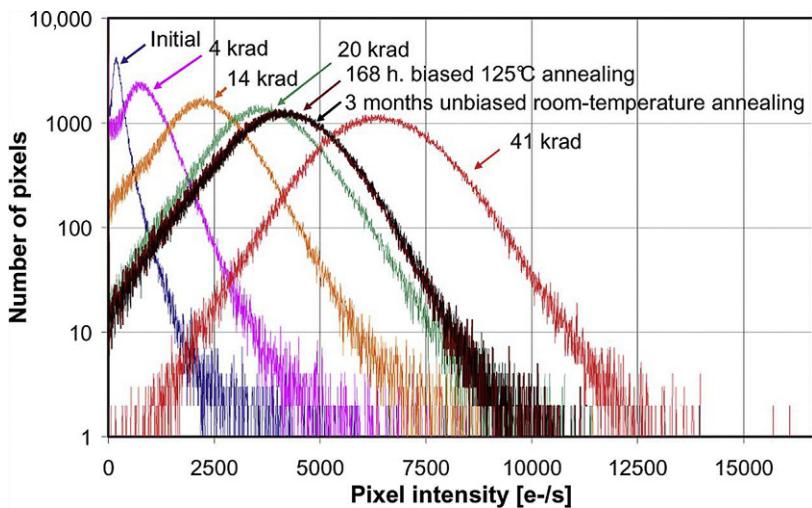
Ionizing radiation causes a number of other effects impacting many parameters of an imaging sensor. Increase in the dark current is universally observed in CIS (Fig. 9.6) and CCDs (Beaumel et al., 2010; Hopkinson et al., 1996) and is the most widely reported effect. Device cooling, where possible, can be used to reduce or eliminate the dark current. An increase in transistor noise and random telegraph signals (RTSs) has also been observed at high TID (Martin-Gonthier et al., 2012). Among the other, more complex effects are the decrease in the maximum output signal caused by charge recombination at interface traps, reduced source follower gain, and degraded linearity. Deterioration of the image lag in pinned photodiode (PPD, or 4T) image sensors, caused by potential "pockets" under the gate spacers, preventing complete charge transfer has also been observed (Goiffon et al., 2012; Rizzolo et al., 2018).

3T pixels can use a number of different radiation-hardened photodiodes by design (Goiffon et al., 2011), eliminating the vulnerabilities from the charge transfer in PPD sensors. Advanced 3T CIS operating at 100 Mrad for extreme radiation environments have been demonstrated (Goiffon et al., 2015).



**Fig. 9.5** Leakage current MOSFETs at increasing TID up to 1 Mrad.

From Shaneyfelt, M. R., et al., 1998. Challenges in hardening technologies using shallow-trench isolation. IEEE Trans. Nucl. Sci. 45(6), 2584–2592. doi: 10.1109/23.736501.



**Fig. 9.6** Increase in the dark current at room temperature after gamma irradiation from a Co-60 source.

From Beaumel, M., Herve, D., Van Aken, D., 2010. Cobalt-60, proton and electron irradiation of a radiation-hardened active pixel sensor. IEEE Trans. Nucl. Sci. 57(4), 2056–2065. doi: 10.1109/TNS.2010.2048043.

The CCDs are susceptible to ionizing damage as well, and they use much thicker gate oxides than CIS. However, CCDs are usually made with oxide-nitride gate dielectric which reduces the amount of voltage shift (Burt et al., 2009). As a result, CCD have been demonstrated to operate successfully to Mrad levels of soft X-rays (Jerram, 2016).

#### 9.4.2.3 Displacement damage

High-energy particles can have sufficient energy to displace silicon atoms out of their lattice positions, creating vacancies (a missing atom) and interstitials (an extra atom within the lattice structure). A silicon atom needs to receive at least 25 eV to be displaced (Lindström, 2003), which translates to 260 keV of kinetic energy for irradiating electrons and 175 eV for protons and neutrons. When the primary displaced silicon atom has kinetic energy above 2–5 keV (Srour et al., 2003; Lindström, 2003) it can in turn displace secondary atoms in a cascade, creating a defect cluster. Neutrons are particularly effective in creating clusters because of the absence of electrostatic repulsion. Vacancies and interstitials can migrate and eventually form stable, electrically active traps capable of capturing and releasing signal carriers through the Shockley-Read-Hall (SRH) mechanism (Sze, 1981). It is these traps in the silicon bulk that are responsible for the different DD effects.

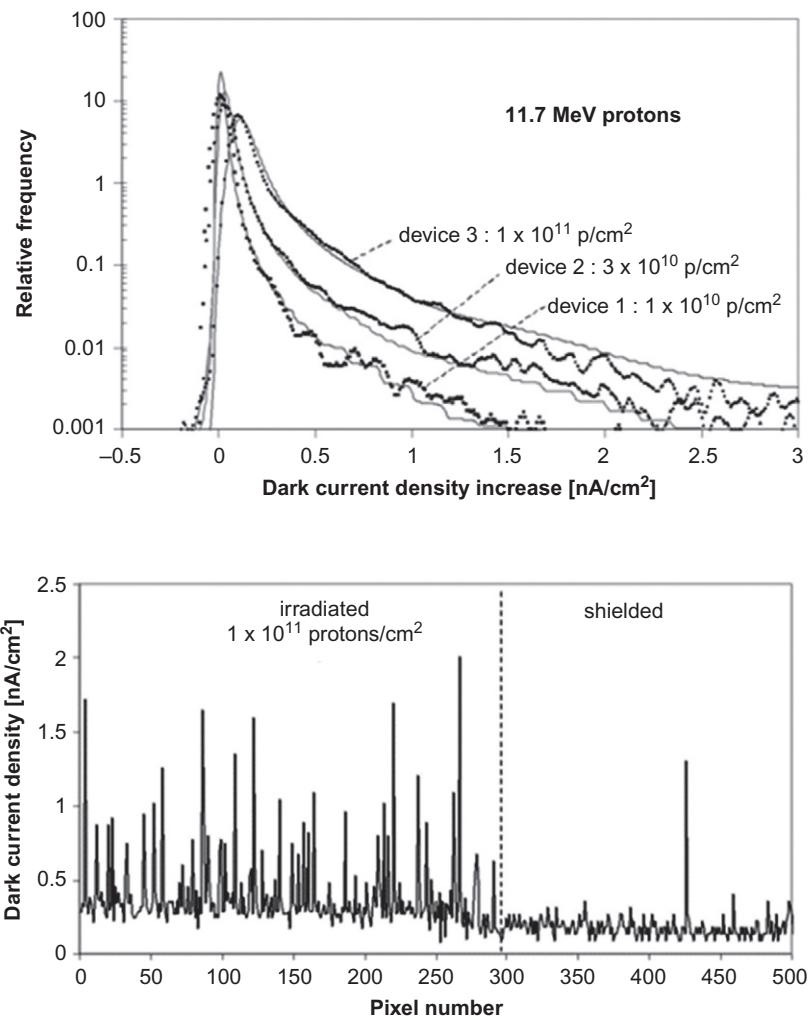
After DD by protons or other heavy particles, the mean dark current of a sensor increases, and a number of “hot pixels” with dark current much higher than the mean are observed (Hopkinson et al., 2004). This is visible in the images and also as a “tail” in the distribution of the dark current (Fig. 9.7). In comparison, gamma-irradiated sensors show only few hot pixels (Beaumel et al., 2010; Virmontois et al., 2014).

The dark current in hot pixels often exhibits RTS behavior in both CIS (Bogaerts et al., 2003; Goiffon et al., 2009; Virmontois et al., 2013) and CCDs (Hopkinson et al., 2008; Smith et al., 2004). Dark current RTS is particularly detrimental when cooling is not possible, and causes excess noise and spurious signals.

Another effect caused by bulk traps is present in sensors using charge transfer. The CCDs are particularly sensitive to DD because the signal charge travels long distances in the device. The capture and reemission of signal charge by radiation-induced bulk traps increases the CTI (Janesick et al., 1989) and sophisticated methods are required to correct the deferred signal (Massey, 2010; Massey et al., 2010). Fig. 9.8 illustrates the effect of the CTI in a proton-irradiated CCD.

Once a trap captures an electron it cannot capture another until it is released by emission. At cryogenic temperatures the emission time constants are of the order of seconds and most traps can be kept filled for a long time. This, in addition to the need to suppress the dark current is the reason why most astronomical CCDs are operated cold, typically in the  $-100^{\circ}\text{C}$  to  $-120^{\circ}\text{C}$  range. A large number of studies on radiation-induced CTI in CCDs are available (Holland, 1993; Hardy et al., 1998; Bebek et al., 2002), and a good overview of the history is provided by Pickel et al. (2003).

Recently, the attention has shifted to cryogenic irradiation of CCDs, aiming to better replicate the operating conditions in orbit where the devices are permanently kept

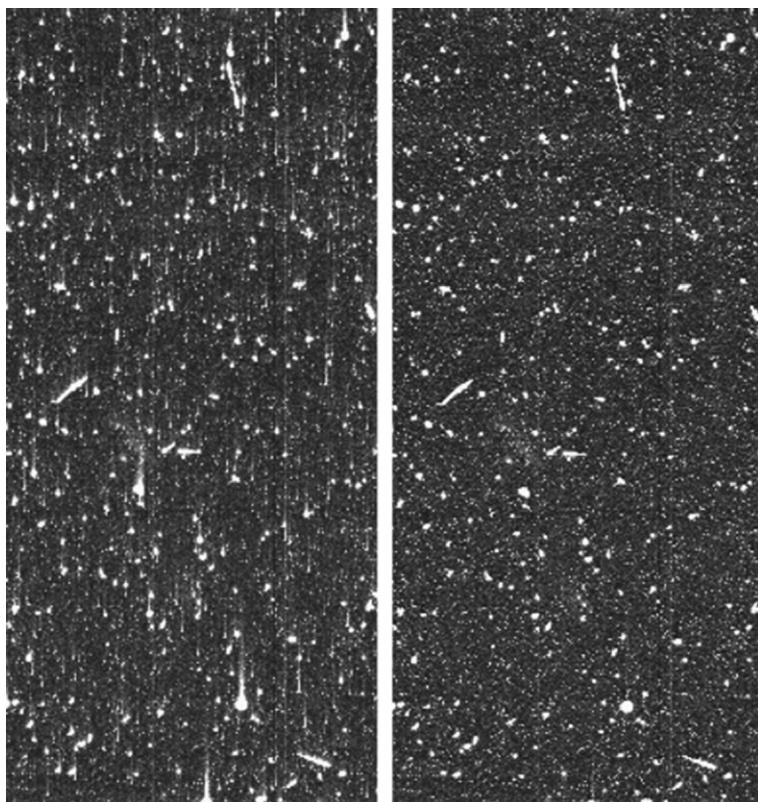


**Fig. 9.7** Dark current and hot pixels in proton-irradiated CIS.

From Bogaerts, J., et al., 2003. Total dose and displacement damage effects in a radiation-hardened CMOS APS. IEEE Trans. Electron Devices 50(1), 84–90. doi: 10.1109/TED.2002.807251.

cold (Bush et al., 2016). Advanced methods capable of characterizing individual traps and the evolution of their types and concentrations (Hall et al., 2014; Wood et al., 2017) are becoming commonplace.

A new class of sensors manufactured in a CMOS process, but incorporating low voltage CCDs have emerged (Marcelot et al., 2015; Rushton et al., 2015; Boulenc et al., 2017), promising efficient charge transfer in combination with the rich



**Fig. 9.8** Uncorrected CTI in proton-irradiated Hubble CCD (*left*), and after correction (*right*). From Massey, R., 2010. Charge transfer inefficiency in the Hubble space telescope since servicing mission 4. Mon. Not. R. Astron. Soc. Lett. 409(1), 109–113. doi: 10.1111/j.1745-3933.2010.00959.x.

functionality of CMOS electronics. Dubbed “CCD in CMOS,” these devices have worse initial CTI compared to “proper” CCDs and similarly suffer from radiation damage, but are competitive for TDI applications which require only few tens of charge transfers and operate at high signal levels.

#### 9.4.2.4 Single event effects

A whole range of effects caused by ionization of a single, high-energy particle is known as SEE (Petersen et al., 2013). These effects are particularly prominent in CMOS circuits due to their structure and functionality, which includes logic circuits, memory, oscillators, voltage references, and many more. Ionizing particles are characterized by the deposited energy, called linear energy transfer (LET) per unit track length (in units of MeV cm<sup>2</sup>/mg) or by the amount of charge generated (in pC/μm).

Heavy ions and energetic protons create dense ionization tracks in silicon and can generate large localized charge, capable of causing a range of SEE. They are classified according to the disruption they cause, as follows (Petersen, 2011):

- Single event transient (SET)—This can be detected in the image areas of CCDs and CIS as spurious signal in the form of a track or a spot spanning a number of pixels (Lomheim et al., 1990; Lalucaa et al., 2013a). In CMOS logic SET can temporarily change the input of a logic element, however, this may not have any lasting effect if the output does not change (e.g., in an AND gate where the other input is at logic low), or the output is not being read at that time in synchronous circuits.
- Single event upset (SEU)—The deposited charge causes lasting effects, most notable in memories where one or several bits can simultaneously change value. In sequential logic the state of flip-flops can change and persist until corrected. This can cause the wrong column or row to be selected during readout, resulting in image artifacts (Lalucaa et al., 2013b; Beaumel et al., 2014).
- Single event functional interrupt (SEFI)—This causes major disruption in the operation of CIS, but is correctable via reset of power cycling. In the case where sequential logic controls sensor readout, for example, implementing a state machine, SEU-induced bit change can make the logic enter an undefined state and “freeze.” SEU and SEFI can be mitigated by design techniques such as majority-vote circuitry and self-scrubbing error correction.
- Single event latchup (SEL)—Most CMOS circuits (with the exception of silicon-on-insulator) have a parasitic p-n-p-n structure created from adjacent n-type and p-type MOSFETs (Sexton, 2003). This structure acts as a thyristor, and can be triggered by the charge created by a traversing heavy ion. Once activated, the parasitic thyristor draws large current directly from the power supplies, and continues to do so until power is removed. If the current is large the SEL can be destructive, as shown for various integrated circuits (Irom et al., 2017). Many image sensors exhibit SEL, but fortunately most SELs are nondestructive—normal operation is resumed after a power cycle. A number of design techniques are available to combat SEL, such as adding guard rings, deep wells (Uemura et al., 2014), reducing the resistance of the transistor n- and p-wells, and the supply voltage. SEL occurrence increases with the temperature (Sexton, 2003), and this offers another way to reduce the impact for cooled sensors.

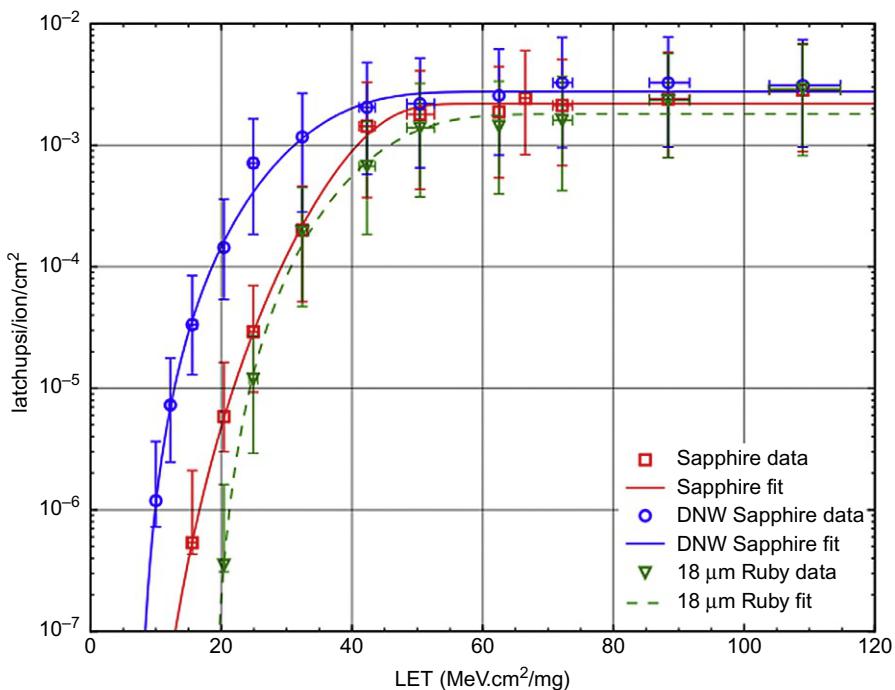
Most of the chip area of a CIS is occupied by the pixels, which contain only n-type MOSFETs and are immune to SEL. However, practically all CIS contain some CMOS circuitry, at least for column and row selection, and that is the part susceptible to SEU, SEFI, and SEL. A study on a sensor suffering from SEL in the decoders (Lalucaa et al., 2013b) shows how it affects the readout and also shows parasitic signals from infrared light, caused by the localized high current.

Fig. 9.9 shows the SEL rate by heavy ions in a CIS at increasing LET (Rushton et al., 2016). This data is useful for determining the SEL threshold and for calculating the total sensitive area in a sensor.

It is important to note that CCDs do not suffer from SEU, SEFI, or SEL because of the lack of CMOS logic circuitry and parasitic thyristors.

#### 9.4.3 Power dissipation

Power dissipation is an important consideration for space-based imaging systems, where there is a constant demand for low-power operation. In this respect CMOS imagers have distinct advantages over CCDs for several reasons.



**Fig. 9.9** Nondestructive SEL cross section in a CIS at increasing LET.

From Rushton, J. E., et al., 2016. Single event effects in 0.18μm CMOS image sensors. In Proceedings of SPIE, 99152Q p. doi: 10.1117/12.2235212.

To minimize the CTI astronomical CCDs are operated at very low temperatures, and many cameras support higher temperatures when not operational (e.g., around 0°C) and for annealing of bulk radiation-induced traps (above +100°C). This normally adds extra power requirements for CCD-based cameras.

There are important differences in the power needed to address one imaging row in CIS and CCDs. The whole of the image area of the CCD has to be driven in order to read out one row, and this involves large amplitude clocks (~10 V) driving the image area gates, with capacitance reaching many tens of nanofarads for large devices. In contrast, only one row needs to be selected and driven at a time in CIS for the equivalent operation, and much smaller capacitance has to be driven.

Also, CCDs use buried channel output transistor for low noise, and it typically runs at 25–30 V drain voltage and few millamps of drain current. Pixel source followers in CIS need only few microamps of drain current and rarely use voltages above 3.3 V. This is possible due to the parallel row readout architecture, which significantly relaxes the settling time requirements. Even with thousands of column source followers and buffers active at the same time, as is for rolling shutter readout, power dissipation in the outputs of CIS is usually lower than CCDs. The exception is perhaps very high-speed CIS incorporating per-column ADCs and hundreds of high-speed digital outputs.

### ***9.4.4 System considerations***

As mentioned earlier, sensor features and functionality that can reduce system complexity, mass, and cost are highly valuable for space applications.

The need to operate scientific CCDs at very low temperatures for good CTI and to provide annealing for radiation damage mitigation, as mentioned above, can add significant complexity. In comparison, the operating temperature of CIS is selected to suppress the dark current and can be much higher than that of CCDs, and annealing is not required due to the negligible radiation-induced CTI.

Typical CCDs do not implement any circuitry on chip except the readout source followers, so clock driving, biasing, analog signal processing, and digitization have to be provided externally. This can occupy large board space and also can dissipate substantial amount of power. To reduce system size the CCD services can be provided with custom integrated circuits (ICs), and normally two are needed—a high-voltage driver chip and an analog chip implementing gain, correlated double sampling (CDS), and A/D conversion.

In contrast, CIS can include everything needed for operation on the same chip and require only one or two low-voltage power supplies. Sensor configuration, control, and data transfer are provided over a small number of digital interfaces. This results in a highly capable and compact system, with advantages over a CCD-based camera due to the reduced size, mass, and power dissipation. The much smaller number of components and connections can also improve reliability. The image sensor can even do some or all of the raw image processing, which further helps reduce power as less data are transferred.

### ***9.4.5 Packaging, reliability, and availability***

#### ***9.4.5.1 Packaging***

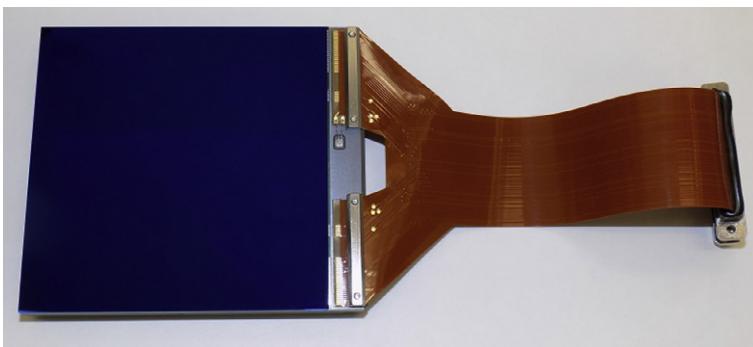
Packages that are used for low-cost terrestrial applications are not generally suitable for space. For space applications devices will frequently have to be used over an extended temperature range and survive the environment in space and during launch. It is also necessary to be able to inspect any solder joints to connect to the detector and hence surface mount components cannot be used.

A number of different materials are used for space applications but EO devices will normally be packaged in multilayer ceramics. These can be made of alumina ( $\text{Al}_2\text{O}_3$ ) if the operating point is close to room temperature or aluminum nitride (AIN) if operation at lower temperatures is required. Aluminum nitride is a better expansion match to silicon and so creates less stress in the package if cooled to low temperature. It is also a much better thermal conductor than alumina, and so for high-speed applications, where there can be high heat dissipation in the sensor, AIN is preferred to ensure that there is only a limited temperature gradient across a detector. [Table 9.1](#) lists the most important thermal parameters of several materials of interest in space packaging.

For operation at very low temperature (for astronomy) silicon carbide is often used as a packaging material. It has excellent thermal and mechanical properties and is a

**Table 9.1** Thermal properties of materials used in image sensor packaging

Material	CTE at room temperature (ppm/ $^{\circ}$ C)	Thermal conductivity (W/m K)	Density (g/cm $^3$ )
Alumina	7	17	3.6
Aluminum Nitride	4.7	150	3.4
Silicon Carbide	2	175	3.1
Silicon	2.5	150	2.3



**Fig. 9.10** Image of the CCD270, a 8-cm $^2$  sensor for use in the PLATO mission.  
Image: Teledyne-e2v.

very good match to silicon, and also has low density. However, its manufacture is a specialized process, so it is only used where necessary, typically for large-area devices that need to be very flat. Silicon carbide packages also require a separate electrical interface, and for space and astronomy a flexi circuit can be used. The use of silicon carbide for large-area sensors has the added advantage that focal planes are frequently built using the same material, ensuring that there is no thermal mismatch between the sensor and the focal plane. In the past, metal packages have been used for low-temperature operation, for example, in the Rosetta Osiris 2 k  $\times$  2 k sensor, but the better thermal and mechanical properties of silicon carbide mean that it is now the preferred choice (Fig. 9.10).

#### 9.4.5.2 Reliability

While electro-optical performance and ease of use are important, the most critical requirement for space applications is reliability. In general sensors for any critical mission will undergo an extensive set of reliability tests before flight models can be delivered. These are similar to any electronic component for space but follow a particular test flow for image sensors defined by ESA9020 standard, or the equivalent NASA requirements, and involve tests in the following categories:

*environmental subgroup*: high-temperature storage and temperature cycling to ensure survival in the space environment;

*mechanical subgroup*: shock and vibration to ensure that the detector will survive the launch;  
*endurance subgroup*: high-temperature endurance tests to ensure that the sensor will remain operational through the mission lifetime;

*assembly capability*: to validate the build standard for a space environment;

*radiation*: to ensure that any degradation from the radiation environment seen in space is allowed for in the performance predictions.

There are actually two levels of tests required, those that are used to qualify a new design, and LATs to verify the performance of a batch (or lot) intended for use in a particular program. The precise details are dependent on the particular program.

As several sensors are required for each of these test streams, a project that will only require the flight of a single device can require the manufacture of 30–40 devices for qualification.

In addition to the qualification of the sensor build standard and the batch or manufacturing lot tests, each sensor built will undergo screening tests on top of the standard electro-optical tests to ensure that there are no early life failures. These in-process screening tests will normally include burn-in at an elevated temperature and temperature cycling.

The use of commercial CMOS foundries presents a problem in that the process is driven to continually improve by commercial constraints, and the changes in manufacturing process are not always fully visible from one manufacturing run to the next. This is not a problem for a given project which will normally require only one or two silicon batches made over a short period of time but can become an issue if repeat runs of the same build are required with a separation of several years and hence requalification may be required.

## 9.5 Space applications

### 9.5.1 Altitude and orbit control systems and star trackers

Star tracker sensors operate in staring mode and monitor a small number of stars at a modest frame rate but with very high positional accuracy of a fraction of the pixel size. Generally there is sufficient signal from the stars so the highest quality electro-optical performance is not required.

Star tracker sensors have been the first type of space image sensor that have mostly moved to use CMOS imagers. This has been the result of many years of funding by ESA for several generations of CMOS imagers. The latest generation is the radiation hard HAS2 sensor ([Hervé et al., 2012](#)) with an array of  $1024 \times 1024$ ,  $18\text{ }\mu\text{m}$  pixels operating in rolling shutter mode. Typically star trackers operate at a small number of frames per second (<10).

Star tracker sensors must image a small number of stars and hence the windowing capability of CMOS imagers is a significant advantage to reduce the amount of unwanted data that is produced. The image of each star is defocused and then the

centroid of the resulting charge cloud is measured down to a small percentage of a pixel. Due to inherent nonuniformity within the pixel the centroiding capability of a CMOS imager is not as good as an equivalent CCD but is good enough for nearly all applications. The ease of use and windowing capability therefore means that CMOS imagers have become the first choice for most star tracker instruments.

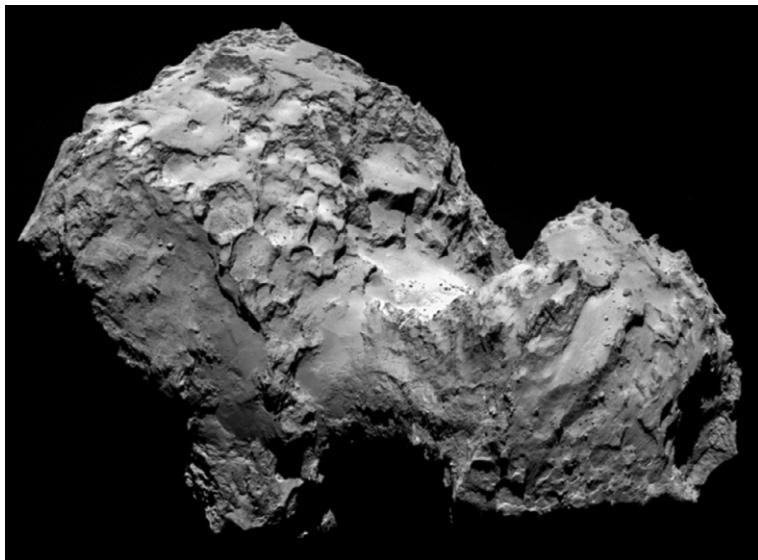
### 9.5.2 Earth Observation and solar system exploration

#### 9.5.2.1 Step-and-stare imaging

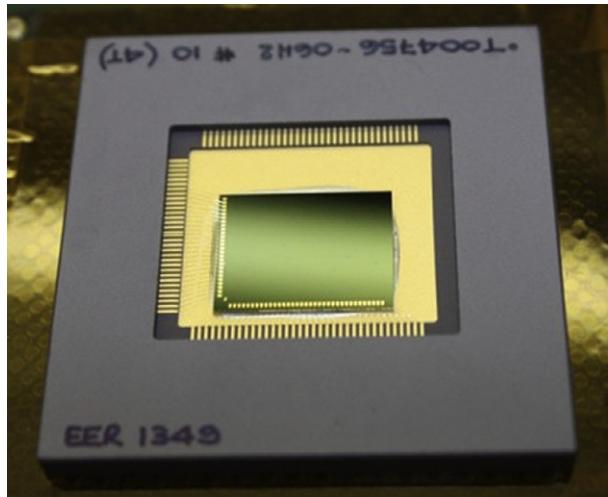
The CCDs have been used in the majority of space missions in Earth and planetary observations so far. However, more and more CMOS sensors are penetrating this field, especially in lower cost spacecraft such as Cubesats. The “step-and-stare” imaging mode, being simpler to implement than TDI, is often used and is also the only mode possible when the spacecraft is not in orbit or images from a great distance.

The OSIRIS camera on the Rosetta spacecraft (Keller et al., 2007) is a great example of an instrument for solar system exploration. The camera uses two identical CCDs in BSI configuration for narrow and wide field imaging, achieving enhanced sensitivity in UV and near-IR. An iconic image of the comet 67P/Churyumov-Gerasimenko, obtained by OSIRIS, is shown in Fig. 9.11.

For the upcoming JUICE spacecraft to Jupiter a CIS has been selected due to the high radiation background around the planet. The JANUS visible camera onboard



**Fig. 9.11** Image of the comet 67P/Churyumov-Gerasimenko taken by the OSIRIS camera on Rosetta, using a  $2\text{ k} \times 2\text{ k}$ ,  $13.5\text{ }\mu\text{m}^2$  pixels, BSI CCD42-40 made by Teledyne e2v.  
Image: ESA.



**Fig. 9.12** CIS115, the CMOS image sensor for the JANUS camera.

Image: Teledyne e2v.

JUICE will use a BSI CIS115 (shown in Fig. 9.12), featuring  $2000 \times 1504$   $7\text{ }\mu\text{m}$  square pixels and four analog outputs. CIS115 uses PPD pixels and exhibits readout noise below  $5\text{ e}^-$  RMS at 4 Mpix/s readout rate from each output. The sensor's radiation hardness has been extensively characterized using gamma, proton, electron, and heavy ion irradiations to simulate the environment around Jupiter (Soman et al., 2016).

### 9.5.2.2 Scanning imagers

As discussed in the previous sections most recent scanning systems now use TDI sensors rather than linear arrays, as the sensitivity and hence the resolution is very much higher when several pixels are added together. To carry out TDI operation with no additional noise, the charge is moved down the sensor at a rate corresponding to the movement of the sensor over the ground. For a CCD this is simple as the inherent mode of operation involved moving charge in the sensor to the register. At the bottom of the sensor the signal is readout through the CCD register. Operating in this way means that many lines can be added together with no additional noise. Over the years CCDs have been used in TDI mode in a large number of systems, for example, in the multispectral visible imaging camera in the RALPH instrument that was used to take the first ever color images of Pluto (Fig. 9.13) (Reuter et al., 2008) and in Pleiades taking extremely good quality, high-resolution images of the Earth (Gaudin-Delrieu et al., 2008) with a resolution of under 0.7 m.

The imager used in RALPH has six independent TDI arrays, each with a width of 5024 pixels and 32 TDI lines that are added together to give the required sensitivity. Two of the TDI arrays are used in panchromatic mode using the whole visible



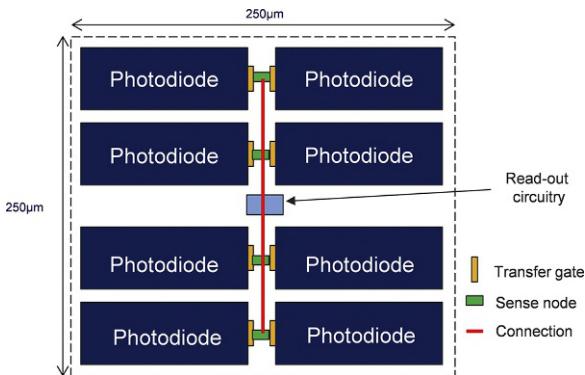
**Fig. 9.13** Image of Pluto from New Horizons (Copyright NASA/JHUAPL/SwRI).

spectrum, while the others are combined with filters to provide blue, red, NIR, and methane channels. In Pleiades the focal plane consists of five panchromatic imagers each with 6000 pixels on a 13- $\mu\text{m}$  pitch, operating with a TDI length of up to 20 pixels giving a total image width of 30,000 pixels. In the case of Pleiades the multispectral (MS) lines are provided by linear sensors rather than TDI imagers. This is possible because they are 4 times the pixel size of the panchromatic pixels.

As the requirements for ground resolution continually increase it has become extremely difficult to achieve the required pixel size and row speed using CCDs.

Until recently, it was not obvious that TDI operation using charge movement was possible using CMOS technology, as there is only a single level of polysilicon available to create the electrodes for control of the charge, and only a low voltage can be applied to any electrode. This makes efficient transfer of charge with zero loss of signal extremely difficult.

However, developments by CMOS foundries have enabled the processing of charge transfer CMOS imagers that effectively have a CCD image area with a CTE of >99.999% and CMOS readout circuits (Korthout et al., 2017). These offer significant advantages to the existing CCD technology in terms of the minimum pixel size and line speed achievable. As the pixel size can now be reduced down to 5  $\mu\text{m}$  or smaller, it is possible to produce a single sensor with a width of up to 24,000 pixels that incorporates the panchromatic and several MS sensors on the same die. In addition, and of equal importance is the fact that the CMOS TDI sensors have digital outputs and operate at low voltages. This means that the focal plane is very much smaller, simpler, and has much lower power consumption. This is critical for ‘New Space’



**Fig. 9.14** MetImage pixel format.

applications requiring large constellations of small low-cost satellites providing high-quality data with very high revisit rates over all parts of the Earth.

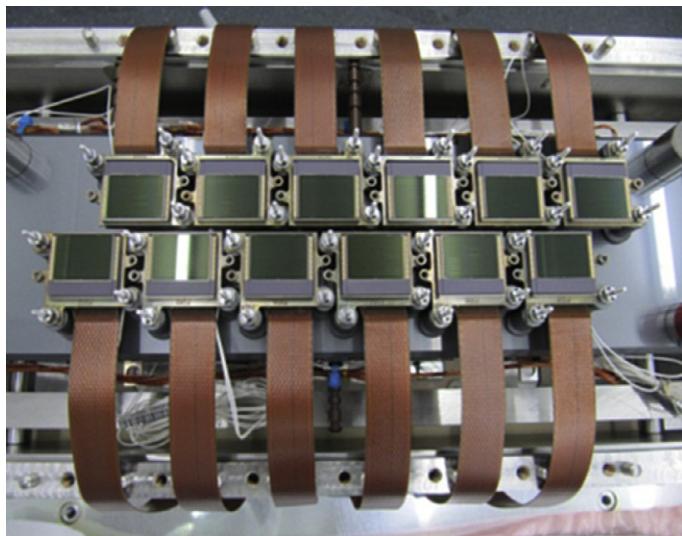
A new generation of sensors for weather satellites (MTG VisDA and MetImage on MetOp SG) are now using CMOS imagers. These present an unusual challenge for CMOS technology in that the pixel sizes are extremely large. In the case on MetImage, the sensor has a pixel size of  $250 \mu\text{m}^2$  (Fig. 9.14). The challenge for such a large pixel is to ensure that all of the charge is removed from the photodiode during the charge transfer operation. This is achieved by combining eight different photodiodes onto one sense node with channel shaping within the pixel to ensure that the charge drifts toward the transfer gate. Both MetImage and MTG operate in scanning mode with a number of different spectral bands (seven in the case of MetImage) and relatively low resolution, but at a high frame rate providing a complete image every 1.8 s.

### 9.5.2.3 Hyperspectral imagers

The main use of hyperspectral imagers is in missions to monitor the health of the Earth. The leading activity of recent years is ESA's Copernicus program which will provide “*accurate, timely, and easily accessible information to improve the management of the environment, understand and mitigate the effects of climate change, and ensure civil security.*”

For Copernicus, ESA is developing a new family of satellites called Sentinels. Many of these satellites have a visible imaging element mostly using imagers operating in a hyperspectral mode. The exception is Sentinel 2 which is an MS imager that uses a focal plane consisting of 12 different sensors (Fig. 9.15), each with 10 spectral bands. For Sentinel 2 the highest resolution band is 10 m, so it is possible to use linear sensors. There are 31,000 pixels across the total instrument in the high-resolution band. The CMOS imager used in Sentinel 2 consists of 10 linear arrays with pixel sizes ranging from  $7.5 \times 7.5$  to  $15 \times 45 \mu\text{m}$ , and has a total of three analog outputs reading at 4.8 Mpix/s (Martin-Gonthier et al., 2010).

The Ocean and Land Color instrument (OLCI) on Sentinel 3 is an example of a sensor that operates in the hyperspectral mode. This images the Earth with a resolution of approximately 300 m in 21 spectral bands and with the launch of the second satellite can now provide updates of the entire planet every 2 days. The detailed data from



**Fig. 9.15** Sentinel 2 VNIR Flight Focal Plane [Airbus (France) and e2v Technologies (UK)].

Sentinel 3 allows agriculture and ocean monitoring in great spectral detail over the range from 400 to 1000 nm (Fig. 9.16).

Sentinel 5P (TroPOMI) also disperses a spectrum across the sensor, although in this case there are four different instruments optimized for different parts of the spectrum. The UV and visible sensors are backthinned  $1\text{ k} \times 1\text{ k}$  CCDs with  $26\text{ }\mu\text{m}^2$  pixels with AR coating and processing optimized for different parts of the spectrum (Jerram et al., 2017). TroPOMI is designed to monitor pollution including ozone, methane, formaldehyde, aerosol, carbon monoxide,  $\text{NO}_2$ , and  $\text{SO}_2$  across the Earth with a ground resolution of around 7 km (Fig. 9.17).

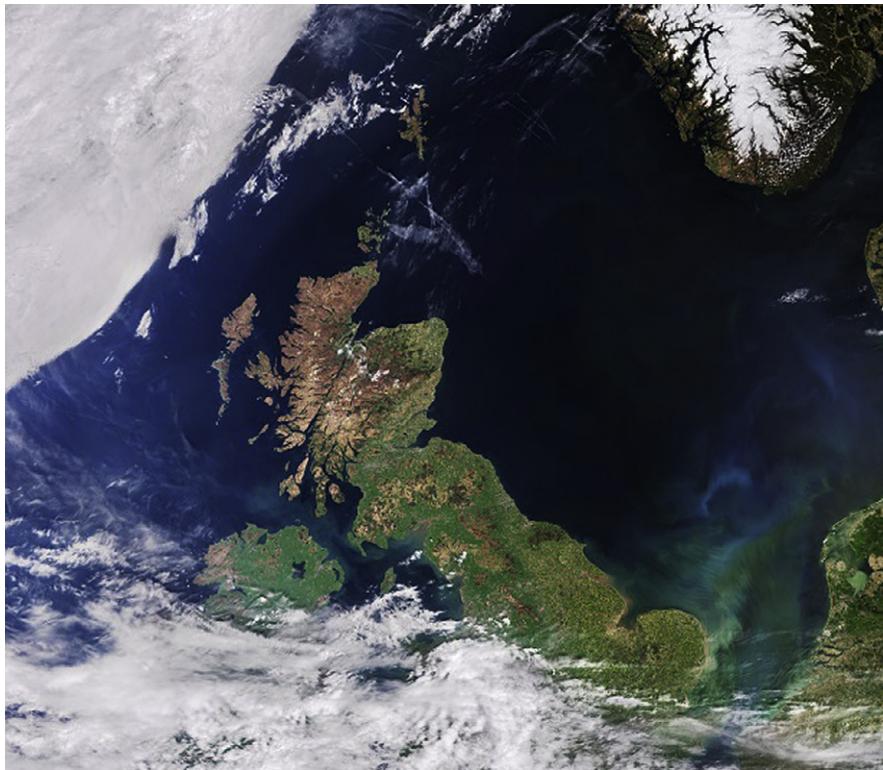
There are many other ESA missions in preparation using sensors that will provide detailed spectral data for EO including FLEX (Floris), Sentinel 4, Sentinel 5, and 3MI. All of these missions will use CCDs.

### **9.5.3 Science and astronomy**

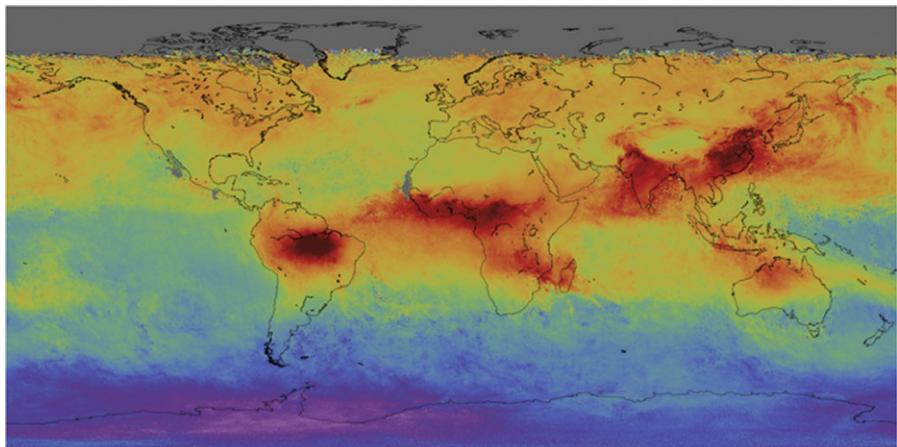
Scientific imaging places high demands on image sensors in terms of image quality, sensor and focal plane size, and cooling requirements. Despite the growth of CIS for almost every application, virtually all high-performance science and astronomy sensors in space are still CCDs.

#### **9.5.3.1 High-performance applications in astronomy**

Very large and even wafer-scale devices are usually required for space-based astronomy. Currently only CCDs have the required image quality and TRL maturity for such missions, despite the drawbacks from radiation-induced CTI and low-temperature



**Fig. 9.16** ESA Sentinel 3 image of the UK showing algal blooms in the bottom right-hand corner of the image.  
Image: ESA.



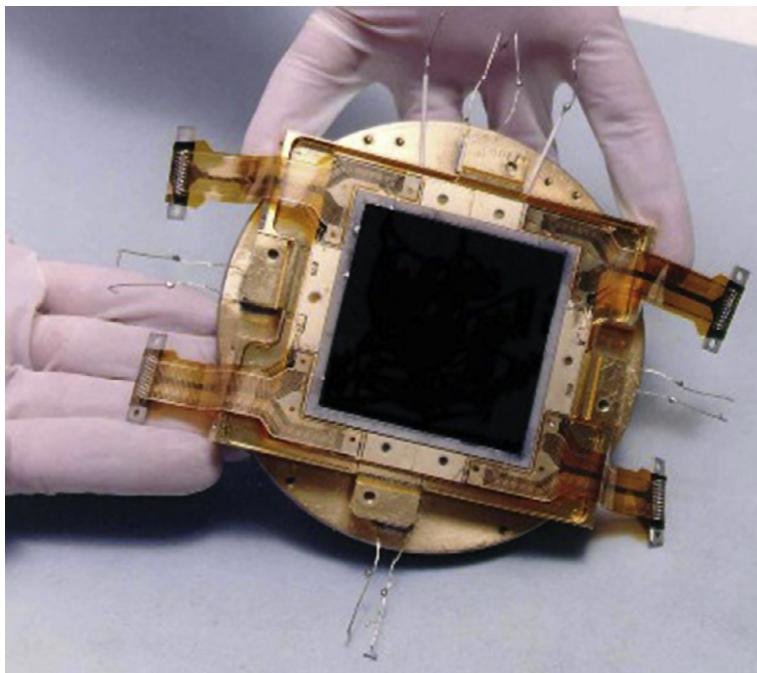
**Fig. 9.17** ESA Sentinel 5P image showing global carbon monoxide pollution across the globe.  
Image: ESA.

operation. The CCDs are generally backside illuminated and feature thick, fully depleted silicon and customized AR coatings for high QE from UV to near-IR wavelengths.

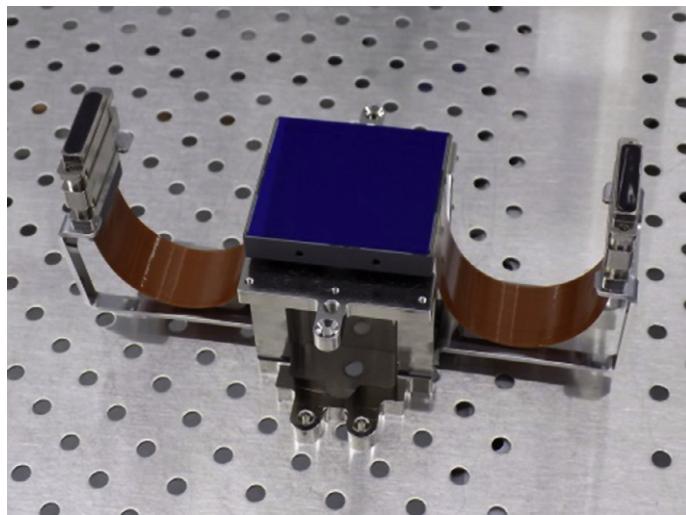
One of the most prominent examples of astronomy CCDs in space is the WFC3 camera installed in the Hubble Space Telescope. It uses two closely mounted BSI CCDs (Fig. 9.18) with high QE in the 200–1000 nm wavelength range, each with  $4\text{ k} \times 2\text{ k}$   $15\text{ }\mu\text{m}$  pixels. The CCDs operate in inverted (MPP) mode at  $-83^\circ\text{C}$  for dark current below  $2\text{ e}^-/\text{pixel/h}$  and feature readout noise of  $3\text{ e}^- \text{ RMS}$ .

The CCDs are very well suited for building of focal plane arrays because they allow close mounting (also known as butting) of devices in one or two dimensions. The gaps between the outermost imaging pixels from adjacent devices can be only few hundred micrometers due to the minimal edge circuitry in CCDs. Euclid, a forthcoming space telescope by ESA will use a focal plane array made of 36 large BSI CCD in its visual imager instrument. Each CCD (shown in Fig. 9.19) has  $4\text{ k} \times 4\text{ k}$   $12\text{ }\mu\text{m}$  pixels and implements thin gate dielectric for reducing the radiation-induced voltage shifts and charge injection for CTI minimization.

The largest focal plane in space belongs to the Gaia space telescope, shown in Fig. 9.20. Physically it occupies an area of  $0.5\text{ m}^2$  and contains a total of 106 CCDs.



**Fig. 9.18** Teledyne e2v CCD43 assembly for the Hubble WFC3.  
Image: NASA/HST.



**Fig. 9.19** The Euclid CCD.

Image: Teledyne e2v.

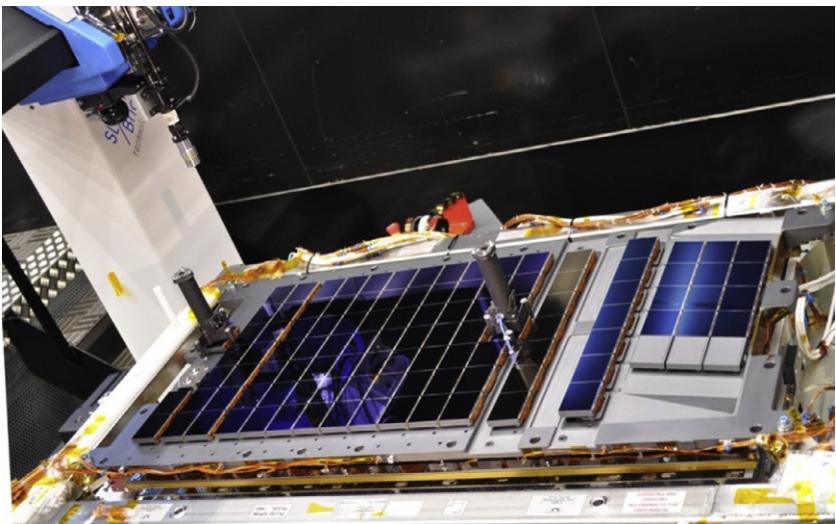
Using TDI mode imaging, Gaia's CCDs perform a number of functions such as astrometry, photometry, and spectroscopy.

### 9.5.3.2 X-ray imaging

X-rays from celestial objects are absorbed by Earth's atmosphere and this requires that imaging is performed from space. A number of telescopes launched over the past two decades (XMM-Newton, Chandra, Suzaku) have provided a new view of the universe in X-rays. Silicon CCDs with relatively large pixels are used in all three telescopes to provide detection capabilities at soft X-rays alongside other instruments. BSI CCDs are preferred and used for extended response at very low-energy X-rays, but designs using "open gate" front side illuminated devices are also possible and used in XMM-Newton (Fig. 9.21).

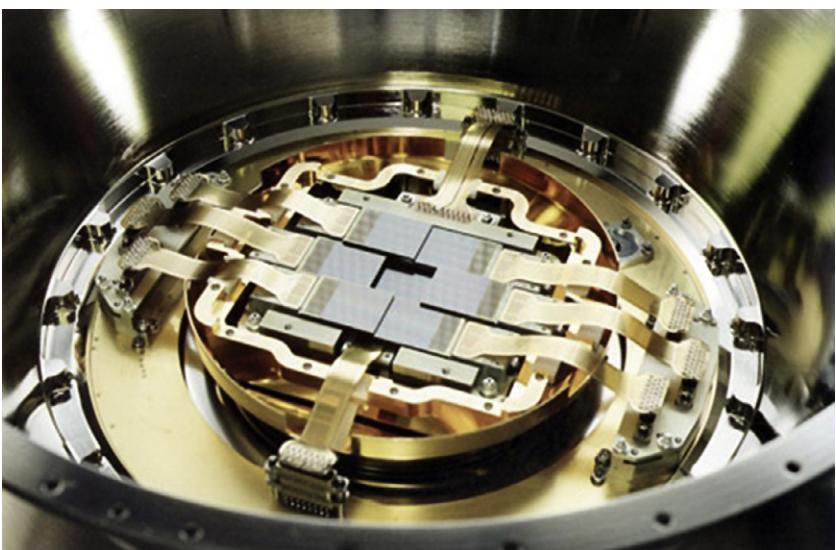
### 9.5.3.3 Sun imaging

Large numbers of spacecraft are in operation studying solar physics and helping predict the space weather. Imaging in UV and extreme UV is required for solar observations, which is possible with BSI silicon sensors. For example, the Atmospheric Imaging Assembly (AIA) onboard NASA's Solar Dynamics Observatory (SDO) uses four BSI CCDs with  $4\text{ k} \times 4\text{ k}$ ,  $12\text{-}\mu\text{m}^2$  pixels to image the Sun in 10 wavelength bands (Fig. 9.22).

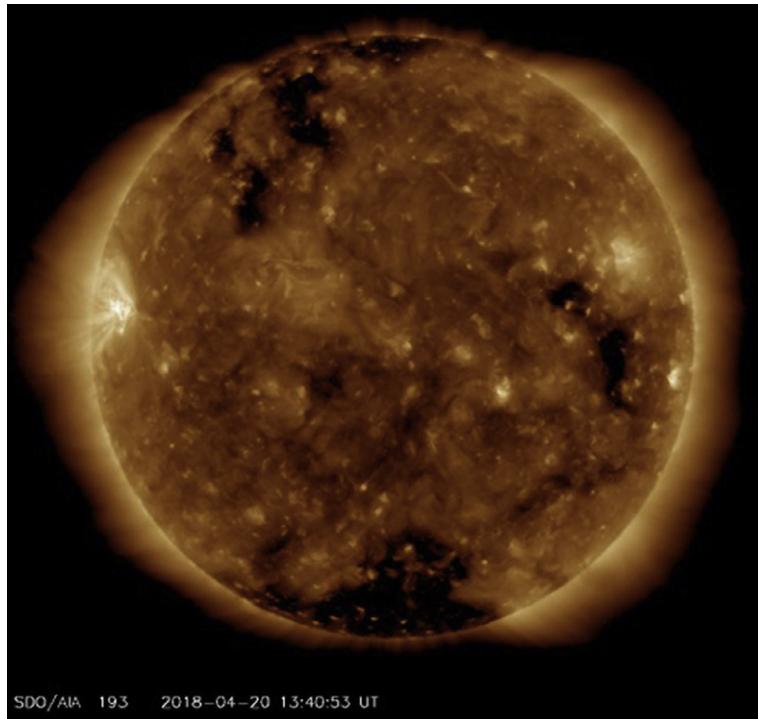


**Fig. 9.20** The focal plane of Gaia.

Image: ESA, <http://sci.esa.int/gaia/>.



**Fig. 9.21** The MOS-CCD EPIC camera on XMM Newton, consisting of seven CCDs with  $600 \times 600, 40\text{-}\mu\text{m}^2$  pixels each.



**Fig. 9.22** Sun's image from SDO's AIA camera at extreme UV (193 Å).  
Image: NASA/SDO.

## 9.6 Conclusion and longer term trends

Space poses unique challenges to image sensors, requiring very high electro-optical performance combined with radiation resistance, low power dissipation, reliability, and long-term stability. The large majority of current high-performance applications currently use CCDs due to their heritage and excellent uniformity and QE. However, in the past 10 years more spacecraft have been equipped with CMOS sensors, and for some applications such as star trackers they are dominant.

The slow trend toward an increased use of CMOS sensors for space imaging will almost certainly continue due to improving image quality, low noise, low power dissipation, and superior radiation hardness to CCDs. Some scientific applications such as astronomy are likely to retain CCDs when large focal planes are needed.

Space remains a conservative market, so in the medium term the CMOS devices used in space are likely to have only a small level of digital processing on chip, mainly to include digital outputs and sequencing. In the longer term the devices used in space will follow the same route as commercial sensors and it is almost certain that there will be an increasing level of integration of processing on the sensor.

## References

- Barth, J.L., Dyer, C.S., Stassinopoulos, E.G., 2003. Space, atmospheric, and terrestrial radiation environments. *IEEE Trans. Nucl. Sci.* 50 (3), 466–482. <https://doi.org/10.1109/TNS.2003.813131>.
- Beaumel, M., Herve, D., Van Aken, D., 2010. Cobalt-60, proton and electron irradiation of a radiation-hardened active pixel sensor. *IEEE Trans. Nucl. Sci.* 57 (4), 2056–2065. <https://doi.org/10.1109/TNS.2010.2048043>.
- Beaumel, M., et al., 2014. Proton, electron, and heavy ion single event effects on the HAS2 CMOS image sensor. *IEEE Trans. Nucl. Sci.* 61 (4), 1909–1917. <https://doi.org/10.1109/TNS.2014.2307759>.
- Bebek, C., et al., 2002. Proton radiation damage in p-channel CCDs fabricated on high-resistivity silicon. *IEEE Trans. Nucl. Sci.* 49 (3), 1221–1225. <https://doi.org/10.1109/TNS.2002.1039641>.
- Bogaerts, J., et al., 2003. Total dose and displacement damage effects in a radiation-hardened CMOS APS. *IEEE Trans. Electron Devices* 50 (1), 84–90. <https://doi.org/10.1109/TED.2002.807251>.
- Boulenc, P., et al., 2017. High speed TDI embedded CCD in CMOS sensor. In: Karafolas, N., Cugny, B., Sodnik, Z. (Eds.), International Conference on Space Optics—ICSO 2016. SPIE, p. 152. <https://doi.org/10.1117/12.2296149>.
- Bourdarie, S., Xapsos, M., 2008. The near-earth space radiation environment. *IEEE Trans. Nucl. Sci.* 55 (4), 1810–1832. <https://doi.org/10.1109/TNS.2008.2001409>.
- Burt, D., et al., 2009. Improving radiation tolerance in e2v CCD sensors. In: Warren, P.G. et al., (Ed.), Proceedings of SPIE. Elsevier, p. 743902. <https://doi.org/10.1117/12.825273>.
- Bush, N., et al., 2016. Cryogenic irradiation of an EMCCD for the WFIRST coronagraph: preliminary performance analysis. In: Holland, A.D., Beletic, J. (Eds.), Proceedings of SPIE. p. 99150A. <https://doi.org/10.1117/12.2234628>.
- Choi, J.-K., et al., 2012. GOCI, the world's first geostationary ocean colour observation platform. *J. Geophys. Res.* 117C09004 <https://doi.org/10.1029/2012JC008046>.
- Endicott, J., et al., 2012. Charge-coupled devices for the Euclid M-class mission. In: SPIE Astronomical Instrumentation 2012.
- ESA PLATO Community n.d. <http://sci.esa.int/plato/>.
- Flamini, E., et al., 2010. SIMBIO-SYS: the spectrometer and imagers integrated observatory system for the BepiColombo planetary orbiter. *Planet. Space Sci.* 58, 125–143. <https://doi.org/10.1016/j.pss.2009.06.017>.
- Gaudin-Delrieu, C., et al., 2008. The high resolution optical instruments for the Pleiades HR Earth observation satellites. In: Proceedings of SPIE, International Conference on Space Optics ICSO2008 vol. 10566. <https://doi.org/10.1117/12.2308258>.
- Goiffon, V., et al., 2009. Multilevel RTS in proton irradiated CMOS image sensors manufactured in a deep submicron technology. *IEEE Trans. Nucl. Sci.* 56 (4), 2132–2141. <https://doi.org/10.1109/TNS.2009.2014759>.
- Goiffon, V., et al., 2011. Generic radiation hardened photodiode layouts for deep submicron CMOS image sensor processes. *IEEE Trans. Nucl. Sci.* 58 (6 PART 1), 3076–3084. <https://doi.org/10.1109/TNS.2011.2171502>.
- Goiffon, V., et al., 2012. Radiation effects in pinned photodiode CMOS image sensors: pixel performance degradation due to total ionizing dose. *IEEE Trans. Nucl. Sci.* 59 (6), 2878–2887. <https://doi.org/10.1109/TNS.2012.2222927>.

- Goiffon, V., et al., 2015. Multi-MGy radiation hard CMOS image sensor: design, characterization and X/gamma rays total ionizing dose tests. *IEEE Trans. Nucl. Sci.* 62 (6), 2956–2964. <https://doi.org/10.1109/TNS.2015.2490479>.
- Hall, D.J., et al., 2014. Determination of in situ trap properties in CCDs using a “single-trap pumping” technique. *IEEE Trans. Nucl. Sci.* 61 (4), 1826–1833. <https://doi.org/10.1109/TNS.2013.2295941>.
- Hardy, T., Murowinski, R., Deen, M.J., 1998. Charge transfer efficiency in proton damaged CCD’s. *IEEE Trans. Nucl. Sci.* 45 (2), 154–163. <https://doi.org/10.1109/23.664167>.
- Hervé, D., et al., 2012. New radiation test results on HAS2 CMOS image sensor. In: *Workshop on Radiation effects on Optoelectronic Devices*.
- Holland, A.D., 1993. The effect of bulk traps in proton irradiated EEV CCDs. *Nuclear Instr. Meth. Physics Res.* A326, 335–343. [https://doi.org/10.1016/0168-9002\(93\)90374-Q](https://doi.org/10.1016/0168-9002(93)90374-Q).
- Hopkinson, G.R., Dale, C.J., Marshall, P.W., 1996. Proton effects in charge-coupled devices. *IEEE Trans. Nucl. Sci.* 43, 614–627. <https://doi.org/10.1109/23.490905>.
- Hopkinson, G.R., Goiffon, V., Mohammadzadeh, A., 2008. Random telegraph signals in proton irradiated CCDs and APS. *IEEE Trans. Nucl. Sci.* 55 (4), 2197–2204. <https://doi.org/10.1109/TNS.2008.2000764>.
- Hopkinson, G.R., Mohammadzadeh, A., Harboe-Sorensen, R., 2004. Radiation effects on a radiation-tolerant CMOS active pixel sensor. *IEEE Trans. Nucl. Sci.* 51 (5), 2753–2762. <https://doi.org/10.1109/TNS.2004.835108>.
- Irom, F., Allen, G.R., Vartanian, S., 2017. Single-event latchup measurements on COTS electronic devices for use in ISS payloads. In: *2017 IEEE Radiation Effects Data Workshop (REDW)*. IEEE, pp. 1–6. <https://doi.org/10.1109/NSREC.2017.8115428>.
- ISO 16290:2013. Space Systems—Definition of the Technology Readiness Levels (TRLs) and Their Criteria of Assessment. Available from [www.iso.org](http://www.iso.org).
- Janesick, J., Elliott, T., Pool, F., 1989. Radiation damage in scientific charge-coupled devices. *IEEE Trans. Nucl. Sci.* 36 (1), 572–578. <https://doi.org/10.1109/23.34503>.
- Jerram, P., 2016. e2v image sensors for high end applications. In: *International Forum on Detectors for Photon Science (IFDEPS)*, Japan Available at: <http://xfel.riken.jp/IFDEPS2016/home.html>.
- Jerram, P., Morris, D., 2016. Recent sensor designs for earth observation. In: Xiong, X.J., Kuriakose, S.A., Kimura, T. (Eds.), *Proceedings of SPIE*. p. 988111. <https://doi.org/10.1117/12.2229038>.
- Jerram, P.A., et al., 2017. Hyperspectral CMOS imager. In: Kadokawa, N. (Ed.), *International Conference on Space Optics—ICSO 2010*. SPIE, p. 67. <https://doi.org/10.1117/12.2309195>.
- Johnston, A., 2011. Space radiation effects and reliability considerations for the proposed Jupiter Europa Orbiter. In: *IEEE International Reliability Physics Symposium Proceedings*, pp. 165–170. <https://doi.org/10.1109/IRPS.2011.5784471>.
- Johnston, A.H., Swimm, R.T., Miyahira, T.F., 2010. Low dose rate effects in shallow trench isolation regions. *IEEE Trans. Nucl. Sci.* 57 (6 PART 1), 3279–3287. <https://doi.org/10.1109/TNS.2010.2085452>.
- Keller, H.U., et al., 2007. OSIRIS—the scientific camera system onboard Rosetta. *Space Sci. Rev.* 128 (1–4), 433–506. <https://doi.org/10.1007/s11214-006-9128-4>.
- Korthout, et al., 2017. A 256 stage charge domain TDI CMOS imager with 12 bit column parallel AD conversion supporting up to 300kHz line rate. In: *Workshop on CMOS Image Sensors for High Performance Applications, Toulouse*.
- Lalucaa, V., et al., 2013a. Single-event effects in CMOS image sensors. *IEEE Trans. Nucl. Sci.* 60 (4), 2494–2502. <https://doi.org/10.1109/TNS.2013.2260355>.

- Lucaa, V., et al., 2013b. Single event effects in 4T pinned photodiode image sensors. *IEEE Trans. Nucl. Sci.* 60 (6), 4314–4322. <https://doi.org/10.1109/TNS.2013.2287751>.
- Lindström, G., 2003. Radiation damage in silicon detectors. *Nuclear Instr. Meth. Physics Res. A* 512 (1–2), 30–43. [https://doi.org/10.1016/S0168-9002\(03\)01874-6](https://doi.org/10.1016/S0168-9002(03)01874-6).
- Lomheim, T.S., et al., 1990. Imaging charge-coupled device (CCD) transient response to 17 and 50 MeV proton and heavy-ion irradiation. *IEEE Trans. Nucl. Sci.* 37 (6), 1876–1885. <https://doi.org/10.1109/23.101204>.
- Ma, T.P., Dressendorfer, P.V. (Eds.), 1989. *Ionizing Radiation Effects in MOS Devices and Circuits*. Wiley.
- Marcelot, O., et al., 2015. Radiation effects in CCD on CMOS devices: first analysis of TID and DDD effects. *IEEE Trans. Nucl. Sci.* 62 (6), 2965–2970. <https://doi.org/10.1109/TNS.2015.2497405>.
- Martin-Gonthier, P., Goiffon, V., Magnan, P., 2012. In-pixel source follower transistor RTS noise behavior under ionizing radiation in CMOS image sensors. *IEEE Trans. Electron Devices* 59 (6), 1686–1692. <https://doi.org/10.1109/TED.2012.2189115>.
- Martin-Gonthier, P., et al., 2010. CMOS detectors for space applications: from R&D to operational program with large volume foundry. In: *Sensors, Systems and Next-Generation Satellites XIV*, Toulouse, France.
- Massey, R., 2010. Charge transfer inefficiency in the Hubble space telescope since servicing mission 4. *Mon. Not. R. Astron. Soc. Lett.* 409 (1), 109–113. <https://doi.org/10.1111/j.1745-3933.2010.00959.x>.
- Massey, R., et al., 2010. Pixel-based correction for charge transfer inefficiency in the Hubble space telescope advanced camera for surveys. *Mon. Not. R. Astron. Soc.* 401 (1), 371–384. <https://doi.org/10.1111/j.1365-2966.2009.15638.x>.
- Petersen, E., 2011. *Single Event Effects in Aerospace*. Wiley-IEEE Press. Available at <http://ieeexplore.ieee.org/servlet/opac?bknumber=6047596>.
- Petersen, E.L., et al., 2013. The single event revolution. *IEEE Trans. Nucl. Sci.* 60 (3), 1824–1835. <https://doi.org/10.1109/TNS.2013.2248065>.
- Pickel, J.C., et al., 2003. Radiation effects on photonic imagers-a historical perspective. *IEEE Trans. Nucl. Sci.* 50 (3), 671–688. <https://doi.org/10.1109/TNS.2003.813126>.
- Reuter, et al., 2008. Ralph: a visible/infrared imager for the New Horizons Pluto/Kuiper Belt mission. *Space Sci. Rev.* 140, 129–154. <https://doi.org/10.1007/s11214-008-9375-7>.
- Rizzolo, S., et al., 2018. Total-ionizing dose effects on charge transfer efficiency and image lag in pinned photodiode CMOS image sensors. *IEEE Trans. Nucl. Sci.* 65 (1), 84–91. <https://doi.org/10.1109/TNS.2017.2778760>.
- Rushton, J.E., et al., 2015. A CMOS TDI image sensor for earth observation. In: *Proceedings of SPIE*, p. 96160R. <https://doi.org/10.1117/12.2187952>.
- Rushton, J.E., et al., 2016. Single event effects in 0.18μm CMOS image sensors. In: *Proceedings of SPIE*, p. 99152Q. <https://doi.org/10.1117/12.2235212>.
- Sexton, F.W., 2003. Destructive single-event effects in semiconductor devices and ICs. *IEEE Trans. Nucl. Sci.* 50 (3), 603–621. <https://doi.org/10.1109/TNS.2003.813137>.
- Shaneyfelt, M.R., et al., 1998. Challenges in hardening technologies using shallow-trench isolation. *IEEE Trans. Nucl. Sci.* 45 (6), 2584–2592. <https://doi.org/10.1109/23.736501>.
- Smith, D.R., Holland, A.D., Hutchinson, I.B., 2004. Random telegraph signals in charge coupled devices. *Nuclear Instr. Meth. Physics Res. A* 530 (3), 521–535. <https://doi.org/10.1016/j.nima.2004.03.210>.
- Soman, M.R., et al., 2016. Electro-optic and radiation damage performance of the CIS115, an imaging sensor for the JANUS optical camera onboard JUICE. In: Holland, A.D.,

- Beletic, J. (Eds.), Proceedings of SPIE—The International Society for Optical Engineering, p. 991515. <https://doi.org/10.1117/12.2234290>.
- Srour, J.R., Marshall, C.J., Marshall, P.W., 2003. Review of displacement damage effects in silicon devices. *IEEE Trans. Nucl. Sci.* 50 (3), 653–670. <https://doi.org/10.1109/TNS.2003.813197>.
- Stassinopoulos, E.G., Raymond, J.P., 1988. The space radiation environment for electronics. *Proc. IEEE* 76 (11), 1423–1442. <https://doi.org/10.1109/5.90113>.
- Sze, S.M., 1981. Physics of Semiconductor Devices. Wiley.**
- Tan, J., Buttgen, B., Theuwissen, A.J.P., 2012. Analyzing the radiation degradation of 4-transistor deep submicron technology CMOS image sensors. *IEEE Sensors J.* 12 (6), 2278–2286. <https://doi.org/10.1109/JSEN.2012.2186287>.
- Uemura, T., et al., 2014. Exploring well-configurations for minimizing single event latchup. *IEEE Trans. Nucl. Sci.* 61 (6), 3282–3289. <https://doi.org/10.1109/TNS.2014.2363666>.
- Virmontois, C., et al., 2013. Dark current random telegraph signals in solid-state image sensors. *IEEE Trans. Nucl. Sci.* 60 (6), 4323–4331. <https://doi.org/10.1109/TNS.2013.2290236>.
- Virmontois, C., et al., 2014. Radiation-induced dose and single event effects in digital CMOS image sensors. *IEEE Trans. Nucl. Sci.* 61 (6), 3331–3340. <https://doi.org/10.1109/TNS.2014.2369436>.
- Wood, D., et al., 2017. Evolution and impact of defects in a p-channel CCD after cryogenic proton-irradiation. *IEEE Trans. Nucl. Sci.* 64 (11), 2814–2821. <https://doi.org/10.1109/TNS.2017.2756019>.

This page intentionally left blank

# Complementary metal-oxide-semiconductor (CMOS) sensors for high-performance scientific imaging

10

R. Turchetta  
IMASENIC, Barcelona, Spain

## 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 designing 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 focus on low-light imaging or space applications, we here review the progress in the detection of higher energy photons and particles. This chapter first briefly summarizes the principle of detection of this type of radiation in silicon, and it then reviews the development of CMOS sensors for both areas. We 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, for example, 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 concludes each of the two main sections dedicated, respectively, to the detection of charged particles and high-energy photons. The chapter finishes with a short review of further reading.

## 10.2 Detection in silicon

In this section we briefly summarize the principles of detection in silicon for high-energy radiation. Although the ways of these two types of radiation interact with silicon are different, the process can be summarized by the loss of  $\Delta E$  energy by radiation, this loss going into the creation of  $N_{\text{eh}}$  electron-hole pair, where the relation between these two quantities is set by the quantity  $W$ :

$$N_{\text{eh}} = \frac{\Delta E}{W} \quad (10.1)$$

$W$  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

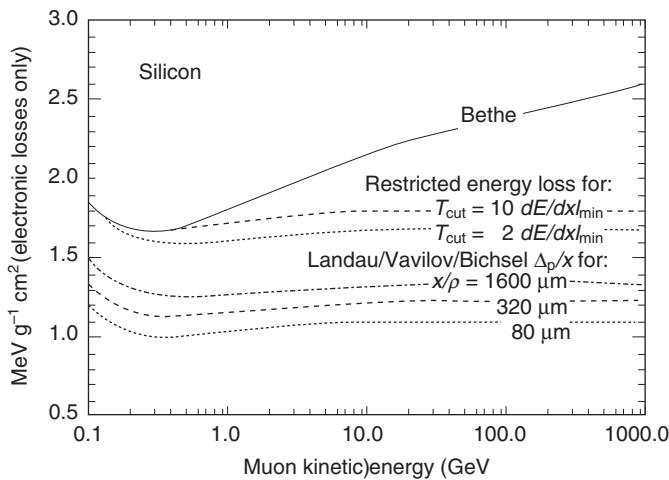
When a charged particle traverses a material, it interacts with that material and loses energy through the electromagnetic interaction. For heavy particles or low-energy electrons, the average energy loss is well described by the [Bethe \(1930\)](#) in the following equation. This is given by

$$-\left\langle \frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad (10.2)$$

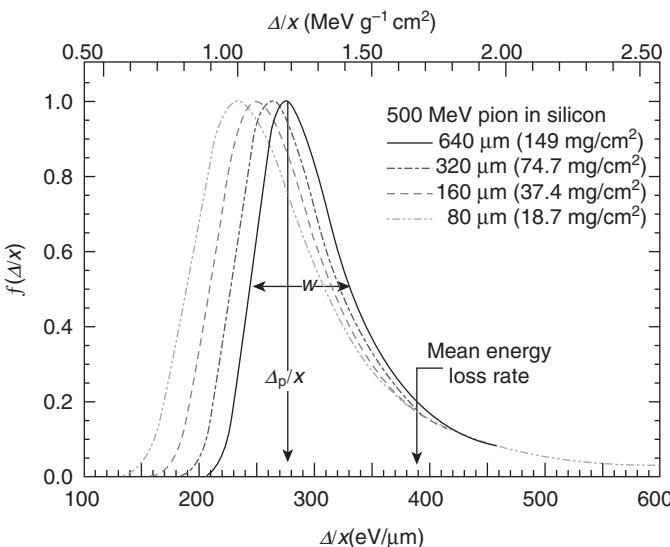
where  $e$  is the charge of the electron,  $z$  is the charge of the incident particle,  $Z$  and  $A$  are respectively, the atomic number and mass of the material,  $\beta$  and  $\gamma$  are the usual relativistic kinematic quantities,  $c$  is the speed of light in vacuum,  $m_e c^2$  is the mass 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 ionization energy loss and with  $N_A$  the Avogadro's number, and  $r_e$  the classic radius of electrons. [Fig. 10.1](#) plots Eq. (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 and 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](#),



**Fig. 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 Beringer, J., et al., 2012. Phys. Rev. D86, 010001. Copyright 2012 by the American Physical Society.



**Fig. 10.2** Energy loss distribution (straggling function) in silicon for 500 MeV pions.  
Reprinted figure from Beringer, J., et al., 2012. Phys. Rev. D86, 010001. Copyright 2012 by the American Physical Society.

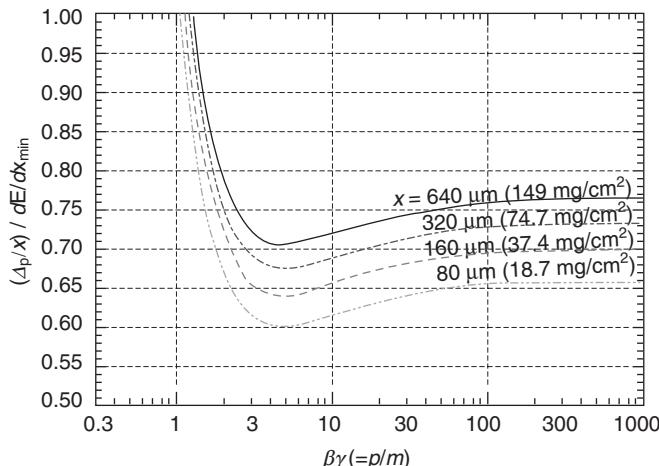
this distribution is highly skewed and so a good parameter to describe is the most probable energy loss ([Bichsel, 1988](#))

$$\Delta_p = \epsilon \left[ \ln \frac{2mc^2\beta^2\gamma^2}{I} + \ln \frac{\epsilon}{I} + j - \beta^2 - \delta(\beta\gamma) \right] \quad (10.3)$$

where  $\epsilon = (K/2) <Z/A> (x/\beta^2)$  MeV for a detector with a thickness  $x$  in g cm<sup>-2</sup> and  $j = 0.200$ .

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  $\Delta$ -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  $\mu\text{m}$ , 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 summarized in [Fig. 10.3](#), which shows the most probable energy loss as scaled



**Fig. 10.3** Most probable energy loss in silicon, scaled to the mean loss of a minimum ionizing particle, for different thicknesses of the absorbing material.

Reprinted figure from Beringer, J., et al., 2012. Phys. Rev. D86, 010001. Copyright 2012 by the American Physical Society.

to the average loss of a minimum ionizing particle. This latter quantity is equal to 388 eV/ $\mu\text{m}$  and corresponds to the creation of 164 electron-hole pairs per micron. It should be noticed that while Eq. (10.1) can be used for the energy 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  $\theta_0$

$$\theta_0 = \frac{13.6 \text{ MeV}}{\beta cp} z \sqrt{x/X_0} [1 + 0.038 \ln(x/X_0)] \quad (10.4)$$

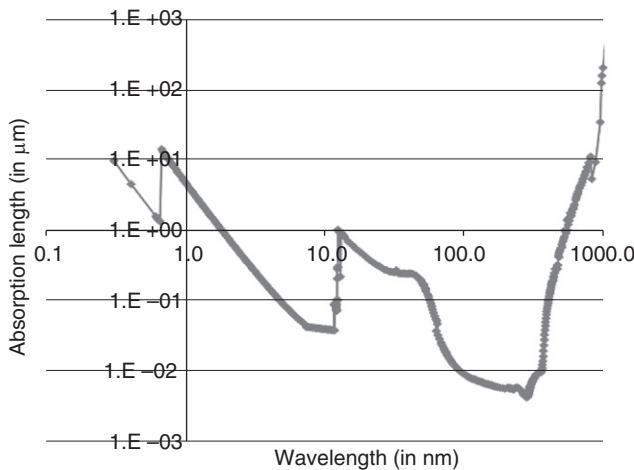
where  $X_0$  is the so-called radiation length of the material. This quantity corresponds to the length over which the electron energy is reduced to  $1/e$  of its original energy by *bremsstrahlung* (Beringer et al., 2012), a phenomenon in which the particle is decelerated and the lost energy goes 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.

Eq. (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 1100 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 nanometers, 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 behavior is shown in Fig. 10.4, which covers the



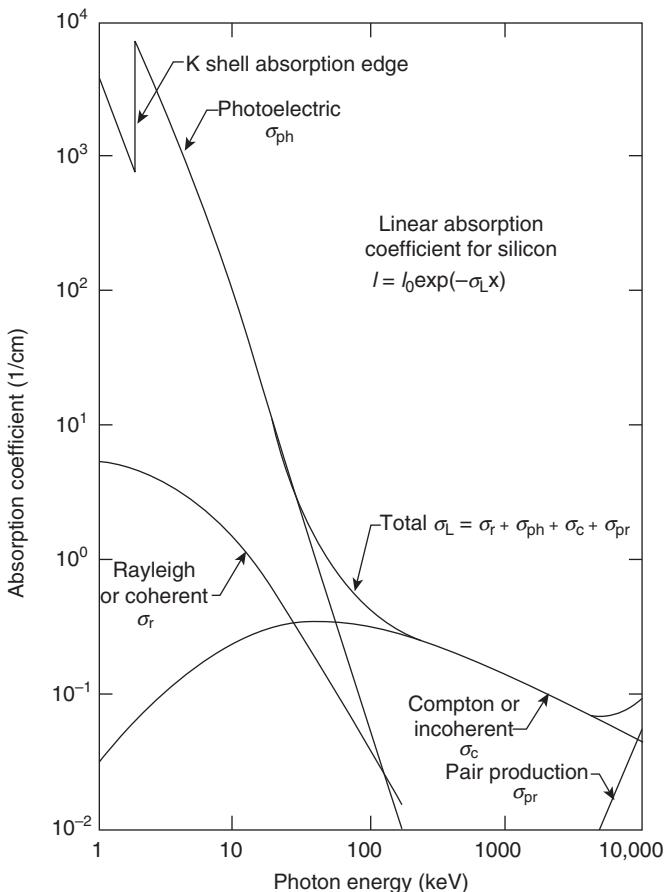
**Fig. 10.4** Absorption coefficient as a function of energy in silicon.

Adapted from Palik, E.D., 1985. Handbook of Optical Constants of Solids. Academic Press, New York.

wavelength range between about 0.2 and 1000 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 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 2 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 of 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, becomes dominant. This effect is important for the measurement of high-energy photons and particles because of their 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



**Fig. 10.5** Linear absorption coefficient in silicon.

From Spieler, H., 1998. Introduction to Radiation Detectors and Electronics. Lecture Notes—Physics 198, Spring Semester 1998—UC Berkeley.

quantum yield  $\eta$ , that is, 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

$$\eta = \frac{E}{W} \quad (10.5)$$

where  $W$  is the same constant that appears in Eq. (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 that and their magnitude is given below. 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  $N_{\text{eh}}$  of electron-hole pairs generated by one photon given by Eq. (10.1), the variance of this number is given by

$$\sigma_{\text{eh}}^2 = F * N_{\text{eh}}$$

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. Similarly, when an increasing number of photons are detected, the photon shot noise increases, and for a sufficiently high number of photons, the photon shot noise is the dominant source. This observation inspired the use of the photon transfer curve for the characterization of image sensors (Janesick, 2007; Bohndiek et al., 2008). This method was originally developed for the test of image sensors for visible light detection. If the sensor can be modeled as a linear system with an output measured in unit AU, which could be volts or ADC numbers, and with an overall conversion gain  $G$ , expressed as AU per electron, given the detection of  $N_{\text{ph}}$  visible light photons, the output of the sensor will be equal to  $G * N_{\text{ph}}$ . When the photon shot noise is dominant, the standard deviation of the output will be  $G * \sqrt{N_{\text{ph}}}$ . Then the plot of the variance of the signal with respect to the average signal will be a straight line with a slope equal to  $G$ , the conversion gain of the sensor. This curve allows extracting  $G$  and other electro-optical parameters by using only a constant, uncalibrated light source. When the quantum yield  $\eta$  is different from 1, the photon transfer method can still be applied, but this time the standard deviation will be equal to  $G * \eta * \sqrt{N_{\text{ph}}}$ . If the noise of the system is low enough to allow single photon detection, for example, in the case of the detection of X-rays, then other methods can be used. For example, radioactive sources with well-known characteristic lines can be used. In this case, a linear fit of the curve sensor output vs photon energy will deliver the overall gain of the sensor.

### 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 in between 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  $10^5 \text{ cm}^{-1}$ , corresponding to an absorption length of

<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 and 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 thallium and structured to provide some channeling to the light, thus improving the spatial resolution (Arvanitis et al., 2009). In some cases a fiber optic 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, 1962; Knoll et al., 1988), or suitable materials converting electrons into charged particles (Aoyama et al., 1992; Maneuski et al., 2008).

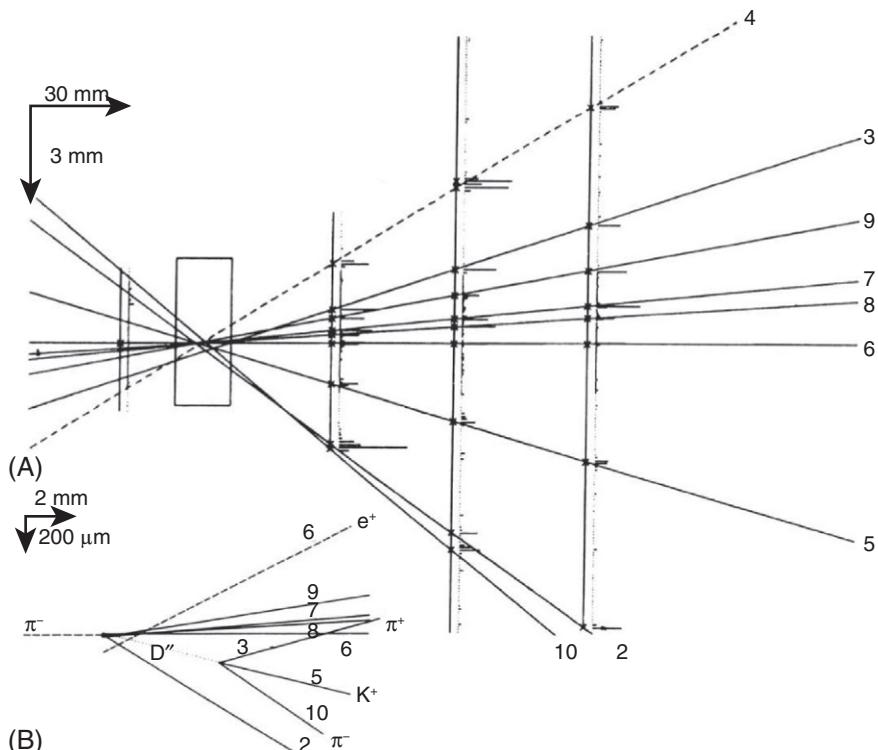
#### 10.2.4 CMOS substrates

Most of the discussions are valid for silicon in general and do not take into account the specific substrates used for CIS. The most common substrate to be used is epitaxial, typically with a resistivity of a few tens of ohm centimeter and a thickness of a few microns. The chosen thickness matches well with 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  $1 \text{ k}\Omega \text{ cm}$  as well as thickness beyond 20  $\mu\text{m}$ . 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, for example, UV or low-energy electrons, substrates need to be backthinned for backside illumination, as shown in Chapter 4. Recently CMOS sensors have been demonstrated for the detection of low-energy X-rays in the project Percival, see Correa et al. (2016) and Khromova et al. (2016).

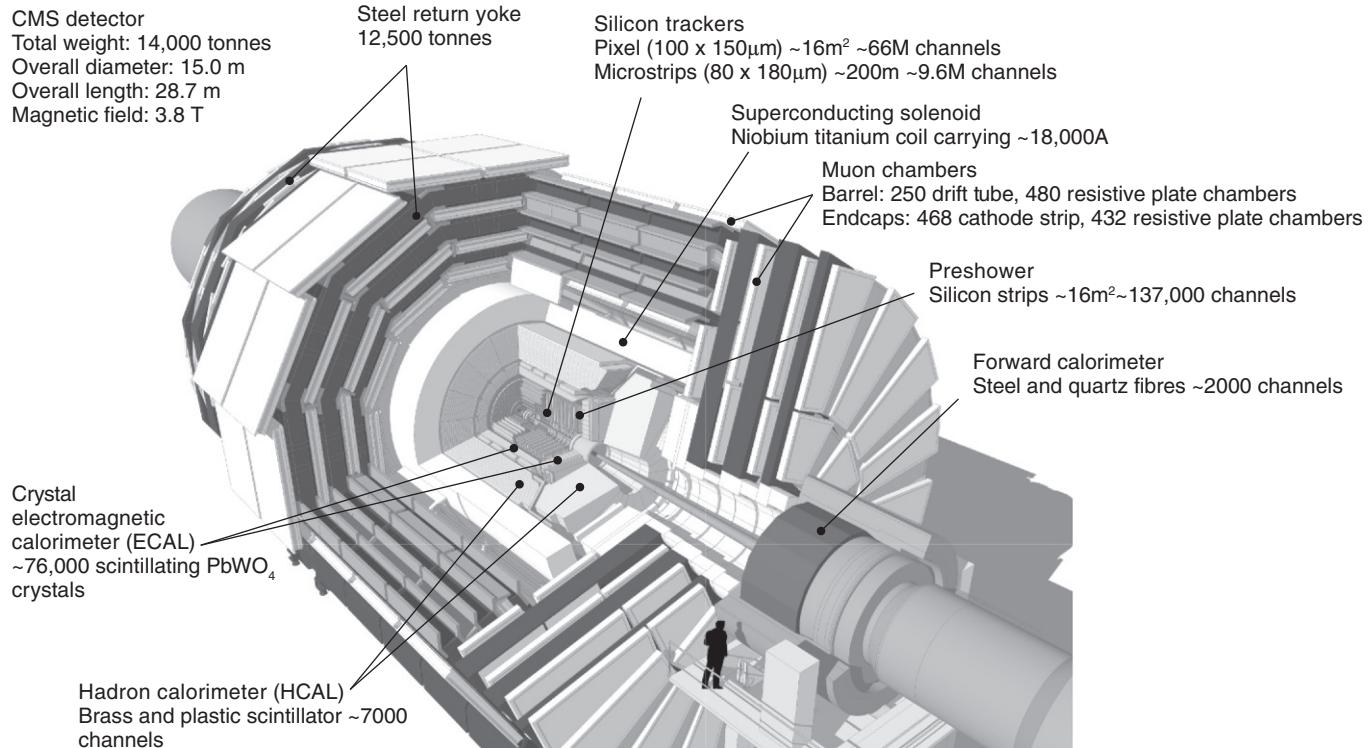
## 10.3 CMOS sensors for the detection of charged particles

### 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,



**Fig. 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). (A) Four planes of one view of the experiment. (B) Enlargement of the vertex area (see rectangle in A) (Belau et al., 1983; Bailey et al., 1985).



**Fig. 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.

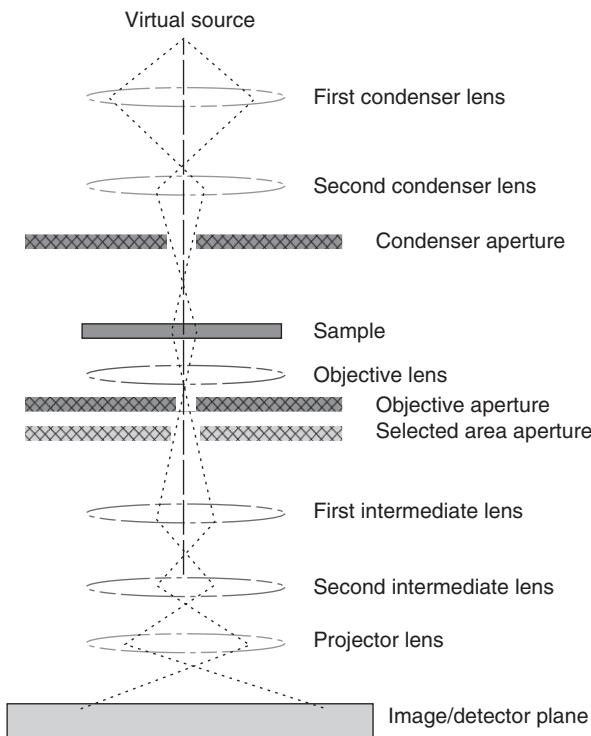
are made to circulate in a circular accelerator and then steered against each other in selected collision areas. Collisions between the elementary particles forming the two beams take place and the resulting particles need to be detected. In both fixed target and collider experiments, the experimental apparatus is composed of a multitude of detectors, with different functionalities. The interaction point are generally closer to 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 CIS 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.

### 10.3.2 Electron microscopy

Two large categories of electron microscopes exist: transmission (TEM) and scanning electron microscopes (SEM). The typical geometry of a TEM 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  $\lambda$  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 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



**Fig. 10.8** Schematic diagram of a transmission electron microscope (TEM).

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 channeling 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-mass ratio  $e/m$  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 MS 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 MS like SIMS (secondary ion mass spectrometry) or MALDI (matrix-assisted

laser desorption/ionization) which are used respectively, for the analysis of solid-state material or biological samples. Imaging time-of-flight MS requires a pixelated 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 CIS 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 radiolabeled 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  $^3\text{H}$ . 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 optimize 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).

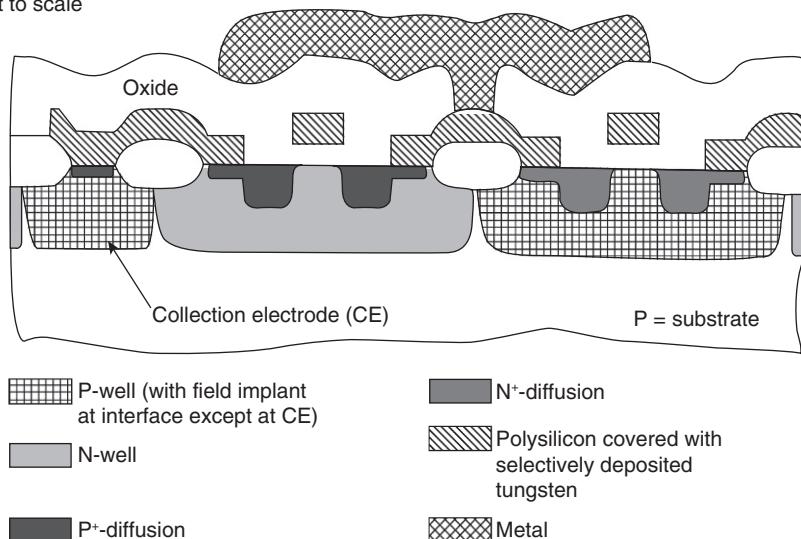
A group at IMEC in Belgium (Vanstraelen et al., 1988, 1989, 1991) also developed fully functional metal-oxide-semiconductor (MOS) transistors on high-resistivity material, obtaining low leakage currents in the pin diodes on the same substrate by having a clean process and relatively low processing temperatures. 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, b). 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, 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 back-side to create the necessary pn junction (Fig. 10.9). On the front-side, N- and P-wells were implemented 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 millimeter size with rectangular  $34 \times 125 \mu\text{m}$  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  $2.0 \mu\text{m}$  (Kenney et al., 1993; Snoeys et al., 1993).

Not to scale



**Fig. 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  $10^{12} \text{ cm}^{-3}$ . P<sup>+</sup> and N<sup>+</sup>-diffusion areas are used to create the source/drain implants of the MOS transistors.

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 vigor. 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 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  $64 \times 64$  pixels at  $25 \mu\text{m}$  pitch, called MIMOSA, meaning minimum ionizing 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  $18.4 \mu\text{m}$  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  $112 \mu\text{s}$ .

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  $\mu\text{m}$  “Opto” process. It features 928 (rows)  $\times$  960 (columns) pixels, 20.7  $\mu\text{m}$  pitch, for a total area of about  $20 \times 23 \text{ mm}^2$ . The epitaxial layer has a thickness of 15  $\mu\text{m}$  and a relatively high resistivity of  $400 \Omega \text{ cm}$ . A fast binary readout and zero suppression are also features of this sensor, whose readout time can be as short as 200  $\mu\text{s}$ . Because of the targeted radiation tolerance of  $\sim 150 \text{ kRad}$  and a few  $10^{12} n_{\text{eq}}/\text{cm}^2/\text{year}$ , radiation tolerant design techniques were used for this sensor.

STAR at RHIC is the first experiment to use CMOS monolithic active pixel sensors, and it has now been followed by ALICE, one of the four experiments at the LHC in CERN. For the upgrade of the so called inner tracking system (ITS) in ALICE (Musa et al., 2012), CMOS sensors were developed in a 180 nm technology. The R&D has just been completed and a  $10 \text{ m}^2$  detector made of monolithic CMOS sensors will be installed in the experiment during 2019–2020. Each sensor used in the upgrade covers a surface of  $3 \text{ cm} \times 1.5 \text{ cm}$ , with a pitch of 28  $\mu\text{m}$  and a total of  $512 \times 1024$  pixels. The in-pixel circuitry is capable of detecting the arrival of particles and generates a “hit” signal as a consequence (M. Mager, 2016).

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 optimization of the detector and the electronics. More complex electronics can be used to condition the signal at high speed with good noise performance and digitize the analogue information locally in the pixel. In most CIS, the detecting junction is formed by an N-doped area, for example, 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  $\mu\text{m}$  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  $48 \times 92$  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  $22\text{ e}^- \text{ rms}$ . 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 MS (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  $70\text{ }\mu\text{m}$  and the number of active devices increased to over 600. A first prototype with  $72 \times 72$  cells has already been demonstrated (John et al., 2012) and a reticle-size sensor, consisting of  $384 \times 384$  pixels has already been manufactured, is working well and is currently being characterized (Sedgwick et al., 2012). All the sensors mentioned here and featuring a deep P-well for protection were manufactured in the 180-nm CIS 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 20 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 behavior 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) and 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  $4096 \times 4096$  pixels at  $14\text{ }\mu\text{m}$  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  $5\text{ }\mu\text{m}$  pixels are integrated in a  $4\text{ k} \times 4\text{ k}$  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). This technique, where each frame only contains single electron hits that can be further analyzed, is called “electron counting.” The name is a bit misleading because it is not just about counting electrons but it is more about detecting the charge footprint they leave in the sensor so that the signal can be further analyzed to increase the spatial resolution beyond the Nyquist limit of the pixel size.

### 10.3.5 Future trends

CIS 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 CIS working in direct detection can now be purchased for TEMs. While becoming mature for using in the field, there are still several R&D programs 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  $\mu\text{m}$  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 the 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). A standard 180-nm CIS technology is used but high-resistivity substrates, in the order of  $1 \text{ k}\Omega \text{ cm}$ , 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  $5 \text{ e}^- \text{ rms}$ . A similar approach is also used in a sensor for a radiation monitor (Guerrini et al., 2012).

Both types of approaches are being considered for future upgrade of LHC experiments, see for example C. Tamma et al. (2016).

## 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 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.

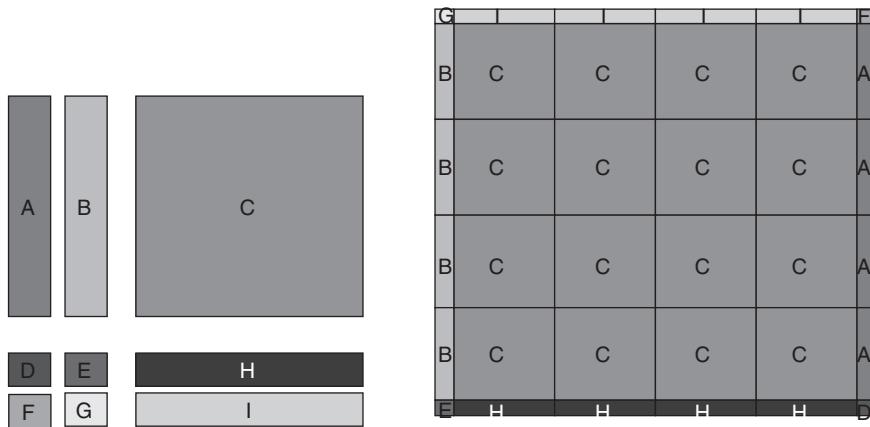
As explained above, CMOS substrates are relatively thin, so that direct detection of X-rays for energies above a few keV is inefficient. For low-energy photons, while the thickness of the substrates is adequate, the absorption from the material covering the surface can drastically reduce the performance 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  $20 \times 20$  mm, reaching  $25 \times 35$  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



**Fig. 10.10** Reticle (A) and sensor floorplan (B) for a stitched sensor. Each block on the reticle is labeled 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.

Adapted from Guerrini, N., Turchetta, R., Van Hoften, G., Henderson, R., McMullan, G., Faruqi, A.R., 2011a. A high frame rate, 16 million pixels, radiation hard CMOS sensor. J. Instrum. 6 (03), C03003. © IOP Publishing Ltd and SissaMedialab srl; doi: 10.1088/1748-0221/6/03/C03003.

known as “stitching” is used (Theuwissen et al., 1991; Kreider et al., 1995). In stitching, the design is organized in a modular way (see Fig. 10.10 showing the typical arrangement of a reticle and the overall sensor floor plan). 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 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 CIS 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  $\mu\text{m}$ , making them effective for detection of energies up to about 20 keV. The silicon drift chamber was proposed by [Gatti and Rehak \(1984\)](#). The charge generated by the radiation is drifted toward 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 nondestructive 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 CIS, the development of CMOS sensors for high-performance scientific imaging is only recently starting to yield results. CIS 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 the recent years ([Claus et al., 2001](#)). The main technology challenges for the development of such sensors are low noise, radiation hardness, and self-triggering, that is, 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 United Kingdom. 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 submicron 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 ranges 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.

## References

- Aad, G., et al., 2008. ATLAS pixel detector electronics and sensors. *J. Instrum.* 3P07007.
- Aoyama, T., Oka, Y., Honda, K., Mori, C., 1992. A neutron detector using silicon PIN photodiodes for personal neutron dosimetry. *Nucl. Inst. Methods* 314 (3), 590–594.
- Arai, Y., et al., 2006. Monolithic pixel detector in a 0.15  $\mu\text{m}$  SOI technology. In: *IEEE Nuclear Science Symposium, Conference Record*. N34–4.
- Arvanitis, C.D., Bohndiek, S.E., Blakesley, J., Olivo, A., 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), 116–126.
- Augustin, J.-E., 2004. Detector concepts for a linear collider. *Nucl. Instrum. Methods Phys. Res. 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. Instrum. Methods Phys. Res. A* 226, 56–58.
- Bak, J.F., Burenkov, A., Petersen, J.B.B., Uggerhøj, E., Møller, S.P., Siffert, P., 1987. Large departures from Landau distributions for high-energy particles traversing thin Si and Ge targets. *Nucl. Phys.* 288, 681–716.

- Ballabriga, R., Campbell, M., Heijne, E.H.M., Llopart, X., Tlustos, L., 2007. The Medipix3 prototype, a pixel readout chip working in single photon counting mode with improved spectrometric performance. *IEEE Trans. Nucl. Sci.* 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. Instrum. Methods Phys. Res. 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. In: *IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, pp. 1169–1173.
- Becksa, K.H., et al., 1997. The DELPHI pixels. *Nucl. Instrum. Meth. A* 386, 11–17.
- Becksa, K.H., et al., 1998a. Commissioning of the DELPHI pixel detector. *Nucl. Instrum. Meth. A* 418, 15–21.
- Becksa, K.H., et al., 1998b. First results from the DELPHI pixel detector. *Nucl. Instrum. Meth. A* 409, 229–231.
- Belau, E., et al., 1983. Silicon detectors with 5 μm spatial resolution for high energy particles. *Nucl. Inst. Methods Phys. Res. A* 217, 224–228.
- Beringer, J., Particle Data Group, et al., 2012. The review of particle physics. *Phys. Rev. D* 86. 010001.
- Besson, A., et al., 2011. CMOS sensors with high resistivity epitaxial layer. In: *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. Molière'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.
- Bohdiek, S., et al., 2008. Comparison of methods for estimating the conversion gain of CMOS active pixel sensors. *IEEE Sensors J.* 8 (10), 1734–1744.
- Bollinger, L.M., Thomas, G.E., Ginther, R.G., 1959. Glass scintillators for neutron detection. *Rev. Sci. Instrum.* 30, 1135–1136.
- Bollinger, L.M., Thomas, G.E., Ginther, R.G., 1962. Neutron detection with glass scintillators. *Nucl. Instrum. Meth.* 17 (1), 97–116.
- Bosiers, J., et al., 2008. Challenges and innovations in very-large CCD and CMOS imagers for professional imaging. In: *Proc. SPIE 6996, Silicon Photonics and Photonic Integrated Circuits.* 69960Z.
- Brouard, M., Campbell, E.K., Johnsen, A.J., Vallance, C., Yuen, W.H., Nomerotski, A., 2008. Velocity map imaging in time of flight mass spectrometry. *Rev. Sci. Instrum.* 79. 123115. <https://doi.org/10.1063/1.3036978>.
- Cabello, J., 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., Winter, M., 2001. Monolithic active pixel sensors for a linear collider. *Nucl. Instrum. Methods Phys. Res. 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. Nucl. Sci.* 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., Weilhammer, P., 1996. A submicron precision silicon telescope for beam test purposes. *Nucl. Inst. Methods Phys. Res. A* 372 (3), 379–384.

- Contarato, D., Denes, P., Doering, D., Joseph, J., Krieger, B., 2011. Direct detection in transmission electron microscopy with a 5  $\mu\text{m}$  pitch CMOS pixel sensor. *Nucl. Instrum. Methods Phys. Res. A* 635, 69–73.
- Correa, J., et al., 2016. Characterisation of a PERCIVAL monolithic active pixel prototype using synchrotron radiation. *J. Instrum.* 11.
- Crooks, J.P., et al., 2007. A novel CMOS monolithic active pixel sensor with analog signal processing and 100% fill factor. In: Proceedings of the IEEE Symposium on Nuclear Science. Vol. 2, pp. 931–935.
- Damerell C. 1995. Vertex Detectors: The State of the Art and Future Prospects, RAL-P-95-008. Available at <http://hepwww.rl.ac.uk/damerell/2003/Vertex-detector-95-008.pdf>.
- Degerli, Y., Deptuch, G., Fourches, N., Himmi, A., Li, Y., Lutz, P., Orsini, F., Szelezniak, M., 2005. A fast monolithic active pixel sensor with pixel-level reset noise suppression and binary outputs for charged particle detection. *IEEE Trans. Nucl. Sci.* 52, 3186–3193.
- Delpierre, P., et al., 1994. Large scale pixel detectors for DELPHI at LEP200 and ATLAS at LHC. *Nucl. Instrum. Methods Phys. Res. A* 342, 233–239.
- Deptuch, G., Winter, M., Dulinski, W., Husson, D., Turchetta, R., Riester, J.L., 2001. Simulation and measurements of charge collection in monolithic active pixel sensors. *Nucl. Instrum. Methods Phys. Res. A* 465, 92–100.
- Deptuch, G., et al., 2000. Design and testing of monolithic active pixel sensors for charged particle tracking. In: IEEE Nucl. Sci. Symp. Conf. Record. vol. 1. pp. 103–110.
- Dierickx, B., Meynants, G., Scheffer, D., 1997. Near-100% fill factor standard CMOS active pixels. In: Proc. 1997 IEEE Workshop on CCDs and Advanced Image Sensors, Bruges, Belgium.
- Dierickx, B., et al., 1993. Integration of CMOS-electronics and particle detector diodes in high-resistivity silicon-on-insulator wafers. *IEEE Trans. Nucl. Sci.* 40, 753–758.
- Downing, M., et al., 2012. Backside-illuminated, high-QE, 3e-RoN, fast 700 fps,  $1760 \times 1680$  pixels CMOS imager for AO with highly parallel readout. In: SPIE Astronomical Telescopes + Instrumentation. International Society for Optics and Photonics, p. 84530C.
- Edelstein, R.S., et al., 2011. Process integration aspects of back illuminated CMOS imagers using smart stacking<sup>TM</sup> technology with best in class direct bonding. In: Proceedings of the International Image Sensor Workshop (IISW).
- Eikenberry, E.F., Brönnimann, C., Hülsen, G., Toyokawa, H., Horisberger, R., Schmitt, B., Schulze-Briese, C., Tomizaki, T., 2003. PILATUS: a two-dimensional X-ray detector for macromolecular crystallography. *Nucl. Instrum. Methods Phys. Res. 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., Henderson, R., 2007. Electronic detectors for electron microscopy. *Curr. Opin. Struct. Biol.* 17 (5), 549–555.
- Faruqi, A.R., Henderson, R., Prydderch, M., Turchetta, R., Allport, P., Evans, A., 2005. Direct single electron detection with a CMOS detector for electron microscopy. *Nucl. Instrum. Methods Phys. Res. A* 546, 170–175.
- Faruqi, A.R., McMullan, G., 2011. Electronic detectors for electron microscopy. *Q. Rev. Biophys.* 44 (3), 357–390.
- FEI, 2009. <http://investor.fei.com/releasedetail.cfm?ReleaseID=399045>.
- Firsching, M., Takoukam Talla, P., Michel, T., Anton, G., 2008. Material resolving X-ray imaging using spectrum reconstruction with Medipix2. *Nucl. Instrum. Methods Phys. Res. A* 591 (1), 19–23.
- Fujita, J., Moritaa, S., Sakuraia, M., 1989. X-Ray Diagnostics for Fusion Plasmas. vol. 7 *Laser and Particle Beams*, pp. 483–486.

- Gatan, 2012. <http://www.gatan.com/K2/>.
- Gatti, E., Rehak, P., 1984. Semiconductor drift chamber—an application of a novel charge transport scheme. *Nucl. Inst. Methods Phys. Res. A* 225 (3), 608–614.
- Granja, C., Jakubek, J., Platkevic, M., Pospisil, S., Vykydal, Z., 2009. Detection and real time spectroscopy of charged particles with the timepix pixel detector. *AIP Conf. Proc.* 1204, 75–79.
- Greinera, L., et al., 2011. A MAPS based vertex detector for the STAR experiment at RHIC. *Nucl. Instrum. Methods Phys. Res. A* 650, 68–72.
- Guerrini, N., Turchetta, R., Van Hoften, G., Henderson, R., McMullan, G., Faruqi, A.R., 2011a. A high frame rate, 16 million pixels, radiation hard CMOS sensor. *J. Instrum.* 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. In: *Proceedings of the International Image Sensor Workshop (IISW)*.
- Guerrini, N., et al., 2012. Design and characterisation of a highly miniaturised radiation monitor. *Nucl. Instrum. Methods Phys. Res. A* 731, 154–159.
- Hara, K., Kochiyama, M., Koikea, K., Segu, T., Shinsho, K., Araib, Y., Fujitab, Y., Okiharah Miurah, M., Motoyoshii, M., 2010. Development of INTPIX and CNTPIX silicon-on-insulator monolithic pixel devices. In: *Proceedings of Science 19th International Workshop on Vertex Detectors*.
- Hartmann, F., 2012. Silicon tracking detectors in high-energy physics. *Nucl. Instrum. Methods Phys. Res. A* 666, 25–46.
- Heijne, E.H.M., Jarron, P., Olsen, A., Redaelli, N., 1988. The silicon micropattern detector: a dream? *Nucl. Instrum. Methods Phys. Res. A* 273 (2–3), 615–619.
- Heitler, W., 1949. Theory of meson production. *Rev. Mod. Phys.* 21, 113.
- Henderson, R., Unwin, P.N.T., 1975. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* 257, 28–32.
- Hoenck, M.E., et al., 2013. The DUV stability of superlattice-doped CMOS detector arrays. In: *Proceedings of the International Image Sensor Workshop (IISW)*.
- Holland, S.E., 1989a. Fabrication of detectors and transistors on high-resistivity silicon. *Nucl. Instrum. Methods Phys. Res. A* 275, 537–541.
- Holland, S.E., 1989b. An IC-compatible detector process. *IEEE Trans. Nucl. Sci.* 36, 283–289.
- Holland, S.E., Spieler, H., 1990. A monolithically integrated detector-preamplifier on high-resistivity silicon. *IEEE Trans. Nucl. Sci.* 37, 463–468.
- Hu-Guo, C., et al., 2009. CMOS pixel sensor development: a fast read-out architecture with integrated zero suppression. *J. Instrum.* 4, P04012.
- Ikeda, H., et al., 2007. Deep sub-micron FD-SOI for front-end application. *Nucl. Instrum. Methods Phys. Res. A* 579, 701–705.
- Janesick, J.R., 2007. *Photon Transfer DN → λ*. SPIE, Bellingham.
- John, J.J., et al., 2012. PImMS, a fast event-triggered monolithic pixel detector with storage of multiple timestamps. *J. Instrum.* 7, C08001.
- Kemmer, J., Lutz, G., Prechtel, U., Schuster, K., Sterzik, M., Strüder, L., Zieman, T., 1990. Experimental confirmation of a new semiconductor detector principle. *Nucl. Instrum. Methods Phys. Res. A* 288 (1), 92–98.
- Kenney, C., Parker, S., Snoeys, W., Plummer, J., Rosseel, G., Aw, C.H., 1993. Performance of a monolithic pixel detector. *Nucl. Instrum. Methods Phys. Res. A* 232, 460–467.
- Khromova, A., et al., 2016. Report on recent results of the PERCIVAL soft X-ray imager. *J. Instrum.* 11.
- Klein, C.A., 1968. Bandgap ionization energies in semiconductors. *J. Appl. Phys.* 39, 2029–2038.

- Kluge, A., et al., 2007. The ALICE silicon pixel detector. *Nucl. Instrum. Methods Phys. Res. A* 582, 728–732.
- Knoll, M., 1935. Aufladepotential und Sekundäremission elektronenbestrahlter Körper. *Z. Tech. Phys.* 16, 467–475.
- Knoll, G.F., Knoll, T.F., Henderson, T.M., 1988. Light collection in scintillation detector composites for neutron detection. *IEEE Trans. Nucl. Sci.* 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 0.20 µm FD-SOI technology. *Nucl. Instrum. Methods Phys. Res. 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., Fox, E., 2009. A wafer-scale CMOS APS imager for medical X-ray applications. In: Proceedings of the International Image Sensor Workshop (IISW).
- Kotlinski, D., 2009. Status of the CMS pixel detector. *J. Instrum.* 4, P03019.
- Kraft, P., Bergamaschi, A., Broennimann, C., Dinapoli, R., Eikenberry, E.F., Henrich, B., Schmitt, B., 2009. Performance of single-photon-counting PILATUS detector modules. *J. Synchrotron Radiat.* 16 (3), 368–375.
- Kreider, G., et al., 1995. An mK × nK modular image sensor design. In: IEDM Dig. Tech. Papers, pp. 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., Sampietro, M., 1996. Silicon drift detectors for high resolution room temperature X-ray spectroscopy. *Nucl. Instrum. Methods Phys. Res. A* 377 (2), 346–351.
- Lechner, P., Pahlke, A., Soltau, H., 2004. Novel high-resolution silicon drift detectors. *X-Ray Spectrom.* 33 (4), 256–261.
- Mager, M., 2016. ALPIDE, the monolithic active pixel sensor for the ALICE ITS upgrade. *Nucl. Instrum. Methods Phys. Res. A* 824, 434–438.
- Maneuski, D., Eklund, L., Kralik, M., Laing, A., O’Shea, V., Pospisil, S., Vykydal, Z., 2008. Evaluation of silicon monolithic APS as a neutron detector. In: IEEE Nucl. Sc. Symposium Conference Record, pp. 2383–2386.
- McMullan, G., Chen, S., Henderson, R., Faruqi, A.R., 2009c. The detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy* 109, 1126–1143.
- McMullan, G., Clark, A.T., Turchetta, R., Faruqi, A.R., 2009b. Enhanced imaging in low dose electron microscopy using electron counting. *Ultramicroscopy* 109 (12), 1411.
- McMullan, G., Faruqi, A.R., Henderson, R., Guerrini, N., Turchetta, R., Jacobs, A., Van Hoften, G., 2009a. Experimental observation of the improvement in MTF from back-thinning a CMOS direct electron detector. *Ultramicroscopy* 109, 1144–1147.
- Mendis, S., Kemeny, S.E., Fossum, E.R., 1994. CMOS active pixel image sensor. *IEEE Trans. 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. *J. Instrum.* 5, C07007.
- Ono, Y., Ishikawa, A., Yamamoto, H., Arai, Y., Tsuboyama, T., Onuki, Y., Iwata, A., Imamura, T., Ohmoto, T., 2013. Development of the pixel OR SOI detector for high energy physics experiments. *Nucl. Instrum. Methods Phys. Res. A* 731, 266–269.

- Pain, B., 2009. Backside illumination technology for SOI-CMOS image sensors. In: Proceedings of the International Image Sensor Workshop (IISW).
- Pain B., Zheng X., 2002. US Patent No. 6,380,572.
- Parker, S., 1989. A proposed VLSI pixel device for particle detection. *Nucl. Instrum. Methods Phys. Res. 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. Instrum. Methods Phys. Res. A* 582, 876–885.
- Peric, I., 2012. Active pixel sensors in high-voltage CMOS technologies for ATLAS. *J. Instrum.* 7(08). C08002.
- Perkins, D.H., 2000. Introduction to High Energy Physics. Cambridge University Press, Cambridge.
- Reshef, R., Leitner, T., Alfassi, S., Sarig, E., Golan, N., Berman, O., Fenigstein, A., Wolf, H., Hevel, G., Vilan, S., Lahav, A., 2009. Large-format medical X-ray cmos image sensor for high resolution high frame rate applications. In: Proceedings of the International Image Sensor Workshop (IISW).
- Richter, R.H., Andricek, L., Fischer, P., Heinzinger, K., Lechner, P., Lutz, G., Wermes, N., 2003. Design and technology of DEPFET pixel sensors for linear collider applications. *Nucl. Instrum. Methods Phys. Res. A* 511 (1), 250–256.
- Rizzo, G., Batignani, G., Bettarini, S., Bosi, F., Calderini, G., Cenci, R., Zoccoli, A., 2007. Recent development on triple well 130 nm CMOS MAPS with in-pixel signal processing and data sparsification capability. In: Nuclear Science Symposium Conference Record 2pp. 927–930.
- Ruska, E., 1987. The development of the electron microscope and of electron microscopy. *Bio-sci. Rep.* 7 (8), 607–629.
- Ruska, E., 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. Nucl. Sci.* 58, 2528–2536.
- Sedgwick, I., Das, D., Guerrini, N., Marsh, B., Turchetta, R., 2013. LASSENA: a 6.7 megapixel, 3-sides buttable wafer-scale CMOS sensor using a novel grid-addressing architecture. In: Proceedings of the International Image Sensor Workshop (IISW).
- Sedgwick, I., et al., 2012. PIImMS: a self-triggered, 25 ns resolution monolithic CMOS sensor for time-of-flight and imaging mass spectrometry. In: Proceedings of the IEEE NEWCAS 2012 Conference.
- Shouleh, N., Yu, Q., Smith, A.L., Jones, T.J., Tombrello, T.A., 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. Stanford University, Technical Report 92-017.
- Snoeys W., 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., Parker, S., 1993. First beam test results from a monolithic silicon pixel detector. *Nucl. Instrum. Methods Phys. Res. A* 326, 144–149.
- Stanitzki, M., et al., 2007. A tera-pixel calorimeter for the ILC. In: Proceedings of the IEEE Symposium on Nuclear Science. vol. 1, pp. 254–258.
- Takahashi, H., Yamashita, Y., Kikuchi, S., Fujita, M., Hirayama, S., Kanou, T., Hashimoto, S., Shigeta, K., Aoki, T., Momma, G., Inoue, S., 2011. A 300 mm wafer-size CMOS image

- sensor for low-light-level imaging. In: Proceedings of the International Image Sensor Workshop (IISW).
- Tamma, C., et al., 2016. The CHESS-2 prototype in AMS 0.35  $\mu\text{m}$  process: a high voltage CMOS monolithic sensor for ATLAS upgrade. In: 2016 IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop (NSS/MIC/RTSD), Strasbourg, pp. 1–5.
- 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. In: IEDM Technical Digest, pp. 167–170.
- Tietz, D., Tietz, H., Nickell, S., Baumeister, W., Plitzko, J.M., 2008. Characterization of a fiber-optically coupled 8k CCD/CMOS device. In: Proceedings of the EMC 2008 14th European Microscopy Congress, Aachen, Germany, pp. 89–90 I1, I1.3.
- Treis, J., Fischer, P., Halker, O., Harter, M., Herrmann, S., Kohrs, R., Wermes, N., 2005. DEPMOSFET active pixel sensor prototypes for the XEUS wide field imager. *IEEE Trans. Nucl. Sci.* 52 (4), 1083–1091.
- Turala, M., 2005. Silicon tracking detectors—historical overview. *Nucl. Instrum. Methods Phys. Res. 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., Winter, M., 2001. A monolithic active pixel sensor for charged particle tracking and imaging using standard VLSI CMOS technology. *Nucl. Instrum. Methods Phys. Res. A* 458, 677–689.
- Turchetta R., Villani G., Prydderch M., 2011. Accelerated Particle and High Energy Radiation Sensor. US Patent No. 7,944,012.
- Vanstraelen, G., Debusschere, I., Claeys, C., Declerck, G., 1988. New concepts for integrated solid state detector electronics. *Nucl. Instrum. Methods Phys. Res. A* 273, 625–629.
- Vanstraelen, G., Debusschere, I., Claeys, C., Declerck, G., 1989. Fully integrated CMOS pixel detector for high energy particles. *Nucl. Instrum. Methods Phys. Res. A* 275, 574–579.
- Vanstraelen, G., Debusschere, I., Claeys, C., Declerck, G., 1991. Design implications of a p-well CMOS technology for the realization of monolithic integrated pixel arrays. *Nucl. Instrum. Methods Phys. Res. 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. Instrum. Methods Phys. Res. A* 604, 370–379.
- Wrigley, C., Yang, G., Pain, B., 2000. High responsivity CMOS imager pixel implemented in SOI technology. In: Proceedings of the 2000 IEEE SOI Conference, pp. 138–139.
- Wunderer, C., et al., 2012. The PERCIVAL soft X-ray imager. In: Presented at the 11th International Conference on Synchrotron Radiation Instruments (SRI).
- Zhao, W., Ristic, G., Rowlands, J.A., 2004. X-ray imaging performance of structured cesium iodide scintillators. *Med. Phys.* 31 (9), 2594–2605.

## Further Reading

Palik, E.D., 1985. *Handbook of Optical Constants of Solids*. Academic Press, New York.

This page intentionally left blank

# CMOS-based optical time-of-flight 3D imaging and ranging

11

R. Lange<sup>a</sup>, S. Böhmer<sup>b</sup>, B. Buxbaum<sup>b</sup>

<sup>a</sup>University of Applied Sciences Bonn-Rhein-Sieg, Sankt Augustin, Germany,

<sup>b</sup>PMD Technologies AG, Siegen, Germany

## 11.1 Introduction to 3D imaging and ranging

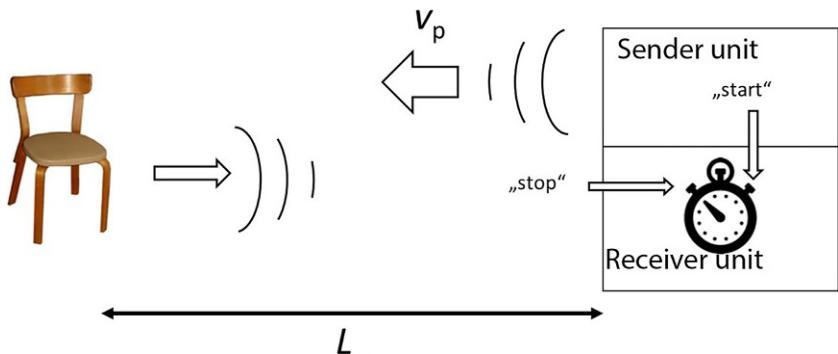
The concept of time-of-flight (ToF) incorporates several different range measurement techniques. All of these have in common, that the turn-around time of a propagating wave with known propagation velocity  $v_p$  is measured, which then represents the double distance between measuring system and target (Fig. 11.1).

$$L = \frac{1}{2} \cdot v_p \cdot t_{\text{meas}} \quad (11.1)$$

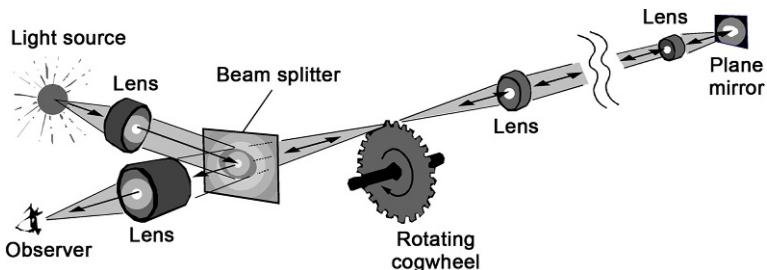
Sonar systems make use of acoustic waves, propagating with speed of sound, while radar systems (radio aircraft detection and ranging) are using electromagnetic radio waves in the range from several megahertz to 100 GHz, propagating with speed of light. Since this book is about imaging, however, we are more interested in optical ToF systems, where the carrier wave is light, predominantly in the near-infrared (NIR) spectrum, making the measurement principle more or less invisible to the human eye. While optical ToF systems are summarized with the terms lidar (light detection and ranging) or ladar (laser detection and ranging) also the term “time-of-flight” or ToF itself seems to establish as designator for such optical measurement systems. Sometimes laser scanners or single laser beam systems are called lidars in contrast to imaging ToF, but this is no universal definition and finally, all those systems are based on the same basic principle.

### 11.1.1 Speed of light

It was Galileo Galilei who is reported to be the first to imagine light as something that travels through free space rather than existing instantaneously. In the early 17th century, he suggested an experiment with two men standing in a distance of 1 km giving signs to each other and thus measuring the propagation time of light. Well, the experiment failed, but what sounds funny from today's perspective, at least lead Galileo to a remarkable conclusion: if light does not exist instantaneously, it must be very fast,



**Fig. 11.1** Time-of-flight principle. The propagating wave can either be sound (sonar), radio (radar), or light (lidar).



**Fig. 11.2** Fizeau's experimental setup used to determine the speed of light in 1849 in the suburbs of Paris.

much faster than the reaction time of men. In 1676, Roemer succeeded in measuring the speed of light by using the departures from predicted eclipse times of a Jupiter moon, strictly speaking also a ToF experiment.

Much closer to today's ToF systems was Fizeau's measuring setup from 1849, which was more precise, reliable, and reproducible than previous experiments (Fig. 11.2). He used a rotating cogwheel as a mechanical shutter for a focused light beam. This pulsed beam was projected onto a mirror, located  $>8$  km away. For an appropriate rotation speed of the cogwheel the complete light packet, which had been transmitted by a gap in the wheel was obstructed by a tooth after being reflected by the distant mirror, so that it could not be seen behind the cogwheel. As the distance and the shutter rate were known, Fizeau could calculate the speed of light. With this early ToF experiment, he obtained an astonishingly precise result of  $3.153 \times 10^8$  m/s (Hoffmann, 1997).

Today, the speed of light ( $c = \lambda \cdot \nu$ ) can be determined even more precisely, for example, by the simultaneous measurement of frequency  $\nu$  and wavelength  $\lambda$  of a

stabilized helium-neon-laser or by the frequency measurement of an electromagnetic wave in a cavity resonator (Von Breuer, 1990). Since 1983, the speed of light has been fixed by definition to  $c = 2.99792458 \times 10^8$  m/s.

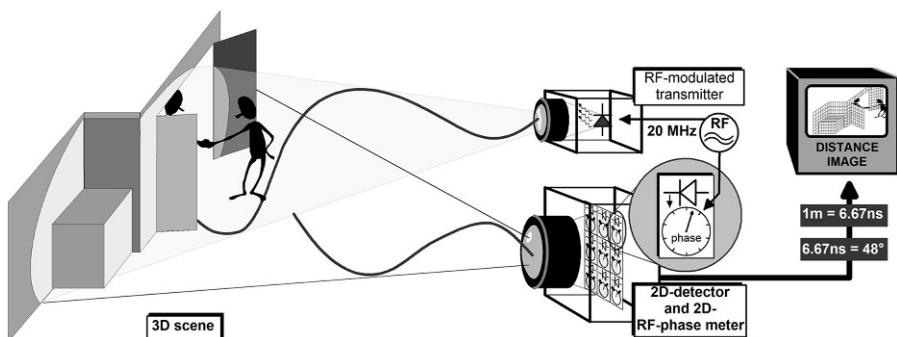
### 11.1.2 Optical ToF measurement

So with this precise knowledge of velocity of light, it is possible to measure distances, assuming one manages to realize a high precision stopwatch. According to Eq. (11.1), such a stopwatch would require roughly 7 ps accuracy for 1-mm distance resolution. Concerning the illustration of the simplified ToF setup in Fig. 11.1, it should be mentioned that in practice active light source and receiver are located very close to each other. This avoids shadowing effects and facilitates a compact setup, in fact one of the advantages of this technology, compared to all triangulation-based approaches, which require a certain baseline between the emitter and the receiver by principle.

The stringent requirements in timing accuracy for the receiver are one reason that, although already in 1903 Hulsmeyer carried out some experiments with electromagnetic radar, only in 1968 was Koechner one of the first to introduce an optical ToF ranging system (Koechner, 1968). Another problem was the lack of high-power (laser-) light sources. These first ToF ranging systems were actually only one-dimensional (1D) systems, measuring the distance to a point. In order to gain full three-dimensional (3D) information, the laser beam needed to be scanned over the scene in two directions. This requires high accuracy laser scanners, which are mostly bulky and sensitive to vibrations.

### 11.1.3 3D ToF ranging

Rather than scanning a laser beam and serially acquiring the range data point wise, one can illuminate the entire scene with modulated light in order to perform a 3D measurement, as illustrated in Fig. 11.3. This, however, necessitates the use of a



**Fig. 11.3** Principle of an imaging ToF camera. Rather than scanning a modulated beam, the whole scene is instantaneously illuminated with modulated light, which is imaged to a ToF image sensor (Lange, 2000).

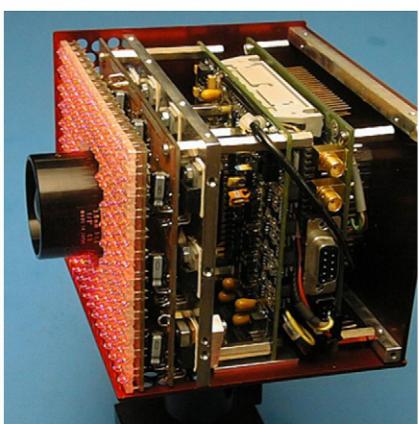
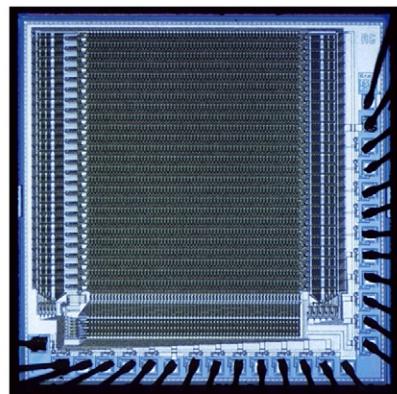
two-dimensional (2D)-electro-optical demodulator and detector to measure the distances of several thousands of points of the observed scene in parallel.

In the 1980s and 1990s, the first attempts in realizing imaging ToF followed the idea of demodulating the active light in the optical domain before it is detected. This can be achieved using large aperture optical modulators, such as Kerr- or Pockels cells (Schwarte, 1995) or microchannel plates (MCP) (Kappner, 1986). For detection, off-the-shelf standard image sensors can be used. There is no need for the imagers themselves to be very fast because they only integrate DC images containing the phase information (distance information, respectively), amplitude information and offset information (background image). An MCP can be switched on and off in <500 ps enabling high demodulation speeds. However, all these large aperture optical (de)modulators suffer from high price and a need for elaborate electronics, since voltage amplitudes in the range of some kilovolts are necessary. Nevertheless, the first commercially available imaging ToF camera was based on the large aperture optical demodulators. This was the Z-Cam of Israeli company 3DV Systems, introduced to the market in 1999 (Iddan, 2001).

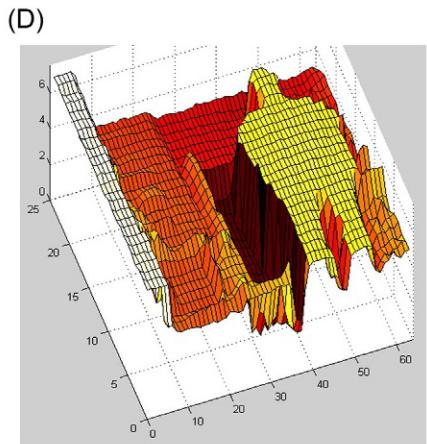
This chapter now is about imaging ToF measurement. The developments in charge-coupled device (CCD) and complementary metal-oxide-semiconductor (CMOS) imaging allowed realizing smart pixel sensors with special functionalities. In case of ToF image sensors, also called demodulation pixel arrays, each pixel contains the complete demodulating receiver of the TOF system. Instead of realizing the demodulation by large aperture optical modulators, the optical demodulation is realized in the smart pixels themselves. The actual correlation process between the received light and the synchronous reference signal, which is the essential process to measure the time delay, is performed in each pixel, immediately after the light is detected. Bandwidth requirements are more or less uncritical for all other electronic components within the following signal processing chain in the range camera system, since the high-speed demodulation process is already done within the detection device.

Introduced in 1995, the lock-in-CCD was the first such smart-pixel array intended to be used for imaging ToF. While basic functionality of demodulation could be demonstrated with the 80- $\mu\text{m}$  pitch pixels (17% fill factor), the device still suffered from low demodulation bandwidth (100 kHz) (Spirig, 1995). Two years later, in 1997, the idea of photonic mixer device (PMD) was introduced. Performing demodulation directly in the light-sensitive pixel area and working with differential push-pull mode was the key to overcome bandwidth limitations and to allow smaller pixel pitches, maintaining reasonable fill factors (Schwarte, 1997a, b). In 1999, the first all solid-state imaging ToF camera based on a ToF imager was then introduced to the public. Key figures of this setup are summarized in Fig. 11.4 (Lange, 1999, 2000).

From this starting point, ToF imaging has then evolved year-by-year until today. Advances in CMOS image sensor (CIS) technology, like microlenses, backside illumination (BSI), field enhancements with special EPI materials, one to one lead to a performance boost for the new ToF generations. But not only CIS technology is the only performance driver, but also developments of light-emitting diodes (LEDs), lasers, and vertical cavity surface-emitting lasers (VCSELs), offering higher optical



Resolution:	64x25
Pixel pitch:	21 x 65 $\mu\text{m}$
Optical fill factor (Pixel):	22%
Demodulation contrast:	40%, 20 MHz, 630nm
Technology node:	2 $\mu\text{m}$ CMOS/CCD process, buried channel option
Lens:	F1.0/2.6mm
Frame rate:	10Hz
Illumination:	900mW, 160 LEDs, 630nm / 820nm
Distance Noise	9cm @ 7.5m, 70% target, 60° viewing angle

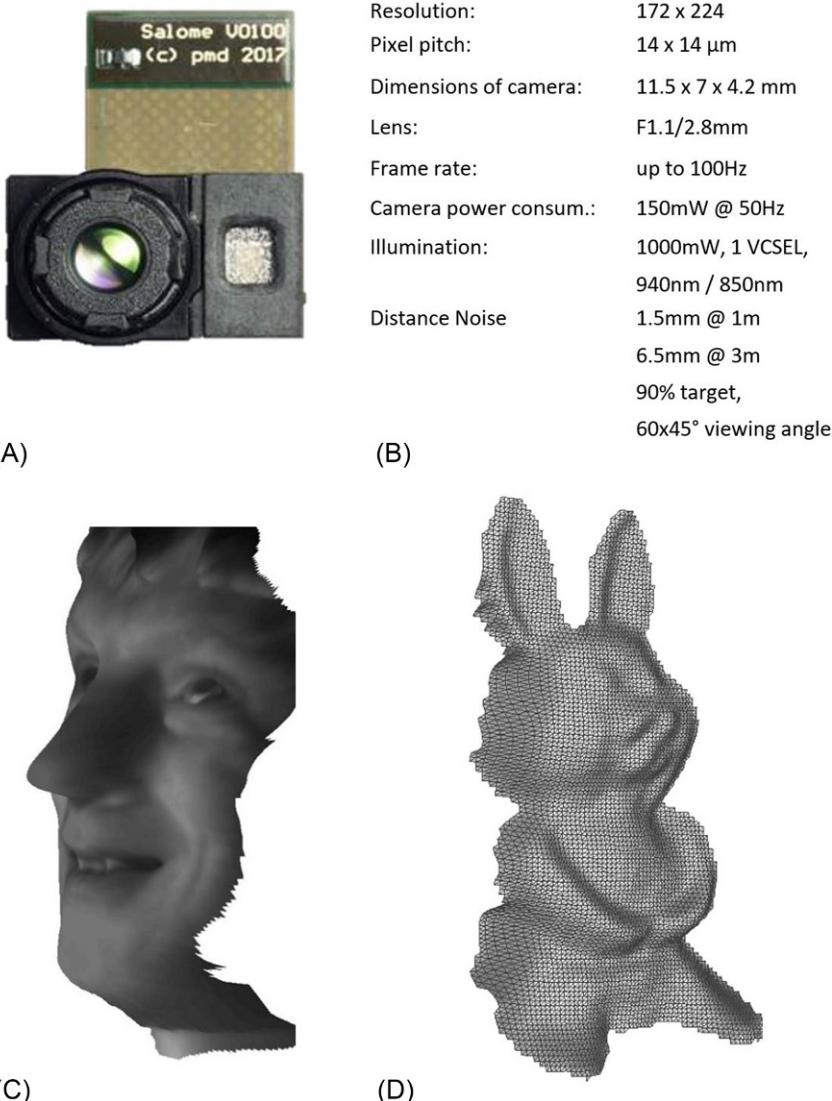


(E)

**Fig. 11.4** Early time-of-flight cameras. (A) First prototype ([Lange, 1999](#)), (B) 64 x 25 pixel time-of-flight imager, (C) world's first 3D ToF camera prototype ([Lange, 2000](#)), (D) key figures, and (E) 3D depth maps.

power at smaller sizes and higher modulation frequencies. Today's VCSELs can easily be modulated beyond 100 MHz with several watts of output power at a footprint in mm range.

[Fig. 11.5](#) illustrates form factor and performance of a state-of-the-art integrated ToF solution of the year 2018, nearly 20 years later than [Fig. 11.4](#).



**Fig. 11.5** Time-of-flight camera module for integration in mobile devices in the year 2018 (pmdtechnologies ag). (A) Image of 3D camera module (size 11.5 x 7.0 x 4.2 mm), (B) key figures, (C) depth map of face acquisition, NIR 2D information overlay, and (D) depth map of toy rabbit.

## 11.2 Concept and design considerations for ToF cameras

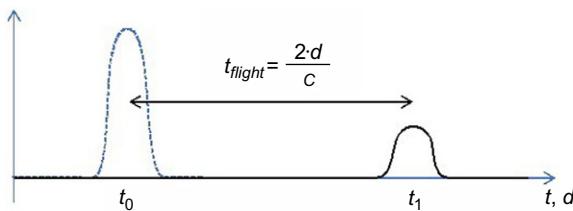
### 11.2.1 Modulation concepts

The basic ToF working principle has been explained in [Section 11.1](#). In fact, there are different methods to choose for the temporal shape of modulated light. [Fig. 11.6](#) illustrates the most prominent methods.

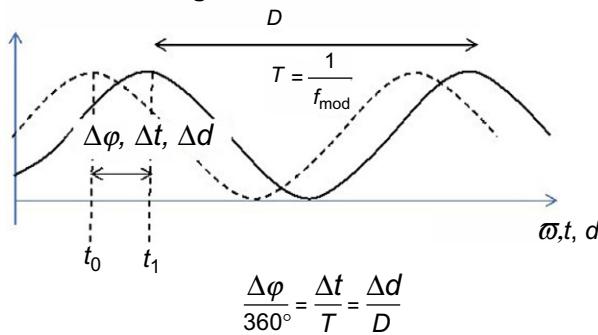
#### 11.2.1.1 Direct ToF

The most obvious modulation scheme is chosen for pulsed ToF. This method is also called direct ToF (or dToF) since the turn-around-time is measured directly. Like described in [Section 11.1](#), for this method very short light pulses are emitted and time measurement is performed in the receiving channel. According to Eq. (11.1), a measurement accuracy of 1 mm requires a timing accuracy of <6.7 ps. This modulation method, therefore, puts high accuracy and bandwidth requirements on both modulation and demodulation timing units. The advantage of pulsed modulation lies in its outstanding robustness against background light. This is because integration times can typically be chosen very short (typically less than a microsecond), short enough that nearly no significant amount of ambient light is collected to occupy part of the sensors dynamic range (DR). On the other hand, eye safety considerations prevent the system designer to choose too much pulse energy for most applications. Thus,

#### Pulsed time-of-flight measurement



#### am-cw time-of-flight measurement



**Fig. 11.6** Comparison between direct (pulsed) ToF and indirect (am-cw) ToF.

the sensor needs to be highly sensitive. Today, dToF systems do not yet play an important role, at least not for imaging ToF. This may change in the near future, especially for applications dealing with strong ambient light conditions.

### 11.2.1.2 Indirect ToF

The most prominent modulation scheme for imaging ToF is the amplitude-modulated continuous wave modulation (am-cw). In this method (see Fig. 11.4) rather than sending a single pulse, the light is continuously switched on and off with a coded HF waveform [e.g., sine, rectangular, or pseudo-noise (PN)], typically of some 10–100 MHz. In this case, the phase-delay between emitted and received (and thus delayed) optical signal is measured in the receiving channel, which again is directly proportional to the time delay and thus directly proportional to the distance. Since for this method, the turn-around-time of light is not measured directly but determined by measuring a phase delay, it is called indirect ToF.

Just the plain am-cw method can cause ambiguous measurement results. One modulation period  $T$  ( $1/f_{\text{mod}}$ ) propagating at speed of light  $c$  forms a wave in the position space with position period  $D$ :

$$D = c \cdot T = \frac{c}{f_{\text{mod}}} \quad (11.2)$$

Thus, the measured phase represents a target distance of

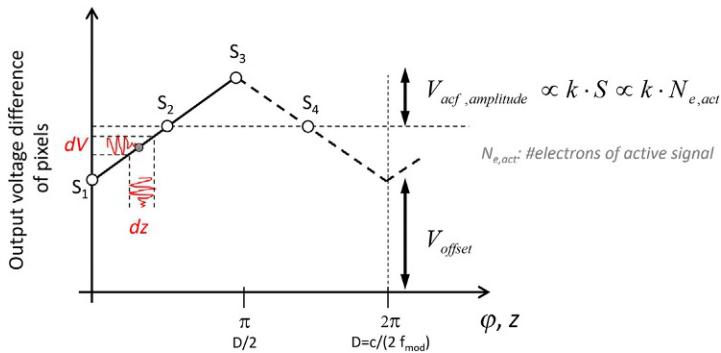
$$dz = \frac{1}{2} \cdot D \cdot \frac{d\phi}{360^\circ} = \frac{1}{2} \cdot \frac{c}{f_{\text{mod}}} \cdot \frac{d\phi}{360^\circ} \quad (11.3)$$

Like in Eq. (11.1), the factor  $\frac{1}{2}$  corrects for the fact that the light needs to travel the distance twice. The modulation frequency is a key figure for indirect ToF methods since it translates the measured phase into a measured distance. It should be chosen high for accurate results and is in practice limited to bandwidth constraints in both sender and receiving channel. Assuming a modulation frequency of 100 MHz, the nonambiguous range  $D$  results in 1.5 m and a 1-mm accuracy requires 0.24 degree of phase accuracy.

It should be mentioned that radar engineering offers a variety of methods to overcome ambiguity limitations of the am-cw modulation, like applying different modulation frequencies, a lower frequency for the coarse measurement and a high frequency for the precise measurement.

### 11.2.2 Performance and limitations

In this chapter, we will discuss the limitations of the am-cw ToF technique. We will see in Section 11.4 that most of the demodulation pixels perform in the electro-optical domain a correlation between the received optical signal and an electrical reference signal. This electrical reference has a defined phase delay with respect to the



**Fig. 11.7** Correlation function for am-cw square-wave modulation for a phase delay of  $\pi$ .  $S_i$  are the respective pixel output values for four consecutive phase measurements with demodulation offset of 0,  $\pi/2$ ,  $\pi$ , and  $3\pi/2$ .

transmitted optical signal. As known from communication technology, the correlation of a sine-function with a sine function yields again a sine while the correlation of a square wave with a square wave yields a triangle-function. Fig. 11.7 illustrates the ideal correlation result of a demodulation pixel (i.e., the difference signal of the two terminals A and B) with respect to the phase offset between electrical demodulation (reference) signal and transmitted optical modulation. This shall be the starting point for the following accuracy discussions.

The sketched case represents the pixel outputs  $S_1$  to  $S_4$  for a target distance corresponding to a phase delay of  $\pi$ . The measured points on the correlation function  $S_1$  to  $S_4$  represent the output values of one pixel. These values are acquired and digitized sequentially with four separate phase images, each with a different demodulation offset of 0,  $\pi/2$ ,  $\pi$ , and  $3\pi/2$  (see Section 11.4). The sampling points  $S_i$  are then used to reconstruct the triangle correlation function, whose maximum extremum ( $\pi$  in this case) represents the measured phase.

The amplitude of the correlation function ( $k^*S$ ) depends on the effective contrast  $k$  (considering pixel demodulation contrast and modulation contrast of light source) and the signal strength  $S$  the pixel receives, which is in fact proportional to the number of generated photons by active light.

The effective phase measurement uncertainty can be derived from the correlation function by taking the absolute value of the slope of the function itself (which is identical for all four sampling points). We will show below how to determine the uncertainty of the pixel output  $dV$  depending on the sensor and scene properties. The resulting phase uncertainty can then directly be read from the correlation function:

$$\frac{d\phi}{dV} = \frac{\pi/2}{k \cdot S} \quad (11.4)$$

Assigning  $dV$  as noise value  $N$  and taking Eqs. (11.2) and (11.3) yields:

$$dz = \frac{D}{4} \cdot \frac{1}{k \cdot \frac{S}{N}} = \frac{c}{8 \cdot f_{\text{mod}}} \cdot \frac{1}{k \cdot \frac{S}{N}} \quad (11.5)$$

This equation gives the corresponding distance noise as a function of the signal-to-noise ratio (SNR) of the acquisition for each of the measurements  $S_i$ . Considering that each of the measurements  $S_i$  contributes to the final result, the final noise value decreases with square root of the total number of measurements  $N_{\text{phase}}$  (4 in the illustrated example).

$$dz = \frac{c}{8 \cdot \sqrt{N_{\text{phase}}}} \cdot \frac{1}{f_{\text{mod}}} \cdot \frac{1}{k \cdot \frac{S}{N}} \quad (11.6)$$

This is the key equation describing the limits of am-cw ToF systems. In the following sections, we will complete this theory by describing how the measurement noise  $N$  is calculated and how we get the measurement signal value  $S$  as a function of the scene and system settings. Finally, it shall be mentioned, that the number of phase shifts is not necessarily 4. As long as the Nyquist-Shannon sampling theorem is obeyed, any number can be chosen, that is, the smallest number of required phase shifts is 2.

### 11.2.2.1 Noise sources

There are mainly three independent noise sources for a properly set up am-cw ToF system which can be summed up geometrically. The first noise source is (1) the read noise of the pixel. This can be described as all the present noise in the absence of any photonic signal. Read noise can be well approximated by the pixel's kTC noise for the example of a source follower pixel readout. The other noise sources are (2) the shot noise of acquired active photoelectrons and (3) the shot noise of acquired ambient light electrons.

Noise sources like ADC-noise, dark current shot noise, or any further noise sources are neglected here, which is valid as long as the overall system is properly designed. Shot noise is described in deep detail in [Janesick \(2007\)](#). Shot noise occurs for all quantized parts, like photons and electrons but also raindrops for example. It goes back to the fact that for very small quantities of let us say electrons, in a certain time the electron may either be there or not, it cannot be divided any further. We will see later, that shot noise is one important physical limitation of am-cw ToF systems. Since it is calculated by the square root of the respective quantity, we will transfer all other signal ( $S$ ) and noise values ( $N$ ) to the charge domain in the following. This will make things easier. We can assume the pixel dark noise as equivalent number of noise electrons:

$$N_{\text{pixel,read}} = \frac{\sqrt{kTC}}{q} \quad (11.7)$$

where  $k$  is Boltzmann constant,  $T$  is temperature in Kelvin,  $C$  is the pixel's charge storage capacitance, and  $q$  is the elementary charge of an electron. Eq. (11.7) is a simplification of the pixel's complete noise behavior, neglecting, for example, the different noise contributions of reset and hold steps. For a more detailed discussion, see Frey (2007).

For the shot noise contributions, we assume:

$$N_{\text{shot,active}} = \sqrt{Num_{e,\text{active}}} \quad (11.8)$$

$$N_{\text{shot,ambient}} = \sqrt{Num_{e,\text{ambient}}} \quad (11.9)$$

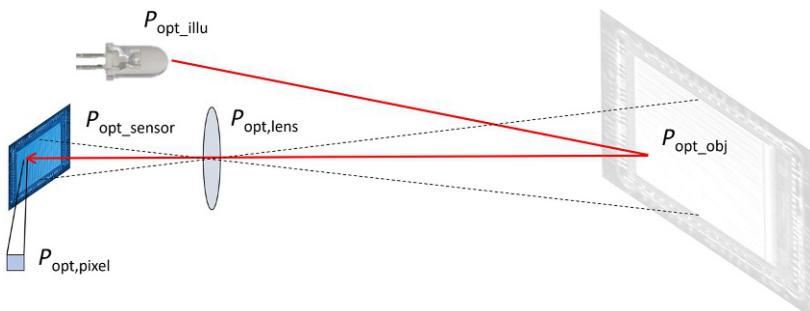
with  $Num_{e,\text{active}}$  and  $Num_{e,\text{ambient}}$  being the respective numbers of integrated electrons generated by active (modulated) or ambient light. Finally, we can calculate the total pixel noise as

$$\begin{aligned} N_{ToF} &= \sqrt{(N_{\text{pixel,read}})^2 + (N_{\text{shot,active}})^2 + (N_{\text{shot,ambient}})^2} \\ N_{ToF} &= \sqrt{\frac{kTC}{q^2} + Num_{e,\text{active}} + Num_{e,\text{ambient}}} \end{aligned} \quad (11.10)$$

We should emphasize that in this noise value also the measurement conditions (target distance, reflectivity, and ambient light) are included, since they directly influence the absolute values of the shot noise values.

### 11.2.2.2 Power budget

After discussion of pixel noise, we will now address how to calculate the number of electrons  $Num_e$  acquired in each pixel, as a function of illumination, target properties (distance and reflectivity), lens properties, and image sensor properties and settings.  $Num_e$  corresponds to the Signal  $S$  in Eq. (11.6) and is required for calculation of the photoelectron shot noise in Eqs. (11.8) and (11.9). Fig. 11.8 illustrates the key elements of the optical power budget.



**Fig. 11.8** Simplified optical power budget for an imaging system with active illumination.

For simplification, we make the following assumptions: (1) perfectly homogeneous illumination in all of the field of view (FoV) of the imager, that is, no light is lost outside this FoV; (2) the target is defined as a perfect Lambert reflector with homogenous reflectivity; and (3) the lens is ideal, that is, no transmission losses nor relative illumination losses take place, thus the illumination is just defined by  $F\#$ .

The number of electrons acquired per pixel first depends on: the optical power received on a particular pixel ( $P_{\text{opt,pixel}}$ ), the effective pixel sensitivity in A/W ( $S_{\text{sens,eff}}$ ) including the optical fill-factor of the pixel as well as the material's optical sensitivity, the light-induced charge collection or integration time ( $t_{\text{int}}$ ), and the elementary charge ( $q$ ):

$$Num_{e,\text{pixel}} = P_{\text{opt,pixel}} \cdot S_{\text{sens,eff}} \cdot t_{\text{int}} \cdot \frac{1}{q} \quad (11.11)$$

The optical power per pixel is a fraction of the total optical power impinging the whole image sensor:

$$P_{\text{opt,pixel}} = \frac{P_{\text{opt,sensor}}}{\#\text{pixel}} = \frac{P_{\text{opt,lens}}}{\#\text{pixel}} \quad (11.12)$$

With above simplifications, the total optical power entering the entrance pupil of the lens ( $P_{\text{opt,lens}}$ ) shall reach the image sensor ( $P_{\text{opt,sensor}}$ ). For the assumed lambert reflector,  $P_{\text{opt,lens}}$  is given by the lens parameters (focal length  $f$ ,  $F$ -number  $F\#$ ), the target distance  $z$  and the totally reflected optical power on the object  $P_{\text{obj}}$ :

$$P_{\text{opt,lens}} = P_{\text{opt,obj}} \cdot \left(\frac{f}{z}\right)^2 \cdot \left(\frac{1}{2 \cdot F\#}\right)^2 \quad (11.13)$$

Finally, for the assumed perfectly shaped beam profile,  $P_{\text{opt,obj}}$  is just the optical output power of the light source projection  $P_{\text{opt,illu}}$  attenuated by the reflectivity  $\rho$  of the target, that is, both are identical for a reflectivity of 1.

$$P_{\text{opt,obj}} = \rho \cdot P_{\text{opt,illu}} \quad (11.14)$$

Summarized, we get the total number of electrons integrated in a single pixel:

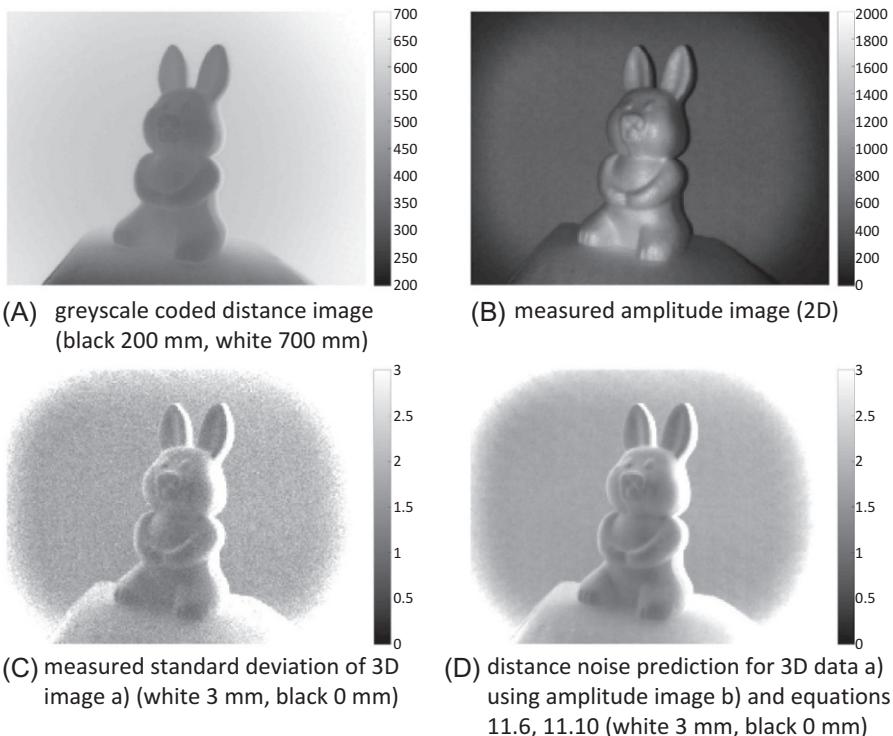
$$Num_{e,\text{pixel}} = \frac{P_{\text{opt,illu}}}{\#\text{pixel}} \cdot \rho \cdot \left(\frac{f}{2 \cdot z \cdot F\#}\right)^2 \cdot S_{\text{sens,eff}} \cdot t_{\text{int}} \cdot \frac{1}{q} \quad (11.15)$$

### 11.2.2.3 Discussion

With Eqs. (11.6), (11.10), and (11.15), we can now precisely predict the noise behavior of a ToF system depending on the key influence factors:

1. pixel performance (effective sensitivity, demodulation contrast),
2. lens properties ( $F\#$ , focal length/viewing angle),
3. target properties (distance and reflectivity),
4. ambient light conditions, and
5. system settings (modulation frequency, integration time, and optical power of the active illumination).

Not only allow these equations to execute a reliable system simulation in the very early system design stage, but also they allow (especially Eqs. 11.6 and 11.10) to deliver a reliability or confidence value for each pixel with every distance measurement. Thus, the following processing chain may decide to what extend to trust the measurement or not. Fig. 11.9 illustrates the power of the noise prediction based on the above derived equations. In addition, it shows the 2D reflectivity (b/w) data which ToF systems always deliver on top of the 3D point cloud. Note that for the illustrated measurement, the illumination was located on the left side of the camera leading to small shadowed regions on the right-hand side of the object (resulting in light shadow on the right side of the object, Fig. 11.9B and increased noise in these regions Fig. 11.9C and D).



**Fig. 11.9** Time-of-flight 3D data (A), time-of-flight 2D amplitude data (B) and comparison of real measured noise (C) with performance prediction (D).

Taking a closer look, considering some extreme boundary conditions allows for simplifications in the algorithms and gives a deeper insight in the limitations and performance characteristics of ToF.

**1. Regular condition: active illumination photon shot noise is the dominant noise limitation**

For the first case, we assume that the ambient light shot noise and pixel dark noise are not being dominant. In this case, Eq. (11.10) simplifies to.

$$N_{\text{ToF}} = N_{\text{shot,active}} = \sqrt{N_{\text{Num}_{\text{e,active}}}} \quad (11.16)$$

So, the SNR in Eq. (11.6) is then simply given by the square root of the total number of integrated photoelectrons coming from the active illumination:

$$\frac{S}{N} \Big|_{\text{regular}} = \frac{N_{\text{Num}_{\text{e,active}}}}{\sqrt{N_{\text{Num}_{\text{e,active}}}}} = \sqrt{N_{\text{Num}_{\text{e,active}}}} \quad (11.17)$$

**2. Sunlight condition: ambient light shot noise is the dominant noise limitation**

Now, ambient light shot noise is assumed to be dominant. This is typically the case if the measurement is performed under full sunlight, here we get.

$$N_{\text{ToF}} = N_{\text{shot,ambient}} = \sqrt{N_{\text{Num}_{\text{e,ambient}}}} \quad (11.18)$$

$$\frac{S}{N} \Big|_{\text{sunlight}} = \frac{N_{\text{Num}_{\text{e,active}}}}{\sqrt{N_{\text{Num}_{\text{e,ambient}}}}} \quad (11.19)$$

Applying some maths and introducing optical power densities (in W/m<sup>2</sup>), that is,  $P'_{\text{opt,illu}}$  and  $P'_{\text{opt,ambient}}$  yields an interesting notation:

$$\begin{aligned} \frac{S}{N} \Big|_{\text{sunlight}} &= \sqrt{\frac{P'_{\text{opt,illu}}}{P'_{\text{opt,ambient}}}} \cdot \sqrt{N_{\text{Num}_{\text{e,active}}}} \\ &= \sqrt{\frac{P'_{\text{opt,illu}}}{P'_{\text{opt,ambient}}}} \cdot \frac{S}{N} \Big|_{\text{regular}} \end{aligned} \quad (11.20)$$

In other words, for strong ambient light conditions, the square root of ratio of ambient light power to active light power on the target directly gives the factor of performance decrease compared to the dark case. Eq. (11.18) also makes clear that for close range situations, as long as the active optical power density is higher than the ambient light power density, distance noise can be similar to the noise present in dark conditions. However, for longer distances (active density decreases with  $z^2$ , while ambient light density remains constant), the distance noise gets higher compared to darkness.

At this point, it should be mentioned that optical filters are usually used to reduce the influence of ambient light on the measurement results. Above equations only refer to the optical power of ambient light in the spectral transfer window of the optical filter.

### **3. Dark condition—long range: active illumination shot noise is the dominant noise limitation**

We get a special situation if we want to determine how the system behaves without ambient light at very long distances. These are the conditions where a low pixel dark noise is important:

$$N_{\text{ToF}} = N_{\text{pixel,dark}} = \frac{\sqrt{kTC}}{q} \quad (11.21)$$

We get for the SNR:

$$\left| \frac{S}{N} \right|_{\text{dark, longrange}} = \frac{\text{Num}_{\text{e,active}}}{\frac{\sqrt{kTC}}{q}} = \text{const} \cdot \text{Num}_{\text{e,active}} \quad (11.22)$$

What is interesting about this special condition is that the SNR here linearly depends on the optical power of the active illumination, whereas in the other conditions SNR varies with the square root of optical power.

#### **11.2.3 2D images of a 3D ToF camera**

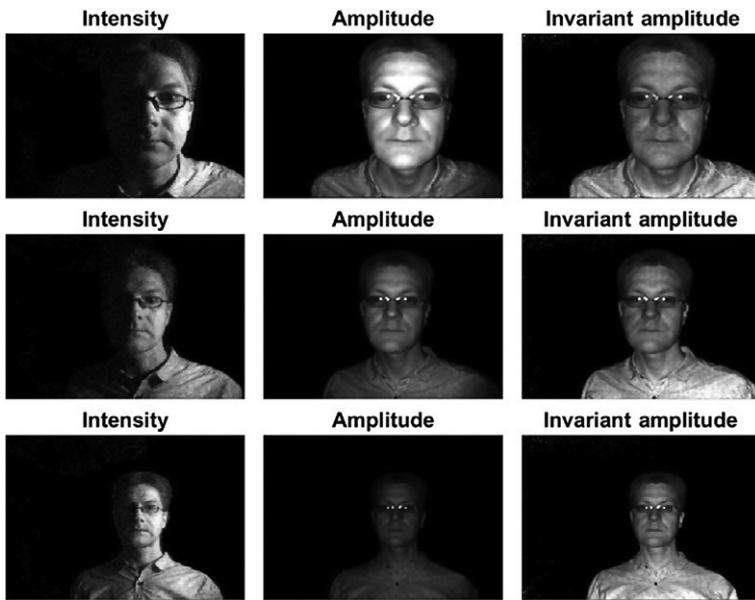
Of course, the ToF imager can also be used for conventional 2D imaging applications. Taking a closer look, it is two different 2D images that a ToF system can deliver: (1) an ambient light b/w image and (2) a special active light b/w image.

Operated without active illumination, it delivers a b/w image that will depend on the spectral bandwidth of the demodulation pixel and the optical filter, if present. Since usually the ToF system is set up to be sensitive in the NIR emission range of the modulated illumination, this ambient light b/w image is mostly also a NIR b/w image. This image can be used for classical image processing applications. It can, however, also be used in power saving modes to abandon the idle operation mode and start the active illumination when an external event is detected.

The active light b/w image has a special property. Namely, it is not the superposition of the ambient light and the active illumination (like for conventional imagers) but it solely represents the reflectance of the active illumination. Thus, it delivers a reflectance image which is completely independent from ambient illumination. This property makes 2D image processing very stable, since the image is no longer influenced by ambient lighting, which is very often not constant overtime or may, for example, cause unwanted shadows.

Finally, since the distance of the target is also measured, the active light b/w image can additionally be corrected for the natural  $1/z^2$  intensity decrease resulting in a true reflectivity measurement that is independent from ambient light conditions and also from the target's distance to the camera. All these properties are illustrated in Fig. 11.10.

Note that the ambient light image only shows the right side of the face, since the left side is shadowed. For the longer distances (lower line), the native amplitude image



**Fig. 11.10** Different types of 2D NIR images acquired with an am-cw ToF camera. *Left column*: ambient light image. *Middle column*: native active illumination b/w image. *Right column*: distance-corrected true reflectivity b/w image.

only offers a dark representation, while the correction with distance data delivers true reflectivity data, independent from ambient light and distance.

#### 11.2.4 Key components

Only few components are required to build a complete 3D ToF system. The main reason for that is that the main system tasks can be realized in the system-on-a-chip (SoC) part of the ToF image sensor as long as it is fabricated in a state-of-the art CMOS process. According to Eq. (11.6), the product  $f_{\text{mod}} \cdot k \cdot \text{SNR}$  needs to be optimized for optimum system performance. Based on these requirements, Fig. 11.11 sketches the impact of the respective key components on the system noise.

##### 11.2.4.1 Image sensor and demodulation pixels

The performance of the demodulation pixels does have huge impact on the overall system performance. A high system accuracy requires high demodulation contrast at high frequencies. At the same time, the effective sensitivity should be as high as possible, especially in the near IR spectral range. Since most systems shall be invisible to the users a wavelength of at least 850 nm is required, while 940 nm is mostly preferred. Effective sensitivity includes high optical fill factors and high material sensitivity. Micro-lenses and BSI are welcome technology steps to achieve the sensitivity goals.

	ToF image sensor	Receiving optics (lens and filter)	Modulated illumination incl. beam forming optics
Modulation frequency $f_{\text{mod}}$	++		++
Contrast $k$ (mod. and demod.)	++		++
Signal strength $S$ ( $\text{Num}_{e,\text{active}}$ )	+	+	++
Ambient light shot noise $N_{\text{shot,ambient}}$		++	
Pixel dark noise $N_{\text{pixel,dark}}$	+		
Lens			
Optical filter			
ToF image sensor			
Beam forming optics			
Modulated light source (VCSEL, EEL or LED)			

The diagram illustrates the optical path of a 3D time-of-flight system. A modulated light source (VCSEL, EEL or LED) emits a beam that is shaped by beam-forming optics and focused by a lens. The beam passes through an optical filter and is directed onto a ToF image sensor. The sensor captures the signal, which is then processed by receiving optics (lens and filter) and sent to a ToF image sensor. The system is shown emitting a beam towards a stylized Earth.

**Fig. 11.11** Key components of an imaging 3D time-of-flight system and their influence on the distance noise.

### Contrast and sensitivity—Figure of merit

When optimizing pixels and CIS technology, it often needs to be decided what is more important: demodulation contrast or sensitivity? The question can be answered by consulting Eqs. (11.6), (11.17), (11.20), and (11.22), respectively, and the answer depends on the application requirements. In all cases, the goal is to maximize the product  $k \cdot \text{SNR}$ . For most applications, the contrast is more important than sensitivity. This

is because not only signal strength linearly scales with sensitivity, but also the shot noise terms increase with the square root of sensitivity [figure of merit:  $k \cdot \text{sqrt}(\text{sensitivity})$ ]. Only if the pixel's dark noise becomes dominant noise source (long range in darkness—use case), sensitivity is of equal importance to contrast (figure of merit:  $k \cdot \text{sensitivity}$ ). In this case, low pixel dark noise is preferred.

## Pixel size

In [Section 11.2.2](#), we discussed that range and accuracy performance of ToF systems directly scale with the number of active photons each pixel captures. From that perspective, the system designer would always like to ask for more sensitive (larger) pixels. For mobile device applications, however, the overall system build size is of highest priority. This is why, like for conventional image sensors, pixel pitches are asked to be smaller and smaller. For lidar applications, however, distance range and measurement performance are more important, allowing larger pixel pitches, with higher overall sensitivity and DR.

### 11.2.4.2 Receiving optics: Lens and filter

The lens of a ToF system is primarily responsible for the signal strength that reaches the sensor. Due to the importance to capture as many active photons as possible, large aperture (low  $F\#$ ) is essential. For equal performance over the entire ToF imager, also relative illumination has to be high over the complete working angles. Since usually, ToF pixel is larger than conventional RGB pixels, sharpness requirements are usually more relaxed.

Another difference, compared to RGB lenses is the need for very low stray light behavior of the lens. For conventional imaging applications, stray light can cause a reduction in image contrast, mostly barely recognizable. In ToF systems, however, stray light not only carries false intensity but also a false distance information. As an example, a bright target in the close range imaged directly next to a dark target in the long range, can easily result in signal strength difference of two to three orders of magnitude. Now stray light of the bright target superimposed on the dark target will obviously lead to measurement errors on the dark target. A proper ToF lens design solves this problem, which in the early days of ToF cameras (using conventional off-the-shelves machine vision lenses) has often been tried to be corrected by a separate calibration for dark and for bright targets (without much success). Finally, anti-reflective coating (ARC) should be applied to ideally all lens surfaces, and it should be considered to match this coating to the wavelength of the active light source.

For every imaging system with active illumination, ambient light usually limits the system performance, because it occupies (significant) part of the sensors DR and, like discussed above ([Eq. 11.20](#)), it contributes shot noise and thus increases measurement noise, in fact the most important side effect for ToF systems. Ambient light is effectively reduced by means of optical filters, where interference filters are superior over organic filters, since they offer higher transmission, sharper spectral edges, and lower thickness.

### 11.2.4.3 Active illumination and beam-forming optics

According to the overview in Fig. 11.9, the modulated active illumination is of comparable importance to the ToF imager itself. The relevant light sources all belong to the category of semiconductor light sources, namely LEDs, edge-emitting semiconductor lasers (EEL), and VCSEL. The ideal light source should offer high optical output power without high conversion losses, it should allow high modulation frequencies with good modulation contrast, should offer a narrow spectral bandwidth, high-temperature range, should be eye safe and cost effective and finally, it should illuminate only the field of view of the ToF image sensor.

LEDs have especially been used in the early days of ToF cameras. LEDs are relatively cost effective and offer today the highest conversion efficiencies. Also, temperature range is superior to the laser sources, while it is more difficult to drive LEDs and they have reduced modulation bandwidth (typically 20–30 MHz are possible with today's devices). This is why today LEDs are mainly used for applications with extended temperature requirements (e.g., automotive) or long-range applications working with moderate modulation frequencies up to 20 MHz.

Edge-emitting lasers do not play an important role yet. While both modulation and spectral bandwidth are excellent, temperature range is limited, cost is still high compared to LED and VCSEL and finally, the system needs to deal with speckles.

VCSELs are currently the most important light source for ToF cameras. They are relatively cost effective due to their planar GaAs manufacturing process which allows testing and yield improvements already during manufacturing. The main advantages over LEDs are (1) modulation bandwidth up to several hundred megahertz, (2) factor 100 smaller spectral width allowing narrow band spectral system layout (thus reducing ambient light), and (3) excellent beam-forming capabilities. Compared to EELs, the VCSELs are smaller in size, easier to handle, cheaper, and speckle free.

Beam-forming optics is usually a diffractive optical element (DOE) or engineered diffuser. This allows large degrees of freedom when defining the beam profile. Not only can the beam shape be ideally adapted to the imager FoV, but also it is possible to compensate for relative illumination losses of the lens, if this is required for the application. LEDs in contrast, mostly emit significant amount of illumination to regions, which are not covered by the imager's FoV, which again can lead to unwanted secondary effects (like stray light or multipath interference).

### 11.2.5 Calibration of ToF cameras

Since the accuracy of am-cw ToF systems mainly lies in the timing circuits, which are barely influenced by ambient conditions, calibration is relatively easy. The most important calibration step is the zero-point calibration, since the control of light source and demodulation signal is always connected with a certain offset, which can easily be measured and compensated.

Also, it should be mentioned that the perfect triangular shape of the correlation function discussed in Section 11.2.2 is only true for the assumption of perfect rectangular modulation and demodulation signals. In praxis, however, all systems are

bandwidth limited and the signal edges are usually washed out. This leads to a non-linearity in the real correlation function which again results in systematic deviations at certain distances.

As long as the used components involved in modulation and demodulation have low tolerances, these distance-dependent offsets can be globally calibrated for examples with a look-up-table (LUT). For even higher accuracies, system individual LUTs can be measured in an initial calibration step.

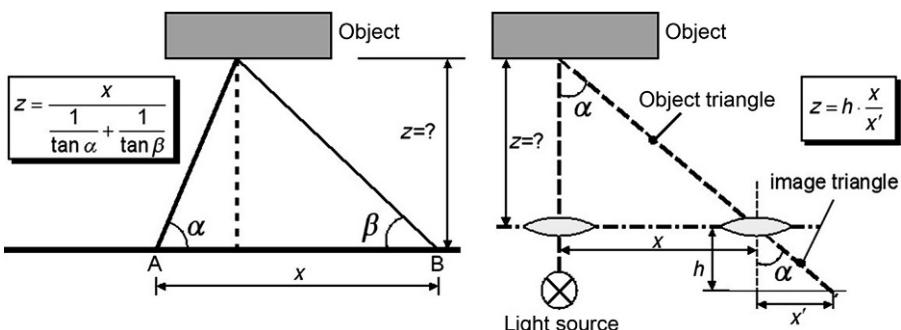
Finally, each pixel just measures a distance to the target. Usually, we are interested in a metric point cloud ( $x, y, z$ ). For this purpose, the directional vectors for each pixel must be known. These can be determined by calibrating the lens distortions, again globally or system individually, for better accuracy.

The calibration of ToF sensors is long-term stable and can be performed with few easy steps. Especially, there is no special requirement for elaborate mechanical alignment and stability.

## 11.3 Comparison of ToF with triangulation-based approaches

In the context of digital 3D image sensing, the most important measurement principles competing with ToF sensors are stereo vision (active or passive) and the structured light method. Both are based on the triangulation approach, a geometric method, where the target point is one point of a triangle whose two remaining points are known parts of the measurement system. The distance of the target can then be determined by measuring the triangle's angles or the triangulation base, like illustrated in Fig. 11.12.

Passive triangulation, the underlying principle of passive stereo, relies on observing the same point from two different sites A and B of known distance  $x$  and measuring the viewing angles  $\alpha$  and  $\beta$  with respect to the base AB. The observed point's distance  $z$  can then be calculated using the following equation:



**Fig. 11.12** The triangulation method passive (left) and active (right) (Lange, 2000).

$$z = \frac{x}{\frac{1}{\tan \alpha} + \frac{1}{\tan \beta}} \quad (11.23)$$

Since each point to measure must be identified from both viewing positions unambiguously, passive triangulation techniques require a scene with high contrast. Stereo vision uses two cameras to observe the scene from different angles. Using 2D correlation, typical object features are found and compared in both images. From the position of each feature's centroid in both separate images, the angles  $\alpha$  and  $\beta$  can be deduced and the distance can be calculated. The accuracy of stereo vision is mainly influenced by the size of triangulation base  $x$  (required to be large for good accuracy) as well as the precise knowledge and stability of the camera's orientation under all environmental conditions.

Active triangulation is used for structure light systems. Here, an active light source is used to project a point to the scene (in case of structured light, a multitude of points), which is observed by an image sensor. Rather than measuring angles directly, active triangulation is based on the similarity of triangles, the object triangle and the image triangle, which is fully defined by the optical axis of the imaging device, the focal length  $h$  of the system and the position of the point projection  $x'$  on the detector. With knowledge of the displacement  $x$  of the light source from the imaging device, the distance  $z$  of the target can be determined as.

$$z = h \cdot \frac{x}{x'} \quad (11.24)$$

For a good distance resolution  $\delta z$ , small absolute distances  $z$ , a large triangulation base  $x$ , and a good local detector resolution  $\delta x'$  are required.  $\delta z$  estimates to.

$$\delta z = \frac{1}{h} \cdot \frac{z^2}{x} \cdot \delta x' \quad (11.25)$$

While for short distances structured light systems can achieve high accuracies, the precision decrease with the square of the distance to sense ( $\delta z \propto z^2$ ) limits the area of application of this technology. It should be pointed out, that this  $z^2$  uncertainty behavior should not be mixed up with the  $1/z^2$  intensity decrease of the projected points, which comes on top.

**Table 11.1** gives a qualitative comparison of typical properties of the discussed 3D imaging principles. What is often misunderstood is the fact, that the real lateral 3D resolution (i.e., the effective number of 3D points), is only identical to the imager's pixel count for ToF systems. Marketing divisions of stereo or structured light-based systems try to sell to customers a higher lateral resolution, identical to the pixel counts of the imagers used. This can obviously not be the case. For structured light, resolution cannot be higher than the number of projected dots and for stereo vision the resolution depends on scene features, which must be oversampled by the image sensor to be detected and is thus far lower than the image sensor's pixel count.

**Table 11.1** Qualitative comparison and discussion of 3D imaging techniques.

	<b>Time-of-Flight</b>	<b>Stereo vision (passive)</b>	<b>Stereo vision (active)</b>	<b>Structured light</b>
Principle	Time-measurement	Geometric measurement	Geometric measurement	Geometric measurement
System size	+	O	O	O
Close range	+	++	++	++
Long range	+	-	-	-
Outdoor performance	O	+	-	-
Computational load	+	-	O	O
Effective resolution	+ (identical with # pixels)	O (depending on scene features)	O (depending on illumination features)	O (<=#projected dots)
Calibration Pro's	+ <ul style="list-style-type: none"> <li>- Easy and long-time stable calibration</li> <li>- Compact systems-no triangulation base required</li> <li>- Long and short range possible</li> <li>- Excellent performance in darkness</li> <li>- Dense point cloud</li> </ul>	- <ul style="list-style-type: none"> <li>- No illumination required</li> <li>- Works well under bright sunlight</li> </ul>	- <ul style="list-style-type: none"> <li>- Combination of active and passive systems</li> <li>- Close range accuracy</li> </ul>	- <ul style="list-style-type: none"> <li>- Close range accuracy</li> </ul>

Con's	<ul style="list-style-type: none"> <li>- Special image sensor required</li> <li>- Modulated illumination required</li> <li>- Eye safety standards to be considered, if lasers/VCSELs are used</li> </ul>	<ul style="list-style-type: none"> <li>- Ambient light required. No functionality in darkness</li> <li>- High contrast scenes required</li> <li>- Fragmented, incomplete point cloud</li> <li>- Triangulation base required</li> <li>- <math>z^2</math> accuracy decrease</li> <li>- Elaborate calibration</li> <li>- Tough mechanical requirements</li> </ul>	<ul style="list-style-type: none"> <li>- Special structured illumination required</li> <li>- Not suited for long range</li> <li>- Fragmented, incomplete point cloud</li> <li>- Triangulation base required</li> <li>- <math>z^2</math> accuracy decrease</li> <li>- Elaborate calibration</li> <li>- Tough mechanical requirements</li> </ul>	<ul style="list-style-type: none"> <li>- Special structured illumination required</li> <li>- Not suited for long range</li> <li>- Fragmented, incomplete point cloud</li> <li>- Triangulation base required</li> <li>- <math>z^2</math> accuracy decrease</li> <li>- Elaborate calibration</li> <li>- Tough mechanical requirements</li> </ul>
-------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## 11.4 CMOS ToF image sensors

### 11.4.1 Basics of the cw-iToF approach

In [Section 11.1](#), we introduced the first realization of imaging ToF by means of demodulation. A light source is intensity modulated at a high frequency, usually between 10 and 100 MHz. Objects within the illuminated scene reflect the light back to the camera which has lens to project the light onto the receiving sensor. Compared to the emitted signal, the received one is attenuated by a factor  $k$  due to: optical losses, phase shifted by  $\varphi$  due to the round trip delay  $\tau$  of the light signal, and is superposed with an offset  $P_B$  which is caused by uncorrelated light. In case of sinusoidal modulation, the received signal can be described by

$$P_r(t, \tau) = P_B(t) + k \cdot P_A \cdot [1 + \cos(2\pi f_m(t - \tau))]. \quad (11.26)$$

With the phase delay  $\varphi$  calculated by

$$\varphi = -2\pi f_m \tau. \quad (11.27)$$

Knowing the phase delay  $\varphi$  one can easily calculate the distance  $d$  using the following equation:

$$d = \frac{c}{2f_m} \cdot \frac{\varphi}{2\pi}. \quad (11.28)$$

with  $f_m$  being the modulation frequency and  $c$  being the speed-of-light. Obviously, the phase delay can only be nonambiguously expressed within  $[0, 2\pi]$  which results in the non-ambiguity rage defined as

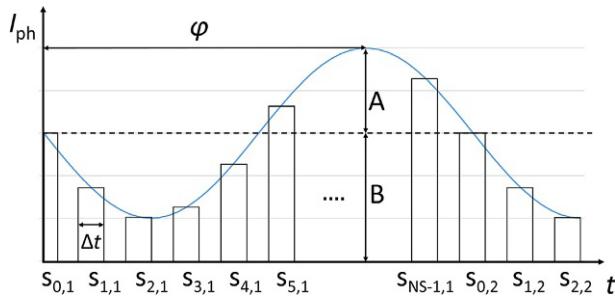
$$D = \frac{c}{2f_m} \quad (11.29)$$

To extract the phase delay of the received signal, two different approaches can be used as explanation model: demodulation by sampling and demodulation by correlation. Both approaches will be introduced in the following subsections and it will be shown that finally both can be described by the very same equations.

#### Demodulation using sampling theory

This method performs a sampling of the received optical signal described by Eq. [\(11.26\)](#), considering  $N_S$  equally spaced samples per period,  $\Delta$  being the sampling time. The sampling procedure is illustrated in [Fig. 11.13](#).

The sampling is repeated over several periods to increase the SNR of each sample. This is of specific importance as it directly translates into the precision of the distance measurement by means of error propagation, as it was derived in [Section 11.2](#). Following this approach, each sample can be expressed as



**Fig. 11.13** Synchronous sampling of the reflected sinusoidal intensity-modulated waveform.

$$S_n = \sum_{l=0}^{m-1} s_{n,l} \quad (11.30)$$

where  $s_{n,l}$  is the  $n$ th sample of the  $l$ th period and  $m = T/T_m$  is the total number of modulation periods. Each sample can be expressed as a convolution of the received signal with a rectangle function (*rect*)

$$S_n = m \cdot K \left[ \text{rect} \left( \frac{t - n \frac{T_m}{N_s}}{\Delta t} \right) * P_r(t, \tau) \right] \quad (11.31)$$

where  $K$  is a general scaling factor. The reconstruction of a periodic signal based on a minimum number of sampling points per period is possible following the theory of discrete Fourier transform (DFT). For sinusoidal modulation only the first harmonic is of special interest, the amplitude  $A$  and phase  $\varphi$  of which can be computed as (Lange, 2000)

$$A = \frac{2}{N_s} \sqrt{\left[ \sum_{n=0}^{N_s-1} S_n \cos \left( 2\pi \frac{n}{N_s} \right) \right]^2 + \left[ \sum_{n=0}^{N_s-1} S_n \sin \left( 2\pi \frac{n}{N_s} \right) \right]^2} \quad (11.32)$$

$$\varphi = -\arctan \left[ \frac{\sum_{n=0}^{N_s-1} S_n \sin \left( 2\pi \frac{n}{N_s} \right)}{\sum_{n=0}^{N_s-1} S_n \cos \left( 2\pi \frac{n}{N_s} \right)} \right] \quad (11.33)$$

The offset  $B$  is expressed as

$$B = \frac{1}{N_s} \sum_{n=0}^{N_s-1} S_n. \quad (11.34)$$

The amplitude is of special interest because its SNR directly relates to the precision of the distance measurement (Lange, 2000). Obviously, in case of no background light, the amplitude  $A$  is expected to be equal to the offset  $B$  component. The process of natural sampling by means of a short-time integration using a rectangular window, as expressed by Eq. (11.31) results in a low-pass effect which attenuates the amplitude of the harmonics. This attenuation is expressed by the demodulation efficiency ( $dme$ ) or demodulation contrast which is defined as

$$dme = \frac{A}{B} \quad (11.35)$$

Beside the low-pass effect resulting from the sampling process, other non-idealities can reduce the  $dme$  further. According to signal theory, natural sampling, as defined by Eq. (11.30), can be expressed as convolution with a *rect* function which means a multiplication with the *sinc* function in the frequency domain. This gives us the upper bound of the  $dme$  to be expressed as

$$dme_{max} = \text{sinc}(\pi f_m \Delta t) := \frac{\sin(\pi f_m \Delta t)}{\pi f_m \Delta t} \quad (11.36)$$

Following the Nyquist theorem, three samples are sufficient to reconstruct a sinusoidal signal. In practice, implementations usually use four samples as this is more convenient to implement and allow for the inherent compensation of non-idealities like mismatches between the receiving channels by the reconstruction algorithm. The equations to calculate the three unknowns of the signal with four samples simplify to

$$A = \frac{\sqrt{(S_3 - S_1)^2 + (S_2 - S_0)^2}}{2} \quad (11.37)$$

$$\varphi = \arctan\left(\frac{S_3 - S_1}{S_2 - S_0}\right) \quad (11.38)$$

$$B = \frac{S_0 + S_1 + S_2 + S_3}{4} \quad (11.39)$$

Depending on how many of the four samples, a demodulation pixel can acquire with one measurement cycle, we distinguish between 1, 2, or 4-tap pixels. If a pixel has just 1 or 2 integration notes, the optical signal has to be subsequently and accordingly sampled. Today, pixels are usually operated as 1-tap. Even so, they incorporate two integration nodes. In this mode of operation, the difference of the two integration notes represents one sample. Again, this is to compensate for any mismatches between the receiving channels.

Using four samples per period satisfies the Nyquist theorem for an ideal sinusoidal modulation as the relation between the sampling frequency and the modulation

frequency is  $f_s = 4f_m$ . As soon as any higher-order harmonics than  $2f_m$  are included in the modulation signal, the phase measurement will include systematic errors.

The process of sampling the received signal can be expressed as convolution with a *rect* function followed by ideal sampling with an infinite Dirac-pulse train (Büttgen, 2007).

$$x(t) = \left[ P_r(t, \tau) * \text{rect}\left(\frac{t}{\Delta t}\right) \right] \cdot \sum_{-\infty}^{\infty} \delta\left(t - l \frac{T_m}{4}\right) \quad (11.40)$$

In frequency domain, the convolution or integration, respectively, results in a low-pass effect which attenuates the harmonics. The multiplication with the Dirac comb which corresponds to a convolution in frequency domain leads to a shift of the frequency spectrum by  $4f_m$ . As a result, the odd harmonics, if present in the modulation signal, will superpose with the modulation frequency at  $f = f_m$  due to aliasing. Since this is the frequency component we use to extract the phase delay, this introduces a phase error. The phase error is systematic and can be compensated analytically in theory. However, in practice, non-idealities of the sampling process are addressed by means of calibration.

#### 11.4.1.1 Demodulation using correlation theory

Alternatively to the demodulation by sampling, cross-correlation can be used as the explanation model to extract the phase of the received signal. The cross-correlation which is defined for real functions as

$$(f \star g)(\tau) = \int_{-\infty}^{\infty} f(t)g(t+\tau)dt \quad (11.41)$$

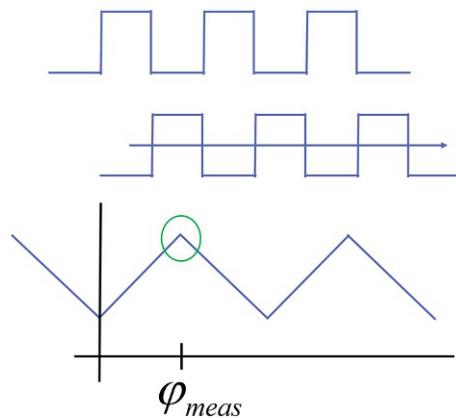
can be used to find how much a signal  $g$  must be shifted to make it identical to  $f$ . The signal  $g$  slides along the  $t$ -axis and the integral of the product is calculated. This is done at each position and when both signals match, the cross-correlation is maximized. Obviously, this makes the approach useful for determining the time delay between two signals which can be expressed as

$$\tau = \arg \max_t ((f \star g)(t)) \quad (11.42)$$

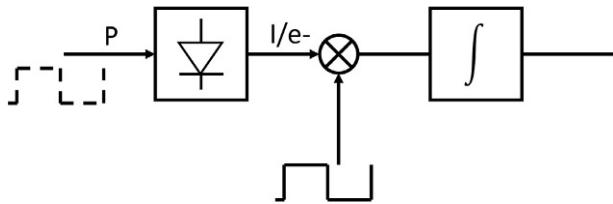
For the sake of simplicity, let us consider square wave modulation. Correlating two square wave signals results in a triangular signal which we refer to as the autocorrelation function (ACF) (Fig. 11.14).

The signal flow of the demodulation through correlation approach is illustrated in Fig. 11.15.

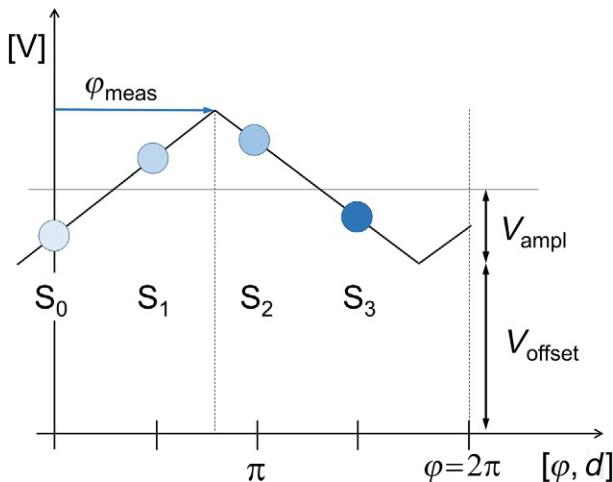
A pixel that incorporates this signal flow acquires one sample of the cross-correlation function. By shifting the reference signal subsequently, the cross-correlation function can be sampled accordingly which is illustrated for four samples in Fig. 11.16.



**Fig. 11.14** Correlation of the emitted and the reflected signals.



**Fig. 11.15** Demodulation by correlation.



**Fig. 11.16** The autocorrelation function defined by three variables.

To fully describe the correlation function, three unknowns have to be calculated: the amplitude, the offset, and the phase. This makes the task to extract the time or phase delay similar to what has been described in the previous section. In fact, the same DFT theory is used which results in the same three equations which have been derived for the demodulation by following the sampling approach:

$$\varphi_{meas} = \arctan \left( \frac{S_3 - S_1}{S_2 - S_0} \right) \quad (11.43)$$

$$V_{ampl} = \frac{\sqrt{(S_3 - S_1)^2 + (S_2 - S_0)^2}}{2} \quad (11.44)$$

$$V_{offset} = \frac{S_0 + S_1 + S_2 + S_3}{4} \quad (11.45)$$

Technically, following the argumentation from section demodulation by sampling the equations to compute  $\varphi_{meas}$ ,  $V_{ampl}$ , and  $V_{offset}$  is defined for the first harmonic only, which is why aliasing will introduce a systematic phase error. In practice, the cross-correlation function is somewhere between a triangular and a sinusoidal function. Again, following the argumentation from section demodulation by sampling, these kind of systematic error can be handled by calibration.

### 11.4.2 Demodulation device approaches

In [Section 11.4.1](#), we explained that we can measure distances by means of demodulation using a continuously modulated optical radiation field. Both explanation models, demodulation by sampling and demodulation by correlation end up described by the same equations for the four-phase algorithm. In fact, both approaches are equivalent from both the computation and pixel architecture points of view. Pixels that incorporate demodulation are synonymously referred to as sampling demodulation or just demodulation pixels.

In [Section 11.3](#), we have seen that the modulation frequency contributes linearly to the figure of merit which is why a high modulation frequency is of specific interest. Usually, frequencies between 20 and 130 MHz are used, resulting in effective sampling rates of between 80 and 520 MHz, respectively, assuming four sampling points per period. To realize such a high-frequency demodulation pixel, an efficient electro-optical shutter mechanism is needed. In general, the requirements for demodulation pixels are the following:

1. To increase the SNR, successively adding short-time integrated sampling points without adding additional noise beside of the photon shot noise results highly beneficial. Using this technique makes the system insensitive to frequencies other than the modulation frequency, and also robust to jitter. That is why demodulation pixels are also called lock-in pixels.

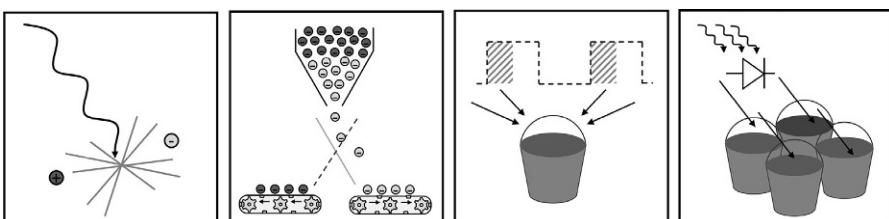
2. The fill factor of the pixel should be high enough to ensure high sensitivity. Assuming a fixed exposure time, the sensitivity contributes with the square root to the SNR and so to the shot noise limited depth resolution scenarios.
3. The intensity of the reflected light decreases with the square of the target distance. In addition, the intensity is further attenuated according to the reflectivity of the target. This is why the pixels require a high DR. A high full well capacity also helps, as uncorrelated light reduces the DR for the active illumination, and so the  $SNR_{max}$ .
4. The pixel must store at least one sampling point. If not all sampling points required to compute the distance can be stored within the pixel in parallel, the sampling points need to be acquired sequentially which can cause motion artifacts. Today, 1 and 2 tap pixels are used for this task, as this is usually a good compromise between pixel functionality and its fill factor.
5. Most of the applications of depth imaging require global shutter sensors. The pixel must be able to store the sampling point after exposure until it is read out. During the readout phase, the signal must not be corrupted by any light the pixel is exposed to.

We can summarize the demands on demodulation pixels by defining four tasks these pixels must be able to carry out which are schematically illustrated in Fig. 11.17 (Lange, 2000).

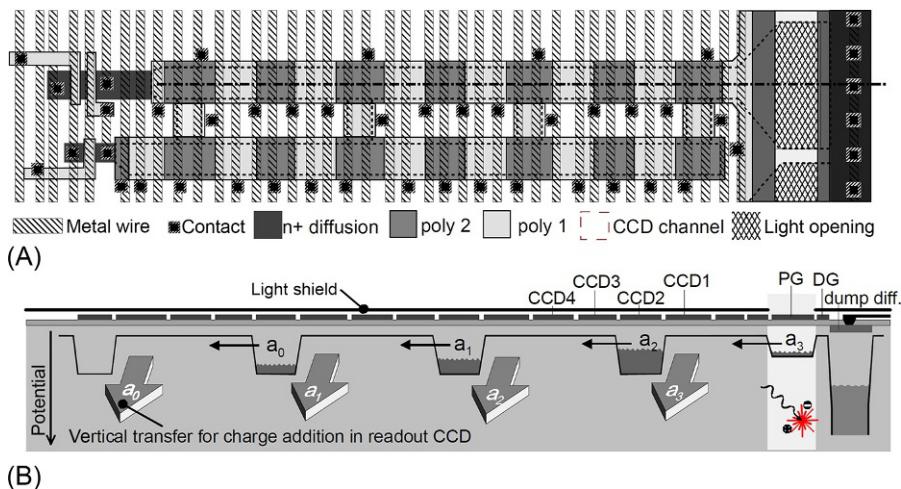
#### 11.4.2.1 CCD-based demodulation pixel

The generation of photoelectrons, their noiseless spatial transportation, separation, and accumulation by means of shift and shutter operation in the charge domain can be realized with the CCD principle (Lange, 2000). In fact, a standard CCD pixel can be operated as one-tap demodulation pixel without any major modifications.

The multi-tap lock-in pixel based on the CCD technology, illustrated in Fig. 11.18, was introduced already in 1995 (Spirig, 1995). The device consists of a light-sensitive photogate that is connected to a four-phase CCD-line. Every CCD-element, which consists of four CCD gates, is connected to another CCD readout line by so-called transfer gates. During operation, the upper CCD-line is clocked at an appropriate speed so that the photogenerated electrons are transported along the CCD-line. After four CCD shifts, which corresponds to exactly one modulation period, each CCD-element carries one



**Fig. 11.17** Basic functions of a demodulation device: (first on the left) generation of electron hole pairs, (second from the left) electron separation, (third from the left) repeated sampling and accumulation (fourth from the left) in-pixel storage of the sampling points.



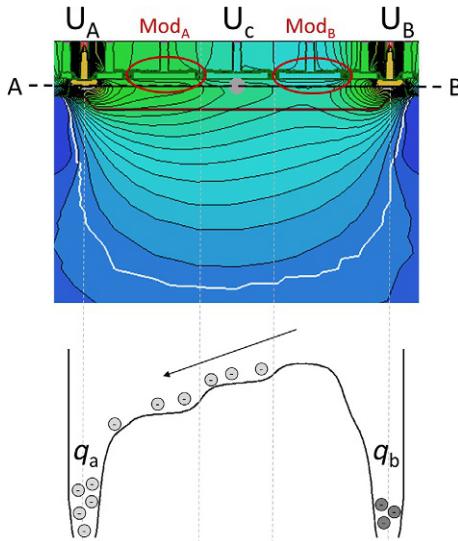
**Fig. 11.18** Multi-tap lock-in-CCD pixel (A) layout view and (B) cross section.

sample point. After these four sample points have been taken, they are transferred to the readout CCD by activating the transfer gates. During this transfer, additional photogenerated electrons are dumped to a sense-node diffusion to avoid parasitic charge from being accumulated in the same node. This procedure is repeated so that the amount of charges collected in the readout CCD-line is increased until a sufficiently high SNR is achieved. Finally, the readout line is clocked to transfer the sample points to the readout stage.

A general problem in using CCD technology is the speed of the charge transfer process. Also the fill factor of 4-tap CCD lock-in pixels is relatively low ([Lange, 2000](#)). To improve both, the speed of the demodulation and the fill factor, different versions of the CCD lock-in pixels have been evaluated in [Lange \(2000\)](#). The technology finally developed to the photogate PMD pixel architecture which will be introduced in the following section.

#### 11.4.2.2 Photonic mixer device

The PMD is a semiconductor device that performs both sensing and demodulation ([Schwarze, 1997a, b](#)). Since the fabrication of a PMD is simply based on a conventional CCD or CMOS technology, a VLSI integration of PMD pixel arrays is possible at the cost and quality comparable to CCD or CMOS-based image sensors. The PMD demodulates by controlling and separating the photogenerated charges into two or more drifting directions. After being separated, they are integrated and processed accordingly. The charge-transfer mechanism of the PMD is comparable to that of CCD. However, there are fine differences such as, for example, the fact that the charge transport in a CCD runs in a single direction, while in the PMD channel the charges move in both directions ([Xu, 1998](#)).



**Fig. 11.19** Cross section and surface potential of the 2Q-PMD device.

Fig. 11.19 shows a cross section diagram of a PMD including the electrostatic distribution within the semiconductor. The PMD device consists of at least two conductive and transparent electrodes which are isolated from the doped substrate by a  $\text{SiO}_2$  or  $\text{Si}_3\text{N}_4$  isolation layers. On the left and right side, the device is bounded by two PN-readout diodes. The two input terminals  $Mod_A$  and  $Mod_B$  are connected to the push-pull modulation voltage  $u_m$  so that applies  $U_{mod_A} = U_0 + u_m$  and  $U_{mod_B} = U_0 - u_m$ .

During the modulation process, the optical active area under the gates is fully depleted. The potential distribution in Fig. 11.19 causes the photogenerated charges to drift toward the surface where the modulated seesaw-like potential gradient causes the carrier to either move to the left or to the right where they are drained to the readout diodes for integration. Thus, the demodulation effect results from the synchronized charge separation caused by the push-pull modulation applied to the modulation gates. Thus, the PMD perfectly satisfies the basic functions of the general demodulation device as introduced in Fig. 11.15. The charge transfer at the surface is dominated by the fringing field effect. Optimization can be done by further shaping the potential distribution at the surface using additional gates and by introducing a buried channel. Both fine tune the potential gradient and so the separation process. The PMD is suitable to operate at high frequencies with high transfer efficiency.

#### 11.4.2.3 Current-assisted photodiode (CAPD)

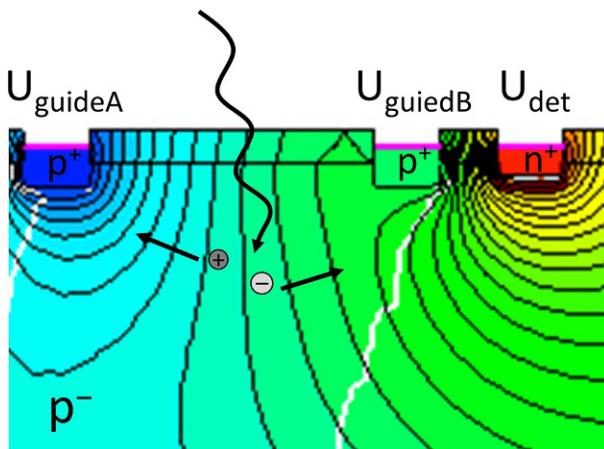
The basic idea of the current-assisted photodiode (CAPD) is to generate a drift field in the neutral zone of the bulk (Van Nieuwenhove, 2005). This will accelerate the photogenerated minority (electrons) carriers and so their detection which enables high

bandwidth photodetectors. Such a drift field can be invoked by applying a voltage over two terminals  $U_{\text{guide}_B}$  and  $U_{\text{guide}_A}$  which induces a majority (hole) current flow through the p-substrate. This can be realized with conventional semiconductor technology which makes the CAPD very easy to transfer to any CMOS or CIS foundry, respectively.

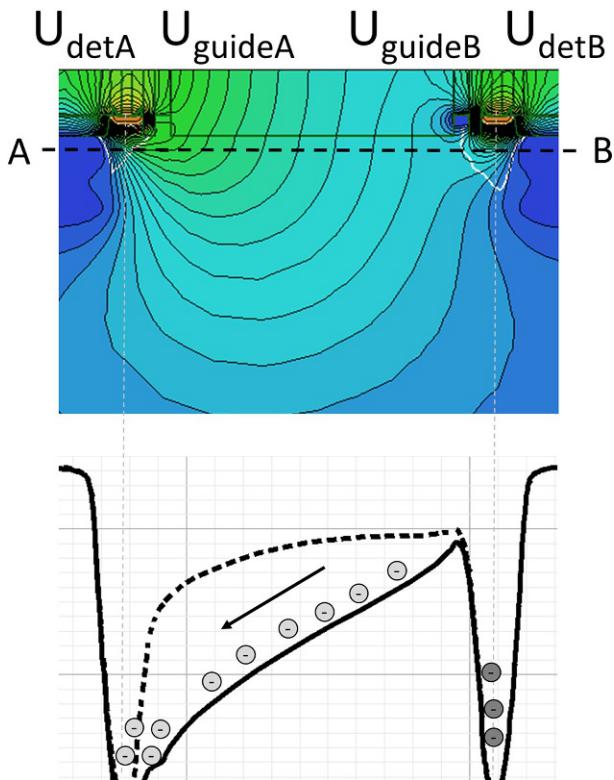
[Fig. 11.20](#) shows how this can be used to separate photogenerated electron-hole pairs according to the structure presented by [Van Nieuwenhove \(2005\)](#). The hole will move toward the  $p^+$  region of the terminal with the lower voltage as part of the majority (hole) current flow. The electron will be accelerated in the opposite direction because of the drift field caused by the majority current flow. Near the built-in  $p^-/p^+$  region, the electron requires some diffusion to reach the depletion drift field of the  $p^-/n^+$  junction and will be detected. A major advantage is the speed and the sensitivity of the detector as photogenerated electrons can be attracted quickly from very deep within the substrate.

Obviously, by modulating the voltages at the two terminals  $U_{\text{guide}_B}$  and  $U_{\text{guide}_A}$  the current flow, the drift field and consequently the direction in which the carriers are being directed can be reversed. By doing so, incident modulated light can be demodulated. A structure that consists of two detector nodes, one at each side of the device, according to [Van Nieuwenhove \(2005\)](#) is shown in [Fig. 11.21](#). This structure implements the general demodulation device which was introduced with [Fig. 11.17](#).

Similar to the PMD, the CAPD can be operated at high frequencies with a high transfer efficiency. The demodulation efficiency significantly depends on the voltages applied over the two terminals. The higher the voltage the higher the drift field that accelerated the carriers. Unfortunately, this increases the majority current flow and so the power consumption of the pixel also. This can become significant for arrays with high pixel counts.



**Fig. 11.20** Basic principle of CAPD detection  $U_{\text{guide}_B} > U_{\text{guide}_A}$ , according to the structure presented by [Van Nieuwenhove \(2005\)](#).

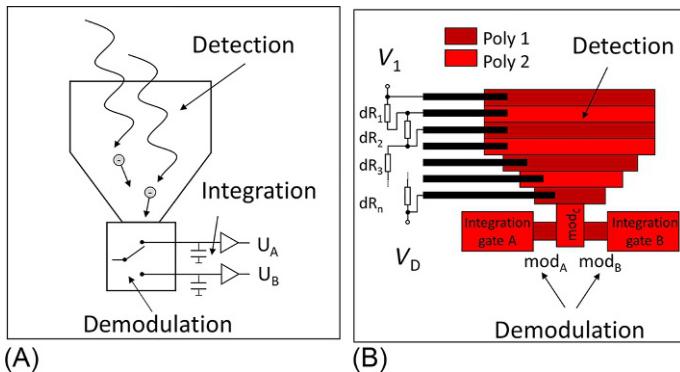


**Fig. 11.21** Cross section and surface potential of a CAPD photonic mixer according to the structure presented by [Van Nieuwenhove \(2005\)](#).

#### 11.4.2.4 Static drift-field demodulation

The idea of the static drift field pixel is to strictly separate the three basic functionalists detection, separation, and integration of the general demodulation pixel ([Büttgen, 2007](#)). **Fig. 11.22A** illustrates the block level view of this pixel concept. Compared to the previously introduced PMD, CAPD, and QEM pixel architectures, the detection and separation regions of the photogenerated carriers are in this case spatially separated. A large detection region is used for the generation of electron-hole pairs. The detected electrons are than accelerated by a static drift field and directed to the demodulation stage where they are separated and integrated into two storage areas by means of sampling or correlation.

The demodulation stage can be implemented similar to the previously introduced lock-in pixel architectures. Since the photogenerated electrons, when reaching the demodulation stage, will have reached the surface of the silicon substrate and are spatially concentrated, the demodulation stage can be very small with minimum additional transport delay. To generate a static drift field within the silicon substrate, a semitransparent and high resistance photogate structure with two terminals can be used. While being transparent in the UV and blue region of the visible light,



**Fig. 11.22** Static drift field demodulation (A) schematic view (B) simplified.

polysilicon is transparent in the NIR which is usually the spectral range where ToF systems are operated. Applying a voltage over the two terminals will invoke a drift field due to the capacitive coupling between the gates and the substrate.

However, the current flow through the polysilicon causes significant power consumption per pixel which is why in Büttgen (2007) another approach was suggested. Here, instead of applying subsequent shift operations to transfer packets of charges throughout the photoactive area, the gates are applied with linearly increasing constant voltages. A simplified layout view of this pixel architecture is shown in Fig. 11.22B. This voltages due to capacitive coupling and usage of buried channel CCDs generates a lateral drift field directing the carriers toward the demodulation stage. The lateral drift-field pixel can operate at high frequencies due to the fact that the major transport mechanism is here based on the drift of photogenerated charge in the large detection region and influenced by the fringing fields in the demodulation region.

Other approaches introduce a static lateral drift field using modified pinned photodiodes (PPD) which eliminates the need for the complex biasing of the adjacent photogates. The drift field along the PPD can be created by means of the pinning voltage being laterally shaped through n-well doping gradients (Durini, 2010a, b; Brockherde, 2012). The doping gradient can also be introduced by means of a spotted or scattered n-well after out-diffusion of the dopants or by means of a narrow channel effect in the n-well (Theuwissen, 2018). The separation of the charges after being generated and laterally drifting through the detector region is realized using transfer gates which are electrically modulated. Various optimizations have been reported recently to optimize the speed of the charge transport as well as its transfer into the storage node (Acerbi, 2017, 2018; Rodrigues, 2017).

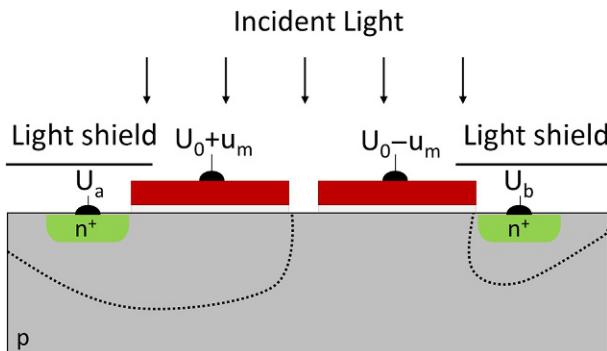
#### 11.4.2.5 Quantum efficiency modulation

The term quantum efficiency modulation (QEM) comes from the idea to control the quantum efficiency of a photodetector by means of a control voltage. In Bamji (2000) various structures suitable to implement this approach are presented. A simple pixel architecture composed of two QEM structures from Gokturk (2004) is illustrated in

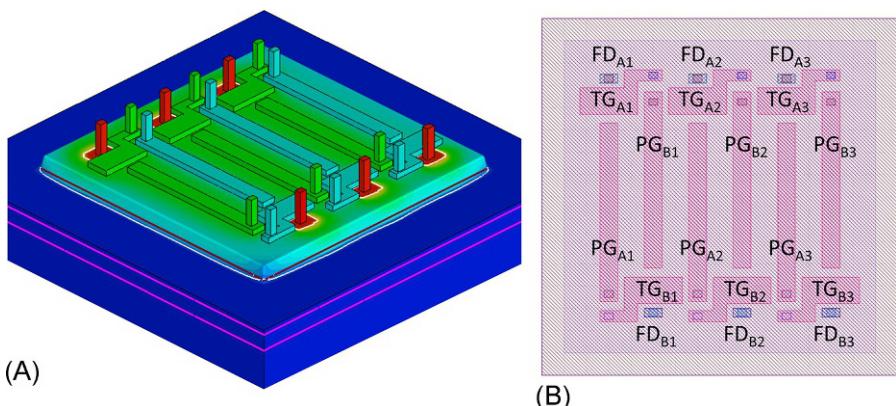
**Fig. 11.23.** The two nodes compete to attract the photogenerated charges collected from within their zones of influence. The size of these zones of influence can be modulated by the two photogates. Hence, the structure implements DQM and so the general model of a demodulation pixel if synchronized with the emitted light.

Obviously, if the two QEM structures are placed tight to avoid photogenerated charges outside the active zone of influence impairing the separation efficiency, the potential distributions within the substrate caused by the two gates will superpose and the structure will evolve into a PMD device as it was introduced previously. Hence, the charge transport to the surface will be based on a drift field while the transport at the surface will be dominated by fringing fields.

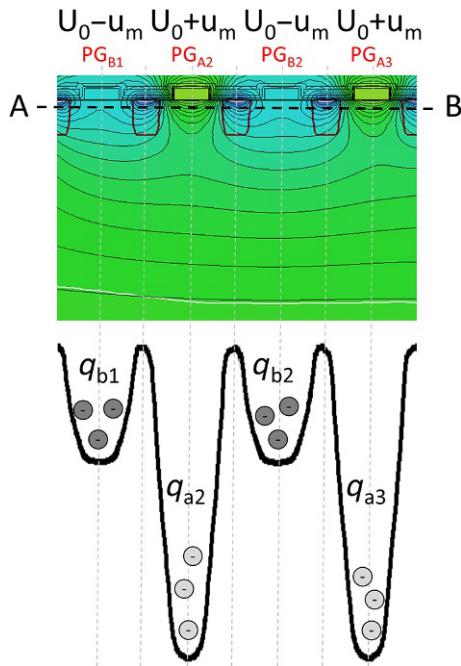
Another implementation presented by Bamji (2015) avoids the lateral transport under the gates completely. Instead the electric fields caused by the poly-gates attract and collect the photo-charges under the gate oxide. **Fig. 11.24** shows a 3D TCAD structure and the associated layout view of the pixel detector with poly-gates  $PG_A$



**Fig. 11.23** Simplified cross section of a pixel implementing QEM.



**Fig. 11.24** QEM pixel architecture (A) 3D TCAD structure, (B) layout view.



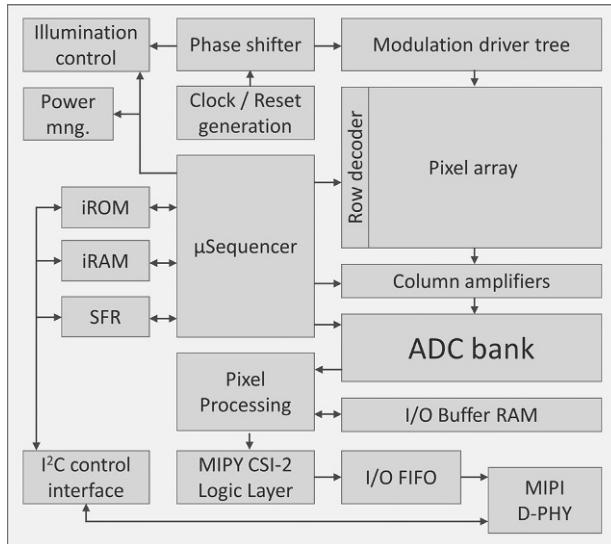
**Fig. 11.25** Potential distribution of the QEM pixel within the substrate.

and  $PG_B$  shaped like fingers. The poly-gates  $PG_{An}$  is at high positive bias while poly-gates  $PG_{Bn}$  are at ground. Under these conditions, the photogenerated electrons  $qa_n$  are collected under poly-gates  $PG_{An}$  which are shown in Fig. 11.25.

Potential barriers created by  $p^+$ -doped regions between the poly-gates isolate the collection zones under the gates. This ensures the charges will not move toward adjacent gates once they are collected by one gate. The potential distribution within the substrate and at the surface is illustrated in Fig. 11.25. Upon collection under the poly-gate, charges will diffuse slowly to the floating diffusion collecting node  $FD_A$  or  $FD_B$ . Since the photocharges always remain under the gate where they have initially been collected this relatively long transport delay is completely decoupled from the critical charge separation operation. This makes this method suitable for high modulation frequencies.

### 11.4.3 ToF Image sensor architecture

In Section 11.4.1, the theory of indirect ToF by means of demodulation was introduced. Different architectures that implement the different functionalities such as detection, separation, repeated sampling, and in-pixel storage of the samples of the general demodulation pixel mode, were presented in Section 11.4.2. Obviously, it requires additional building blocks to make up a sensor that implements ToF imaging.



**Fig. 11.26** Sensor architecture block diagram.

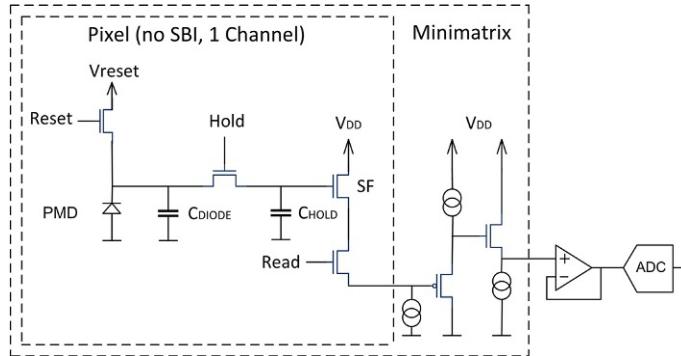
A block diagram of the ToF SoC architecture is illustrated in Fig. 11.26. This section will detail the major building blocks of a highly integrated ToF image sensor.

#### 11.4.3.1 Pixel electronics

Actually, an important function, the in-pixel storage of the sample is normally performed using a simplified approach using for it the readout node, normally in form of a floating diffusion. In practice, many concepts for the pixel electronics have been proposed, some of which include specific additional functionality to extend the DR, cancel  $kTc$  noise, or to suppress ambient light. Whether to implement these in-pixel functionality or not is usually a trade-off decision driven by pixel size, fill-factor, and application requirements specifications, respectively.

All ToF pixels have in common that they implement a global shutter concept. This is due to the fact that distortions caused by motion of the camera or objects in the scene are usually not acceptable in these applications. Also, the pixel complexity increases because the phase-shifted electrical modulation signal applied to the pixel has to be multiplexed row wise in a rolling approach. Fig. 11.27 exemplarily shows the pixel electronics of one channel of the very basic 4T global shutter active pixel that can be used with the demodulation device approaches that were introduced in Section 11.4.2.

To reset the detector, the *Reset* and *Hold* switch is closed here to charge both capacitors  $c_{DIODE}$  and  $c_{HOLD}$  to the reset voltage  $V_{reset}$ . During integration, the *Reset* switch is opened and the modulation clocks, for example,  $Mod_A$  and  $Mod_B$  for a PMD-based pixel as shown in Fig. 11.19 are modulated 180° out of phase at the modulation frequency. During the entire integration, the *Hold* switch remains closed because both



**Fig. 11.27** Pixel and readout circuitry.

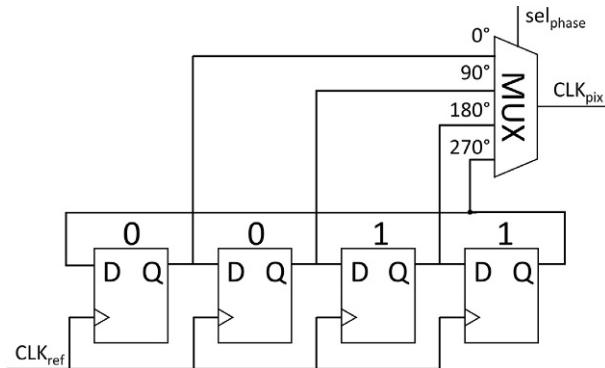
capacitors  $c_{DIODE}$  and  $c_{HOLD}$  combined make up the integration capacitance  $c_{int}$ . At the end of the integration, the *Hold* switch is opened which results in the collected charges to be distributed between  $c_{DIODE}$  and  $c_{HOLD}$  in a way that the voltage across both capacitors remains equal. After that, the *Reset* switch is closed to drain any parasitic light hitting the detector. The combination of *Hold* and *Reset* functions implements a read-after-exposure global shutter functionality. For the readout, the *Read* switch is closed and the source follower *SF* is driving the signal into the readout chain.

#### 11.4.3.2 Amplifiers and ADC

Because of the need to capture at least four samples to compute one distance value as described in [Section 11.4.1](#) ToF puts requirements on the readout electronics similar to those on high frame-rate cameras. Multifrequency algorithms, to overcome the problem of the unambiguity range when using high modulation frequencies, additionally increase the frame-rate requirements on raw data level. For each raw frame, the pixel array has to be read out sequentially which reduces the time for the exposure because of the read-after-exposure global shutter approach. Because of this and to avoid motion artifacts, a minimized readout time is required. Depending on the spatial resolution of the sensor, the bandwidth of the readout chain can easily be required to be as high as 100 MB/s to 2 GB/s with bit width between 10 and 12. In contrast to rolling shutter operated 2D image sensors column parallel ADC single slope ADC are rarely used for ToF sensors because of the high conversion rate. Instead SAR converters are primarily used with architectures with greater number of slower thus smaller ADC is preferred to a smaller number of fast ADC. This is because the settling time, and so the current required to drive the input stage, can be reduced which optimizes the overall power consumption of the readout path.

#### 11.4.3.3 Clock generation and phase shifting

The clock and reset generator provides the system clock and reset for the digital core just as in any other SoC. However, it additionally creates the pixel array and light source clocks from a high-speed reference clock usually generated by a high-speed



**Fig. 11.28** Phase shifter schematic diagram.

phase-locked loop (PLL) circuit. According to the demodulation theory introduced in [Section 11.4.1](#), the pixel or illumination clock is required to be shifted in phase to sequentially capture at least four samples to calculate the distance. [Fig. 11.28](#) shows a circuit that generates four modulation signals which are shifted by 0, 90, 180, and 270 degree accordingly.

The major advantage of this circuit is that it is very simple to implement. However, it requires the reference clock  $clk_{ref}$  to be four times faster than the modulation frequency  $clk_{pix}$ . Also, the four flip-flops need to be balanced precisely to ensure exact phase delay and signal shape of the modulation signal because. This is because of the higher order harmonics of the modulation signal being affected. As described in [Section 11.4.1](#), these harmonics create aliasing which causes a systematic error in the computation of the phase. This effect can easily vary with temperature and process corner and needs to be addressed by calibration. In practice, more complex clock generators are used that support more complex modulation schemes like PN sequences or harmonic cancellation ([Bamji, 2015](#)). Advanced circuit design concepts like dynamic flip-flops are used because of the very high frequencies at the edge of the CMOS technology nodes involved ([Bamji, 2015](#)). Also, selectable duty cycles are used to minimize the aliasing and maximize the geometric dme as explained in [Section 11.4.1](#).

The generated clock signals are fed to the illumination block which drives and controls the laser or LED output, respectively. The pixel clock is fed into a modulation driver which is implemented as balanced clock tree to ensure optimal phase delay distribution over the whole array. Nonoptimal phase distribution gets visible in the so-called fixed pattern phase noise which is a pixel individual phase offset and must be compensated during the phase computation.

#### 11.4.3.4 Pixel processor

The pixel processor receives data from the ADC and performs pixel level data processing. Depending on the amount of on-chip RAM, the performed operations can be as simple as pixel binning or filtering in the digital domain, saturation

detection, or data normalization. If the RAM is large enough to buffer at least one raw image on-chip high dynamic range (HDR) exposure sequences can be implemented to extend the DR of the sensor (Bamji, 2018). In practice, it is a trade-off between on-chip functionality and on-chip memory size because SRAM can significantly increase the silicon area and so the cost of the sensor.

After the pixel processing, the data are transferred off-chip through the pixel interface which can be parallel or serial. Depending on the interface standard, the pixel processor performs scrambling, serialization, and coding of the data. Usually, the data stream is extended with additional data like frame counters or chip temperature.

#### 11.4.3.5 Micro-sequencer

The sequences control the operation of the image sensor. A typical frame capture requires the clock generator being programmed to provide the correct modulation frequency and phase shift, respectively. After the PLL is locked, the pixel is reset and the exposure is started for a configured exposure time. After that, the array is read out row wise and the pixel data are converted to digital values, processed and transmitted off-chip. During this frame sequencing, the various functional units are not necessarily required to operate all the time. Instead, the sequencer controls the units involved being powered and clocked as needed which is essential for low-power applications. The entire procedure is called “one measurement sequence” because one raw data frame is produced. Based on the particular application multiple sequences which differ in settings like modulation frequency, exposure time or phase shifter setting are used to compute the final depth data stream.

To ensure maximum flexibility and late binding for a particular application, the sequencer should be configurable in a wide range to operate different sequence schemata. For maximum flexibility, the sequencer is usually implemented as application-specific instruction set architecture (ASIP). The control sequence to operate the sensor is implemented as a series of microinstructions downloaded into the instruction RAM or hard coded in the instruction ROM. This firmware-based approach allows the sensor to be adapted as application requirements change or new applications come up.

#### 11.4.4 Summary and future trends

In Section 11.4.1, we introduced indirect ToF where the phase delay between an emitted and reflected continuously modulated light wave is determined. Sampling and cross-correlation theory has been presented to explain how to compute the phase delay and so the distance of the reflecting object to the sensor. Both approaches result in the very same n-phase algorithm defined by Eqs. (11.32)–(11.34) which is of no big surprise because for both approaches the same mathematical operations are involved: multiplication, summation, and shifting of the signals. In fact, there is a third approach based on the homodyne mixing or homodyne detection theory which again performs the same mathematical operations and again results in the same three equations (Gokturk, 2004).

Different architectures that optimally implement the functionality of the general demodulation pixel model in terms of quantum efficiency, dme and bandwidth have been described in [Section 11.4.2](#). Other demodulation pixel approaches have been reported over the years but never exceeded an academic or predevelopment level, respectively, which is why they have been omitted ([Kim, 2011, 2012](#); [Oh, 2012](#)). Another demodulation pixel approach not mentioned in [Section 11.4.2](#) is based on single-photon avalanche diodes (SPAD) which has been successfully commercialized ([Rae, 2017](#)). However, in practice, SPAD-based sensors were limited to small spatial resolution until recently because of the high pixel level electronics. Higher spatial resolutions have been reported by [Al Abbas \(2016\)](#) where smaller pixel sizes have been achieved through separating the detector element and pixel electronics on two wafers which are stacked using hybrid bonding technology.

Until now, the development of ToF image sensors was driven by lower-volume markets in the industrial space which was why cost-efficient CMOS nodes have been used and evolution was more or less limited to EPI tuning. However, there is a growing interest to bring 3D sensing to mobile devices ([Böhmer, 2016](#)) which will, similar to 2D image sensors, dramatically accelerate the evolution of ToF technologies. The demand for higher resolutions which can only be cost efficiently addressed with smaller pixels will be a challenge. Smaller pixels usually result in a lower full-well capacitance and so will reduce the DR  $SNR_{max}$ . These challenges will require to migrate into cutting edge CIS processes which offer the whole portfolio of 2D image sensor options like BSI, in-pixel memory, minimal low-noise size transistors and efficient ADC circuitry and pixel level stacking. This is essential because every photon counts ([Fossum, 2016](#)).

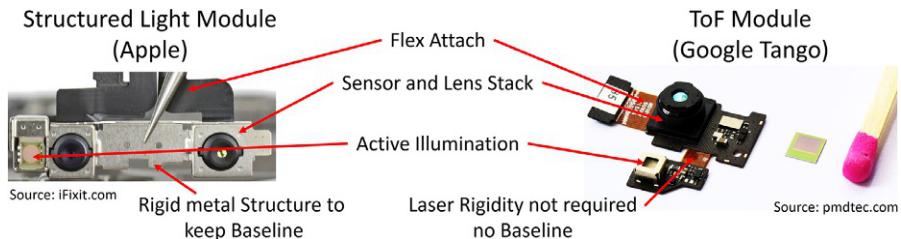
## 11.5 Applications

The launch of the iPhone X in September 2017 with a 3D “selfie” sensor has made 3D imaging a strongly growing and important market. Every mobile company is evaluating 3D imaging, planning both front and rear facing 3D sensors for their platforms, and a wide number of other consumer segments are expected to follow.

The iPhone X, however, was not the first successful 3D imaging platform. The Microsoft Xbox Kinect in November 2010 was the first consumer device that introduced consumers to 3D imaging, and it had the fastest adoption rate of any new electronic device at the time, surpassing even the iPad. It was assumed then that all types of electronic devices with 3D sensors would quickly follow.

However, mass adoption of 3D did not happen after the Kinect. 3D imaging was still too big, expensive, and required a lot of power. In the following 7 years, however, those issues have been solved, the supply chain has matured, and the proliferation of computer vision into areas such as autonomous driving and in-car monitoring have driven 3D technology into more use cases. In addition, emerging markets such as augmented reality (AR), personal robotics, and interconnected monitoring devices [“internet of things (IoT)’] will all require 3D imaging to come to their full potential, promising large growth in this area for years to come.

As “Big Data” and analytics increase, the number of machine-monitored cameras will increase as computers are assigned to more vision monitoring tasks. The ultimate



**Fig. 11.29** Comparison between SL and ToF [Photos via iFixit ([Yole2, 2017](#)) and pmdtechnologies ([Yole1, 2017](#))]. The ToF module is clearly smaller and delivers the same data quality in the near up to 75 cm range with 2 × pixel resolution, and clearly better depth data quality for higher distances.

computer-controlled device—the self-driving car—will need 3D imagers on both the inside and outside to accurately interpret both the occupants and the car’s continually changing surroundings.

Comparing the different 3D technologies available, especially the 3D ToF approach seems to be an interesting candidate due to a number of reasons. As ToF systems are very flexible in their usage (by means of electronically chaining respective measurement modes), their potential applications are manifold. Even an integrated 3D ToF camera can serve a bunch of different use cases—from short range to long-range measurements, from lower depth resolution at high fps to higher resolution with lower fps and somewhere in between. This scalability “on the fly” is not possible with stereo or structured light approaches.

On top of these advantages, 3D ToF systems show higher robustness (mechanically as well as with respect to measurement conditions), lower-power consumption as well as smaller size and bill-of-materials (BOM) cost than competing technologies, such as stereo, structured light, depth-from-motion, and others (Fig. 11.29).

### 11.5.1 Mobile market

Although the iPhone X launch got more press for their 3D front facing “selfie,” mobile phones with 3D imagers in rear-facing mode were in the market almost a year before the iPhone X launch.

However, the Apple feature garnered more press since its use case was clearer (Face ID) and due to the importance of the Apple brand. However, in the long run rear-facing 3D sensors will be even more important due to their potential use in AR, a megatrend in mobile ecosystems.

#### 11.5.1.1 World-facing 3D

Google Tango for Android was the first offering, which was effectively canceled and merged into Google’s ARCore AR platform. Apple rumored to be working on world-facing intro in 2019, probably using internally developed ToF technology. AR headsets which are expected to reach the market in the next 3 years.

### 11.5.1.2 *Front-facing “selfie” 3D imaging*

The iPhone X introduced the world to the world’s first 3D selfie imager. While some of the features have actually been introduced before with the regular selfie camera (such as Face ID using 2D imaging by Samsung), using 3D allows for a more robust solution and additional features.

### 11.5.1.3 *Apps and functionality*

The 3D feature on the iPhone uses structured light technology, claiming to produce  $200 \times 150$  or 30 K voxels/depth points, which turns out to be less performance as shown by recent teardowns (below 20 K depth points). Recall that the 3D depth for a structured light device is dependent on the projector, not the receiver, which in this case has 1.4 million pixels. 20 K points at arm’s length is sufficient voxel density for face capture, but will have limited functions for anything beyond a meter, so the same design cannot be used in the future for a world-facing 3D camera.

The functionality introduced by Apple in the 3D selfie can be classified into three broad areas:

- Face ID/Face unlock—the feature will allow users to unlock devices simply by looking at their phones and swiping up, essentially letting iPhone owners to use their face as a password.
- Facial animation/emojis/morphing—according to Apple, this feature “uses more than 50 different muscle movements to mirror your expressions onto an emoji,” so allows a user to face control an animation.
- Picture improvement for selfies (lighting/“bokeh”)—by using the depth data, the background behind the user can be brought into focus or blurred, effectively allowing the user to change the FoV. The 3D data also allow the user to improve the lighting on a selfie.

Surprisingly, the iPhone X does not yet offer background substitution, which is an obvious and easy feature to have when having a 3D map when using video chat. It is unclear if this will be added in the future.

Moreover, 3D is expected to replace 2D + accelerometers due to the increased reality and immersion by the user, but the higher cost of 3D has made phone manufacturers decide to develop the AR ecosystems before launching full 3D. With the 3D sensor in place, other features that were shown on the Tango platform include improved computational photography, 3D scanning of people and objects and indoor navigation/simultaneous location and mapping (SLAM).

### 11.5.1.4 *AR and mobile phone evolution*

While handsets can be used for AR experiences, many are predicting that the handset will move to a wearable device (or accessory) on the head that provides a head-up display (HUD), an updated version of Google Glass. This next generation will require 3D sensors to create an accurate HUD as well as gesture control, as seen in AR devices today in the enterprise market.

Of course, consumer AR was attempted already in 2013 with Google Glass. Although a bold experiment the product did not catch on for a variety of factors: first, the technology was too early, and second, both application (app) availability and user acceptance of “always-on-apps,” providing individualized, situation-specific information, was relatively low at that point of time. Today, people can be seen using their cellphones while walking, driving, at the gym, and nearly every other activity.

Now that people live through their cellphones even while doing other things, acceptance of mobile AR will come more naturally. In addition, the technology will be better, smaller, and more stylish. 3D sensors, which allow for gesture interaction in front of the device on a virtual interactive display, will be required for this segment.

#### ***11.5.1.5 Consumer mobile AR entries***

A technology “who’s who” have all publicly announced or hinted about being in the AR market including Microsoft, Facebook, Apple, Google, Magic Leap, and most major Asian vendors.

Even Apple’s Tim Cook, usually secretive about product direction, stated on Apple’s November 2017 earnings call “We’re already seeing things that will transform the way you work, play, connect, and learn. Put simply, we believe AR is going to change the way we use technology forever.”

#### ***11.5.1.6 Consumer AR as a phone accessory***

Most likely, the mobile AR glasses will be launched as an accessory, much like the iWatch and other personal devices. Like the iWatch, AR glasses will have their own display, own processor, and own sensors, and will be tethered to the iPhone.

This strategy keeps the costs down for initial consumer adoption. Once the market is better penetration versions that have a direct tether to the internet can be introduced. As for when this is happening, development kits are already being released to developers already by Magic Leap and others, and as noted above Apple and Android will use their mobile 2D AR aps as a base development platform. Mass consumer introduction is not expected before 2020, and perhaps a year or two later, and rumors have an Apple AR introduction in 2020.

#### ***11.5.1.7 3D sensors and consumer AR headsets***

As noted each AR headset needs a 3D sensor so the attach rate will be 100%, compared to mobile handsets, which will be an option for both front and rear facing so will have a much smaller penetration rate. By the time these mobile AR units are launched in the 2020 timeframe, the 3D module supply chain will be well established for mobile selfies, and no issues are seen for volume or size constraints. The main issues will be power consumption and outdoor use (sunlight resistance).

Both structured light and ToF are addressing the power issue by becoming more sensitive and efficient, allowing lower-power laser emitters to be used to get the same performance.

For outdoor use, this issue is solved today by ToF modules, in production in automotive, but still an issue for structured light. This is one reason why ToF has a long-term advantage over structured light, and ToF is used almost exclusively today in the enterprise AR market.

### 11.5.1.8 3D sensors and mobile volume

As noted above, the use of 3D sensors in mobile devices falls into three main categories:

- Mobile selfie—penetration starting with iPhone X in 2017 with followers starting in 2019.
- Mobile rear facing—penetration forecasted to start in 2019.
- Mobile AR—penetration forecasted to start in 2020.

While penetration rates of 3D sensors are expected to explode in the next 3 years, it should be noted that the mobile handset market itself is considered mature with 2%–3% annual growth through 2022, so 3D module penetration will provide an upside of those who participate in its supply chain.

Using these statistics, market researcher *Yole Développement* forecasts the 3D mobile market using three different scenarios for penetration of 3D ([Yole1, 2017](#)). Using the middle forecast value of 20% penetration of all mobile devices (phones and AR), and assuming handset/headset shipments reach 1.8B that year, this would be a total of 360 M 3D camera modules starting by 2022. This of course could be as much as 50% larger if the optimistic forecast levels are met ([Fig. 11.30](#)).

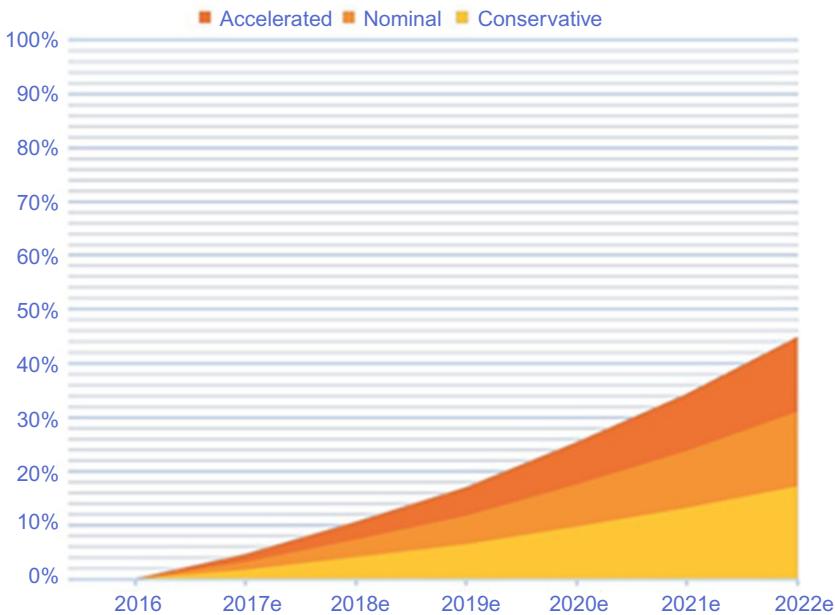
### 11.5.2 Automotive market

Excitement about autonomous cars, ride sharing, and electric cars have put automotive into the front pages of newspapers, the top of social media feeds, and made it a hot market for acquisitions and venture capital funding.

#### 11.5.2.1 Key trends in automotive

There are three megatrends in automotive today:

- Electric cars—both start-ups and established auto brands are pushing into this market and trying to create differentiation, accelerating the design cycle to mimic the consumer electronics market rather than the traditional automotive market. This is pushing experimentation and adoption of interior 3D for gesture control and driver monitoring.
- Advanced driver-assistant systems (ADAS)—ADAS is providing drivers with ever more information about both the interior and exterior of the car, driving sensor adoption of all kinds.
- Autonomous vehicles—autonomous vehicles are simply robots that carry people, and 3D vision is required for both exterior maneuvering, as well as monitoring what is going on inside the vehicle.



Source: Yole, used by Permission

**Fig. 11.30** Penetration of 3D imaging into mobile, a 20% penetration rate would translate to about 360 M imagers by 2022 (Yole1, 2017).

Each trend is driving 3D sensing demand for both inside and outside the vehicle for the following reasons:

- Interior—for human-controlled cars, sensors are needed to allow the user to have better access to the car, and for new user interactions such as gesture control. For computer controlled cars, 3D sensors are needed to monitor the driver and passengers, especially for “handing over” back and forth from an automatic driving mode to human driving control.
- Exterior—a 3D map of the exterior of the car, both long range and short range, are needed in both driver assist as well as autonomous driving.

### 11.5.2.2 *Interior automotive 3D sensing*

As noted above, automotive interior 3D sensing falls into two broad categories:

- Gesture control—active uses by the driver or passenger to control aspects of the car, usually the entertainment, communications, and environmental controls.

This feature was pioneered by BMW and introduced in their 700 series in 2015, and fanned out to their 500 series in 2017. Other BMW models expected to follow in 2018 and 2019, and other luxury car makers will introduce versions in 2018. Gesture is then expected to fan-out to midrange cars through 2020. ToF is the only technology considered for this use case today due to sunlight tolerance and ability to use LED as a light source. Laser sources are not yet auto qualified, and may not be for several years.

The main ToF suppliers today are Sony, with pmd expected to make significant inroads in 2018. Panasonic is also trying to enter the market with an automotive ToF sensor.

Overall, the volumes in-cabin gesture will be in the low 100 s of K, with the main revenue not coming from the sensor, but from the system and the gesture control software. In this category, Sony is well positioned due to their acquisition of SoftKinetic in 2015, as well as Valeo, who acquired gesture control software company Gestigon in 2016.

- In cabin monitoring (ICM)—passive observation of the driver and passengers, usually for safety reasons, and soon for “hand-off” of driver to computer driving and back.

ICM is starting implementation rollout with standard 2D sensors or other types of sensors for simpler tasks in 2018. As features expand, 3D will supplement 2D to obtain full data of what is going on in the cockpit. In some cases, the 3D sensor used for gesture control in the above case will be leveraged to add additional passive monitoring. The 3D sensor providers will largely be based on those that got a foothold in automotive gesture: Sony and pmd.

As with gesture, the main revenue will not come from the sensor, but from software, and in this use case the data. Business models are being proposed to capture and monetize the data taken in the cockpit. Due to the analytics potential, Google and other major data players are actively investing in this segment along with the exterior sensing.

### **11.5.2.3 Automotive interior vision rollout**

While BMW and other luxury brands are expanding gesture control to more lines, most car manufacturers are skipping gesture and going to ICM since ICM is needed for autonomous driving for handover to/from driver.

ToF seen as the go-to choice in this segment, so key players Sony and pmd are well positioned to penetrate this segment which will see 1–2 sensors per vehicle.

- Looking at a volume perspective, annual automotive production in 2017 is about 100 M units worldwide. Total automotive volume may actually flatten in the next few years due to the increase use of ride sharing.
- Interior 3D sensor penetration for gesture control and ICM will climb to high single digits in next 5 years, so growth volume should be approaching 8–10 M sensors/modules a year. Assuming a \$10 sensor ASP, this will be a \$80–\$100 M. Total available market by 2022, all held by ToF vendors.
- Inflection point to much higher penetration will occur when autonomous driving hits, but is close to 10 years out (see the following section).
- Most 3D sensor volume will be Sony and pmd ToF sensors, with ToF Panasonic entering and competing within 2 years.
- Additional entries into the market are difficult due to automotive quality requirements, which take time and investment.
- Light source vendors will be LED rather than laser for next few years as laser struggles to get automotive qualified due to temperature and other automotive conditions.

#### 11.5.2.4 Exterior automotive 3D sensing

Cars already have a number of exterior sensors, including ultrasonic sensors, 2D sensors, and even simple forms of RADAR. 3D sensing technologies will be added as cars become more autonomous since the car will need to get a 3D map of both the exterior and interior. The key to forecast when a current 3D technology is to start being used is based on whether the sensor is doing ranging or interpretation.

- Ranging—today's exterior ultrasonic and simple RADAR sensors provide a yes/no if something is in the way. They provide a warning to the driver or ADAS system, which can then take action, but not provide no information about what the object is, how big it is, if it is moving, etc.
- Interpretation—when computers are doing the driving, they need data to interpret the surroundings, which today's exterior sensors cannot do. The data for interpretation require a 3D depth map, and require 3D sensors.

#### 11.5.2.5 Exterior automotive 3D technologies

No single technology will cover a full 360 3D depth map around the car for all distances. Instead multiple sensors will need to be used, and development efforts are pointing toward different 3D technologies for different ranges.

- Automotive long range  $>5\text{M}$ —LIDAR is considered the solution for long-range outdoor depth map generation  $>5\text{ m}$ .
- Automotive short range  $<5\text{M}$ —closer to the car other technologies will need to be used. Today, backup and blind spot sensors are mostly ultrasonic, but ToF will make inroads as intelligence is added and a full 3D map is needed.

Besides distance, the other consideration for a technology is its FoV or “cone” that sees. For long range, only a narrow band is required to see where the car is going, so a narrow FoV, which favors LIDAR. For short range, where everything around the car short distance is needed, a wide FoV is required, which favors ToF.

#### 11.5.2.6 Automotive driving and 3D sensor use

While the term “autonomous driving” is used a lot, there are actually various levels of automation as defined by the Society of Automotive Engineers (SAE). Their defined levels—ranging from no automation to full automation—are used in the industry to talk about what level of computer control is embedded in the car. In general, the higher the level, and the higher the level of computer control, the more 3D vision is required since the computer needs a complete layout of the outside world.

Level 1, such as cruise control, has been around for some time. Level 2, which would be cruise control with steering support, such as lane change notification, is increasingly common. Level 3, with the computer in charge with human support, has early adopters in brands such as Tesla. Levels 3 and 4 with no human control are only in trial phases, with some promised transit services (i.e., with a planned, fixed route) being

proposed for rollout over the next 2 years. Mass driving to any destination with autonomous driving, however, is expected to be many years out.

#### ***11.5.2.7 Automotive driving and interior 3D sensor use***

As noted above levels 3 and 4—autonomous cars with human drivers being able to step in or a computer being able to step in if a driver is incapacitated, means that the driver needs to be monitored, either to see if they are able to take control, or if the computer needs to take over.

A level 5 fully autonomous car with no driver will need to know if passengers are seated, buckled, etc. before taking off, or if the robot-driven car needs to slow or stop if an issue inside the car.

#### ***11.5.2.8 Automotive driving and exterior 3D sensor use***

Most analysts have high Max/Min variance, but either way true penetration will not happen until the mid to late 2020s. Until then, both long-range LIDAR and short-range ToF for exterior 3D automotive sensing will be a development effort.

Autonomous driving will not take off for another decade, and the current investment in LIDAR is not so much for a return on the sensor volumes itself, but in the future data collection and data mining, including by insurance companies, governments, etc.

Exterior 3D is still in an early development stage with lots of entrants in LIDAR for long range, and early development in ToF for short range. LIDAR players are too numerous, and the market too early to forecast; ToF will be the main players doing interior sensing today: Sony, pmd, and Panasonic.

#### ***11.5.3 Emerging and secondary markets***

Like 2D imaging, mobile volume and revenue is expected to dominate the 3D market. Automotive, while not as large as mobile, will have higher margins and is harder to design out once designed in due to the lifecycle of vehicles. Automotive also has a certain prestige attached to it in 2018 as press, venture capital, and the public's imagination are captured by autonomous driving and what it means for society.

Many imaging companies have publicly stated they will focus on mobile and automotive for these reasons, but there are a number of secondary and emerging markets where 3D imaging will be used and could be significant segments for the technology in the future.

#### ***11.5.4 Virtual reality***

Virtual reality (VR) is a segment that received a lot of hype when Facebook acquired Oculus for 2 billion USD in 2014. The reality has not lived up to the hype as this is now seen as a niche market for gamers and the big volume will be in AR. To understand these two related, but separate markets, here is a breakdown of the two:

**AR vs VR:**

- VR—completely covers the user’s eyes and displays a computer-generated display in the user’s field of vision. The user is completely immersed in the computer-generated world. One of the reasons VR has not taken off is that it isolates the user from both the environment as well as prevents sharing the experience with others. While 3D will help fix the first problem, the personal isolation involved with VR will likely keep it to a niche market.
- AR—the user can “see through” the display to the real world, and computer-generated images are superimposed or projected. The real world and computer-generated graphics are mixed, giving raise to the alternative phrase “mixed reality.” In AR, the “real world” as well as other people are still accessible to the user. In addition, it is much easier for users to share the same experience if, for example, the same projection is provided on two users’ devices. The users can both see each other as well as the virtual object from their own perspective, giving them the ability to share the experience.

### **11.5.5 Further VR/AR segmentation**

Both VR and AR devices need processing power, and that processing can either be a part of the unit, making it “mobile,” or it can be required to be tethered to a PC or other dedicated processing unit that is stationary, making it tethered. If we back up and call VR and AR as “head-mounted displays” (HMD), the market and market potential can be segmented as follows:

- Tethered VR—examples: Oculus Rift, HTC Vive, PlayStation VR—these VR units require a dedicated PC or gaming console and are aimed at gamers. While Oculus got lots of press attention with its acquisition by Facebook, Sony’s PlayStation VR has jumped to the market lead with an estimated 3 million units sold to date. Due to its limited appeal as described above, this market is seen as a niche market for gamers, and perhaps some units into enterprise for deep simulation and training.
- Mobile VR—examples: Google Cardboard, Samsung VR—in this use case the mobile phone is the video platform for VR. The phone is placed in a holder for the head—which can be as simple as a piece of cardboard or more elaborate to add more features—and the user as a VR experience. Despite pushes by Google, Samsung, and others, the acceptance and penetration rate is still rather small. Other than gaming—which has better experiences on a tethered unit—there is no “killer app” for this segment. And since it both isolates the user plus creates an awkward appearance in public, people are not using mobile VR during “linger time” with their cellphone. This is seen a niche segment in the years to come as mobile VR takes this place.
- Console AR—examples: Microsoft Hololens, Daqri Smart Classes—console AR systems are high-end systems costing thousands of dollars targeted at enterprise users. With either a dedicated PC or high-end dedicated embedded processor, these units are out of reach for consumers but are bought by corporations for training, field work, and other work applications. Companies selling these units make more money on the software, support and subscription than they do on the units, and total volume will be in the 100 s of K, perhaps going into the low millions if picked up by a large number of enterprises with lots of field people (FedEx, etc.).

Mobile AR—this is discussed in the mobile section, and is seen by many as the path that mobile phones will take as users are “always on.” This will be the main AR market, and if adopted will push the other HMD technologies further into niche markets.

### **11.5.6 Why 3D for VR**

As discussed in the mobile AR section, users with HMD want to use their hands to interact with the computer-generated environment, and 3D sensors are the best way to provide gesture control as well as map the “real world” surroundings and bring it into the virtual world, either for user interaction or to, say, warn a VR wearer that they are getting too close to a chair as they play their game in the virtual world.

Today’s VR systems are still using hand controller to manipulate within the virtual world, and usually 2D cameras or other sensors to warn users if they are getting too close to real objects. However, starting in the 2020 timeframe several VR systems will start incorporating depth sensors into the system to capture the hands and real-world objects, bringing virtual version of the user’s hands into the VR world.

### **11.5.7 3D players for VR players and market potential**

The two main VR players—Sony and Oculus—effectively have their own internal 3D sensing technology. Sony of course will lean toward using Sony sensors, most likely their ToF product line. Oculus has its own internal structured light division and would likely lean toward that as a solution, but would still need to get the actual silicon sensors from a silicon house. Microsoft’s Hololens also uses its own internal ToF sensor from its Canesta acquisition.

### **11.5.8 Robotics**

Robotics is one of the oldest, well-established markets for 3D imaging. To date robotics have been driven by industrial and factory automation concerns, which having the cost points, investment dollars, and cycle times to help develop early 3D imaging technology.

The robotics market to date has been small compared to consumer volume standards, and it is largely serviced by specialty 3D and machine vision companies, and largely not serviced or supported by the 3D companies outlined here. However, the push of robotics into lower end units for retail, and even lower end markets for consumer, is both creating an extended market for robotics, and bringing the consumer 3D players into this space.

Right now, there are three main segments developing:

- Industrial/factory automation—the traditional robotics market, with volumes in the 100 s of K, creating robots that is in the thousands to millions of dollars.
- Service and retail—these include service facing robots (which are effectively an extension of a kiosk) as well as work reduction machines which are being developed for cooking and automation of fast food and other retail. Most vendors are adopting 3D vision as a part of their roadmap to better interact with customers and the environment. Overall volume short term will be in the 100 s of thousands per year, but could expand rapidly in the coming years if this model is accepted by consumers. Consumers of Japan, for example, will likely accept it faster than those of the United States.

- Home robotics sub segments—robots bought by consumers are seen as a growing segment, and can be subdivided into two areas:
  - Work reduction (vacuuming, lawn mowing, etc.)—vacuum robots are the largest within this category at about 2 million units a year. Vacuum robots today primarily use ultrasound sensors but are slowly migrating to 3D as it gets cheaper and the robots become more advanced and map the environment (navigation) instead of just avoiding obstacles when they get nearby (random path).
  - Companion/entertainment—companion robots are seen as a growing segment over the next decade and will use 3D sensors if they are to both navigate as well as interact and recognize gestures with humans. They are starting to ship now but will take some time to be mass adopted.

### **11.5.9 3D robotics opportunity in next 5 years**

The only market that matters for real volume in robotics is consumer robots, and that market is still some ways out before it sees >1M units a year in 3D sensors.

The industrial robotics space will provide steady volume and moderate growth to a wide range of specialty companies, with retail and service robots providing an additional few hundred K annual volume opportunity growth over the next few years, but 3D providers will be very fragmented for this segment.

Within consumer robots, vacuum robots will see increased use of 3D, but at penetration rates well below 50% of the vacuum robot total available market, making that market opportunity <1M/year by 2022. Low-end structured light and the ToF players will pick up market share here.

Companion robots will be a good long-term growth prospect for 3D, but market adoption is speculative and probably no mass market adoption in next half decade. These robot designers will try to ride the mobile 3D module supply chain, and buy from the same structured light and ToF vendors supplying into mobile.

### **11.5.10 Drones (unmanned aerial vehicles or UAV)**

Penetration of 3D imaging into drones is currently limited, but is increasing as autonomous flying capabilities are added to drones, similar to autonomous cars. Adding 3D sensing will allow the computer take over completely for tasks like takeoff, landing, and obstacle avoidance. DJI had two of the first drones in the market with 3D: the DJI Mavik Air with stereoscopic vision, and the Phantom 4 which used two ToF sensors.

Like exterior automotive, drones are looking at two different use cases for classes of 3D sensors:

- Close range/slow speed—Less than 5 m, for when the drone is hovering, landing, or slow moving. ToF and stereo are being utilized for this use case.
- Long range/fast speed—More than 5 m and as far as 30 m. The only real option at this point is stereo for this use case.

3D is superior to color sensors for doing these tasks, but the current 3D cost vs range and power are limiting their penetration. This should improve in the next 2–3 years.

The extended goal for drones is to have “selfie drones” that follow people automatically, as well as interact using gesture control, but this will be some time out. These types of units are effectively flying autonomous vehicles, and will have the same 3D sensor needs as autonomous cars.

Current human following drones rely on GPS, which is inaccurate, or tracking using a wearable band or phone on the person rather than vision. These create acceptable experiences, but would be improved by full 3D automation.

3D sensors will find themselves in the mid to high-end units that offer higher levels of automation. As automation increases, the penetration into drones by 3D will increase. Also, due to the multiple modules per unit, the unit rate will be faster than the penetration rate.

Looking at the numbers, the 2017 drone market is estimated at about 3 million units, of which there were about 500K 3D sensors shipped in a dual-sensor setting, so 250K drones had 3D in 2017, or about an 8.3% penetration rate. The drone market will likely flatten out now that the novelty has worn off, plus increased regulation for hobbyists and prosumer users will flatten that segment growth.

Commercial drones will have a high growth rate but off a small base, and delivery drones are likely >5 years out for any significant volume. This segment will be dominated by stereoscopic, but will have some ToF penetration.

### **11.5.11 *Other markets in development***

There are several markets doing R&D with 3D today which will not be large in the short term, but could be significant segments on the long-term horizon.

### **11.5.12 *Retail***

The first are a variety of use cases in retail, which is looking to provide more automation and tracking to improve and streamline the user experience.

- Automated shopping/customer tracking—Amazon Go and other players are developing shopping experience that have no checkout, but require tracking the shopper and shelf. This development uses 3D sensors throughout the store to track where the customer is and what they’ve purchased. Test stores are already open, but will be some years out before deployment as well as mass consumer acceptance.
- Retail analytics/shelf monitoring—one step back from automated customer checkout, this is monitoring customer and shelves to track habits, similar to tracking a user on a website (where they browse, if they take an item and put it back, etc.). Alternatively, there are shelf scanning robots already in place that work usually when the store is closed to check stocking levels of the shelf, replacing people who did this previously.
- People counting (malls, stores, etc.)—this is already in use today with very low end sensors or even turnstiles. 3D would improve the tracking plus be able to add some demographics (child, etc.) but keep privacy since 3D does not disclose personal details.
- Clothes/shoe fitting—this area uses 3D scanners to measure people for clothes, shoes, glasses, etc. Several niche companies, start-ups and even a few stores already exist, but will not move into mainstream for another 5–10 years.

### 11.5.13 IoT/home appliances

The IoT is a buzzword meant to mean a variety of things, most commonly any consumer device that is hooked up some way to the internet and to the cloud. Today, home assistants such as Amazon's Alexa and Google Home are getting a foothold as the IoT "hub," from which other devices in the home will be controlled by voice.

One of the areas these assistants will expand to in the home is security and people monitoring, for doing things like automatically changing the thermostat or lighting depending on who is in what room. This type of functionality will require 3D sensors to provide an accurate representation of who is in the home, if they are a resident or intruder, and controlling devices and environment based on that person's profile.

The first IoT device with 3D doing many of these things is LightHouse. Their home security unit includes a 3D sensor ToF sensor from pmd to allow segmentation and identification of objects and individuals. Home devices from Amazon and Google are expected to follow with 3D sensor in the coming years, and this could turn into a significant market for 3D by 2022.

### 11.5.14 What else?

Industrial security, medical, and many other minor segments are all developing with 3D as segments become more autonomous, and a better interpretation of the real world is required that is not available using regular 2D sensors.

## References

- Acerbi, F., 2017. Transfer-gate region optimization and pinned-photodiode shaping for high-speed TOF applications. In: 2017 International Image Sensor Workshop (IISW), Hiroshima, Japan.
- Acerbi, F., 2018. Optimization of pinned photodiode pixels for high-speed time of flight applications. *IEEE J. Electron Devices Soc.* 6, 365–375.
- Al Abbas, T., 2016. Backside illuminated SPAD image sensor with  $7.83\mu\text{m}$  pitch in 3D-stacked CMOS technology. In: 2016 IEEE International Electron Devices Meeting (IEDM), San Francisco, CA, United States.
- Bamji, C., 2000. Systems for CMOS-Compatible Three-Dimensional Image Sensing Using Quantum Efficiency Modulation. US 6580496B2.
- Bamji, C., 2015. A  $0.13\mu\text{m}$  CMOS system-on-chip for a  $512 \times 424$  time-of-flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC. *IEEE J. Solid State Circuits* 50, 1.
- Bamji, C., 2018. 1Mpixel 65 nm BSI 320MHz demodulated TOF image sensor with  $3.5\mu\text{m}$  global shutter pixels and analog binning. In: IEEE International Solid-State Circuits Conference, San Francisco.
- Böhmer, S., 2016. Integrated time-of-flight technology—a scalable solution for depth sensing. In: 3rd SEMICON Europa Imaging Conference, Grenoble.
- Brockherde, W., 2012. Solid-state ToF sensors. In: 6th Fraunhofer IMS Workshop CMOS Imaging Low-Light Imaging, Duisburg, Germany.
- Büttgen, B., 2007. Extended Time-of-Flight Optical 3D-Imaging to Extreme operation Conditions. Sierke Verlag, Göttingen, Germany.

- Durini, D., 2010a. Lateral drift-field photodiode for low noise, high-speed, large photoactive-area CMOS imaging applications. *Nucl. Instrum. Methods Phys. Res. A* 626 (2), 470–475.
- Durini, D., 2010b. Lateral drift-field photodetector based on a non-uniform lateral doping profile photodiode for time-of-flight imaging. In: 6th Conference on Ph.D. Research in Microelectronics & Electronics (PRIME 2010), Berlin, Germany.
- Fossum, E.R., 2016. The quanta image sensor: every photon counts. In: Photon-Counting Image Sensors, A special issue of *Sensors*. 1424–8220.
- Frey, J., 2007. Entwurf und Untersuchung von hochauflösenden 3D-Bildsensoren in CMOS-Technologie. Ph.D. Dissertation University of Siegen, Germany.
- Gokturk, S.B., 2004. A time-of-flight depth sensor—system description, issues and solutions. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop, Washington, DC.
- Hoffmann, B., 1997. Einsteins Ideen. Spektrum Akademischer Verlag.
- Iddan, G.J., 2001. 3D imaging in the studio (and elsewhere ...). In: Photonics West 2001—Electronic Imaging, Proceedings of SPIE 4298, San Jose CA.
- Janesick, J.R., 2007. Photon Transfer DN to Lambda. SPIE Press, ISBN: 9780819467225.
- Kappner, H.A., 1986. Verfahren und Anordnung zur dreidimensionalen optischen Erfassung von Objekten. European Patent No. 0 192 993 A1.
- Kim, S.-J., 2011. A  $640 \times 480$  image sensor with unified pixel architecture for 2D/3D imaging in  $0.11\mu\text{m}$  CMOS. In: 2011 Symposium on VLSI Circuits—Digest of Technical Papers, Kyoto, Japan.
- Kim, S.-J., 2012. Image Sensor Based on Depth Pixel Structure. US 20120175685 A.
- Koechner, W., 1968. Optical ranging system employing a high power injection laser diode. *IEEE Trans. Aerosp. Electron. Syst. AES-4* (1), 81–91.
- Lange, R., 1999. Time-of-flight range imaging with a custom solid-state image sensor. Proc. SPIE 3823, 180–191.
- Lange, R., 2000. 3D Time-of-Flight Distance Measurement With Custom Solid-State Image Sensors in CMOS/CCD-Technology. Ph.D. Dissertation, University of Siegen, Germany.
- Oh, M.S., 2012. Backside-illumination  $14\mu\text{m}$ -pixel QVGA time-of-flight CMOS image. In: 10th IEEE International NEWCAS Conference.
- Rae, B., 2017. Fully Industrialised Foundry SPAD in an Optimised 130nm CMOS Imaging Technology. Image Sensors Europa, London.
- Rodrigues, B., 2017. Indirect ToF pixel integrating fast buried-channel transfer gates and gradual epitaxy, and enabling CDS. In: 2017 International Image Sensor Workshop (IISW), Hiroshima, Japan.
- Schwarte, R., 1995. A new active 3D-Vision system based on rf-modulation interferometry of incoherent light. In: Photonics East-Intelligent Systems and Advanced Manufacturing, Proc. SPIE, vol. 2588, p. 126.
- Schwarte, R., 1997a. New electro-optical mixing and correlation sensor: facilities and applications of the photonic mixer device (PMD). In: Proceedings of Sensors, Sensor Systems and Sensor Data Processing, Munich, Germany. vol. 3100.
- Schwarte, R., 1997b. Pseudo-noise (PN) laser radar without scanner for extremely fast 3D-imaging and navigation. In: MIOP '97, Stuttgart.
- Spirig, T., 1995. The lock-in CCD—two-dimensional synchronous detection of light. *IEEE J. Quantum Electron.* 31 (9), 1705–1708.
- Theuwissen, A., 2018. Time-of-Flight for 3D-Imaging. Harvest Imaging, London, UK.
- Van Nieuwenhove, D., 2005. Novel standard CMOS detector using majority current for guiding photo-generated electrons towards detecting junctions. In: Proceedings of Symposium IEEE/LEOS, Mons.

- Von Breuer, V., 1990. dtv-Atlas zur Physik, Tafeln und Texte', Bd.1, dtv, ISBN 3-423-03226-x.
- Xu, Z., 1998. Investigation of 3D-Imaging Systems Based on Modulated Light and Optical RF-Interferometry (ORFI). Ph.D. Dissertation, University of Siegen, Germany.
- Yole1, 2017. 3D imaging and sensing—2017 report. In: YOLE Développement, Lyon, France.
- Yole2, 2017. Camera module industry—2017 report. In: YOLE Développement, Lyon, France.

## Further reading

- Bellisai, S., 2013. Single-photon pulsed-light indirect time-of-flight 3D ranging. *Opt. Express* 21 (4), 5086.
- Kim, S.-J., 2010. A CMOS image sensor based on unified pixel architecture with time-division multiplexing scheme for color and depth image acquisition. *IEEE J. Solid State Circuits* 47(11).
- Lange, R., 2001. Solid-state time-of-flight range camera. *IEEE J. Quantum Electron.* 37 (3), 390–397.
- Lee, T.Y., 2011. A  $192 \times 108$  pixel ToF-3D image sensor with single-tap concentric-gate demodulation pixels in  $0.13\text{ }\mu\text{m}$  technology. In: 2011 IEEE Electron Devices Meeting (IEDM), Washington, DC.

This page intentionally left blank

# CMOS sensors for fluorescence lifetime imaging

12

Robert K. Henderson<sup>a</sup>, Bruce R. Rae<sup>b</sup>, Day-Uei Li<sup>c</sup>

<sup>a</sup>University of Edinburgh, Edinburgh, United Kingdom, <sup>b</sup>ST Microelectronics, Edinburgh, United Kingdom, <sup>c</sup>University of Strathclyde, Glasgow, United Kingdom

## 12.1 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 devices (EMCCD), intensified CCD (ICCD), or most recently scientific CMOS (sCMOS) sensors (Hynecek, 2001; 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 CMOS or 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 potential to assist researchers in improving our understanding of the processes underlying normal cellular function, and their alteration in disease states.

### 12.1.1 Outline

The intense scientific interest in FLIM resides in the fact that the fluorescence lifetime of a fluorophore depends on its molecular environment and not on its concentration (Lakowicz, 2006). As a result, effects at molecular scale can be studied independent 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 is currently acquired by two main techniques broadly classified into time-domain and frequency-domain techniques. Time-domain FLIM is performed by time-correlated single-photon counting (TCSPC) or by gated image intensifiers. Frequency-domain FLIM uses gain-modulated photomultiplier tubes (PMTs) 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 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 integrated circuits capable of performing high-speed digital computation are also 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 are enabled entirely by solid-state implementations.

### **12.1.2 Sources of information**

A comprehensive textbook on FLIM techniques was published in 2025 by Wolfgang Becker entitled “Advanced Time-Correlated Single-Photon Counting Techniques” (Becker, 2015). This book covers all aspects of the technology and applications of time-resolved imaging. An excellent review article (Suhling et al., 2015) also contains valuable material including some more recent solid-state sensor implementations. The standard reference in the wider area of fluorescence techniques is by Lakowicz entitled “Principles of Fluorescence Spectroscopy” (Lakowicz, 2006). A textbook has appeared recently surveying all single-photon imaging technologies including CMOS edited by A. Theuwissen and P. Seitz entitled “Single-Photon Imaging” (Theuwissen and Seitz, 2011). Much useful information can be found on the websites of the various time-resolved imaging equipment vendors, for example, B&H, Picoquant, Andor, Lambert Instruments, LaVision, Hamamatsu, and Olympus.

## **12.2 Fluorescence lifetime imaging**

### **12.2.1 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 the imaging of specific labeled biological samples, more recently with the advent of deoxyribonucleic acid (DNA) sequencing there has been considerable interest in microarray 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.

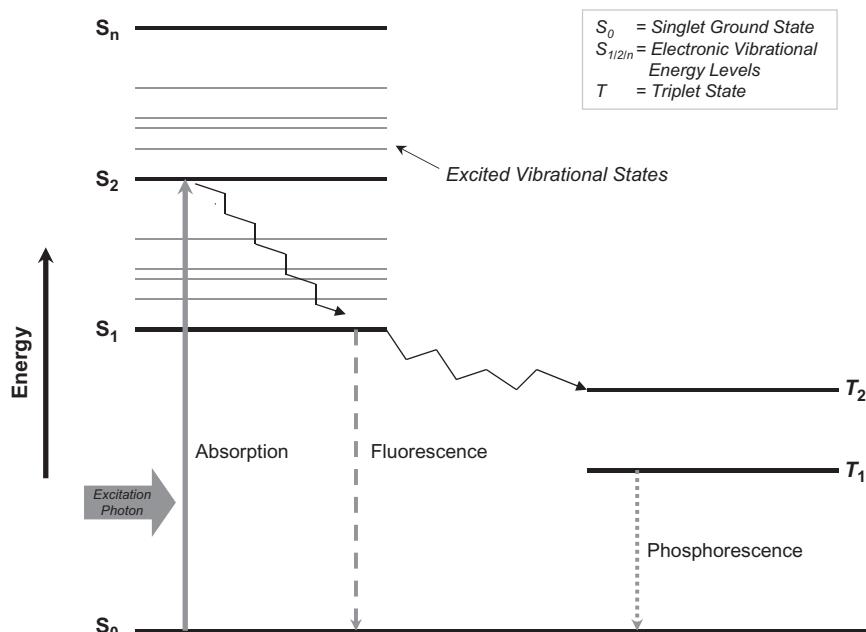
Time-resolved fluorescence analysis is the measurement of the temporal properties of a fluorophore sample. Fluorescence lifetime detection provides a method for differentiating between spectrally overlapping samples which exhibits different lifetime

properties (Cubeddu et al., 2002). The sensitivity of a sample's lifetime properties to the microenvironment 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.

### 12.2.2 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. 12.1).

A sample's fluorescence lifetime,  $\tau$ , is determined by the rate at which the sample leaves the excited state (Eq. 12.1). The transition can occur via two mechanisms, either by fluorescence emission (at rate  $\Gamma$ ) or by competing non-radiative processes (represented collectively as  $K_{nr}$ ).



**Fig. 12.1** Jablonski diagram. An electron can be promoted 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.

$$\tau = \frac{1}{\Gamma + \sum K_{nt}} \quad (12.1)$$

A fluorophore's quantum yield ( $\Theta$ ) is the ratio of emitted photons to the number of absorbed photons. This is represented by Eq. (12.2).

$$\Theta = \frac{\Gamma}{\Gamma + \sum K_{nt}} \quad (12.2)$$

For a given excitation light intensity, a fluorophore's brightness (molecular brightness,  $q$ ) can be calculated if the molecular absorption coefficient ( $\epsilon$ ) is known, Eq. (12.3).

$$q = \epsilon \times \Theta \quad (12.3)$$

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 Eqs. (12.2) and (12.3), if the fluorescence intensity changes this will usually result in a change in sample lifetime. Due to the fact fluorescence intensity is a composite property of a sample, dependent on sample quantity and concentration as well as instrument setup, 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 are as follows (Lakowicz, 2006):

- intersystem crossing
- collisional or static quenching
- solvent effects
- resonance energy transfer

Fluorescence intensity is related to lifetime according to Eq. (12.4) (for a mono-exponentially decaying sample). The equation assumes that the sample has been excited by an infinitely short ( $\delta$ -function) light pulse. The time-dependent intensity at time  $t$ ,  $I(t)$ , is given by.

$$I(t) = I_0 \exp\left(\frac{-t}{\tau}\right) \quad (12.4)$$

Fluorescence lifetime is independent of fluorophore concentration but dependent on the sample's local environment. Thus, lifetime detection allows precise quantitative data about both 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 labeling (Pepperkok et al., 1999), quantitative detection of ion concentrations and oxygen and energy transfer characteristics using fluorescence resonance energy transfer (FRET) (Prasad, 2003).

### 12.2.3 Lifetime measurement techniques

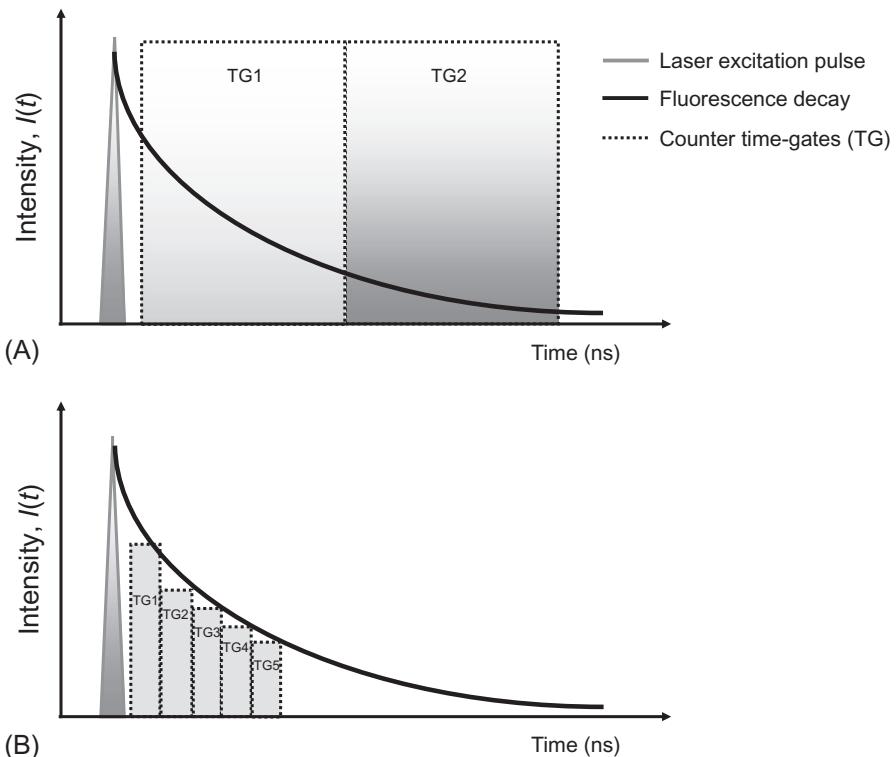
There are two main techniques for measuring the fluorescence lifetime of a sample: the frequency-domain and time-domain methods. In the frequency domain a sample is excited by an intensity modulated light source. The fluorescence emission is modulated at the same frequency, but with a phase shift due to the intensity decay law (Eq. 12.4) of the sample (Lakowicz, 2006; Chodavarapu et al., 2005) and a reduction in the modulation depth. In the time domain the intensity decay of a fluorescent sample is directly measured as a function of time, following absorption of a short excitation pulse.

The FLIM is achieved by two methods: wide-field imaging and point scanning (detector or light source). Wide-field imaging makes use of 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 one-pixel image 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 synchronization pulse from the pulsed light source (Fig. 12.2). 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) 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 minimizes 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: MCP-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

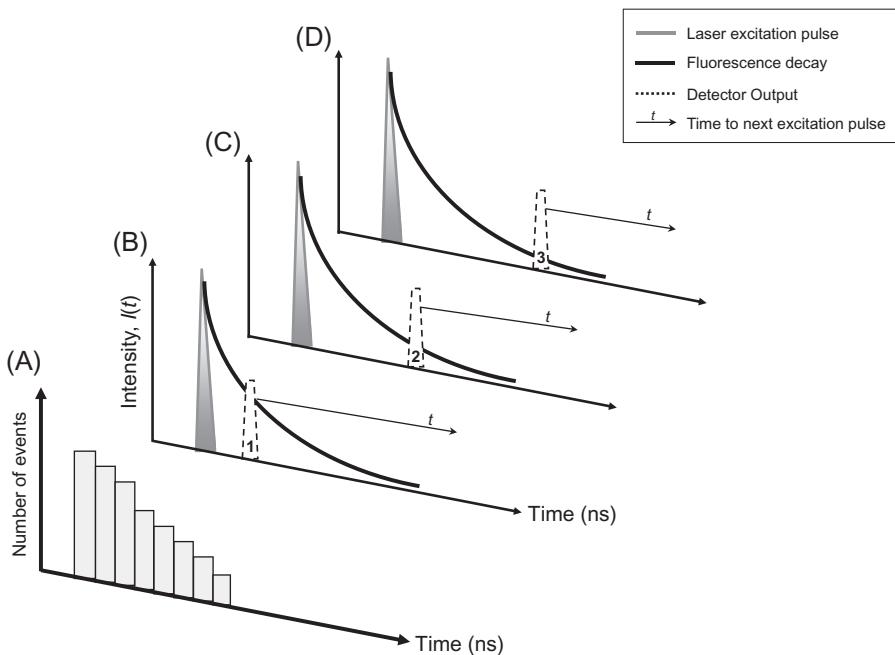


**Fig. 12.2** Reverse start-stop TCPSC principle. The sample of interest is excited by a short laser pulse and the time between the first fluorescence photon detected and the subsequent excitation pulse is recorded. A histogram of photon arrival time (A) is generated by repeating this measurement many times (B–D).

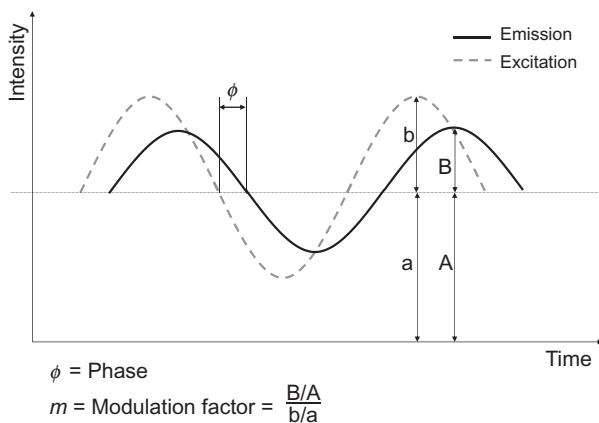
and delicate. These devices achieve single-photon sensitivity through electron multiplication, triggered by an initial electron-hole pair caused by an incident photon.

An alternative method of capturing fluorescence lifetime information in the time domain is to use gated detection (Fig. 12.3). 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 two-dimensional (2D) detector. Such a system achieves electron multiplications between multi-channel plates situated



**Fig. 12.3** Time-Gated FLIM principle. Fluorescence decay is captured using a series of two (A) or multiple time gates (B).



**Fig. 12.4** Frequency-domain fluorescence lifetime principle.

in front of the imager. Time gating is achieved by providing a short electrical pulse which enables the multi-channel plates only during the period of interest.

Fig. 12.4 shows the principle of frequency-domain fluorescence lifetime estimation. A sample is excited through a sinusoidal illumination source [usually a light-emitting diode (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 ( $\tau_\phi$ ) and modulation ( $\tau_m$ ) can be derived from measurements of the phase shift and the modulation depth, respectively. The phase shift ( $\phi$ ) between the excitation and emission signals is a function of the excitation frequency ( $\omega$ ), and varies over the range 0–90 degrees across excitation frequencies. The modulation lifetime can be calculated from the difference between the emission and excitation amplitudes (Eq. 12.7). 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 the range 0–1 across excitation frequencies according to Eq. (12.6).

$$\tau_\phi = \frac{\tan(\phi)}{\omega} \quad (12.5)$$

$$m = \frac{A_{em}I_{ex}}{I_{em}A_{ex}} = \frac{1}{\sqrt{1 + \omega^2\tau_m^2}} \quad (12.6)$$

$$\tau_m = \frac{1}{\omega} \sqrt{\frac{1}{m^2} - 1} \quad (12.7)$$

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).

#### 12.2.4 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 time resolution can be obtained (in the order of picoseconds) (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 miniaturization. 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 onto on-chip capacitors. Currently, 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 (femtoseconds 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 microsystem 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.

### 12.2.5 Applications

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 are as follows:

- 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  $\text{Ca}^+$  or  $\text{Cl}^-$  concentration can be observed. Probes such as Quin-2 demonstrate dramatic changes in lifetime in response to changes in local  $\text{Ca}^{2+}$  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, by 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.
- Fluorescence 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 which exhibit fluorescence lifetime changes in the presence of nitro-aromatic vapor. 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).

## 12.3 CMOS detectors and pixels

### 12.3.1 Introduction

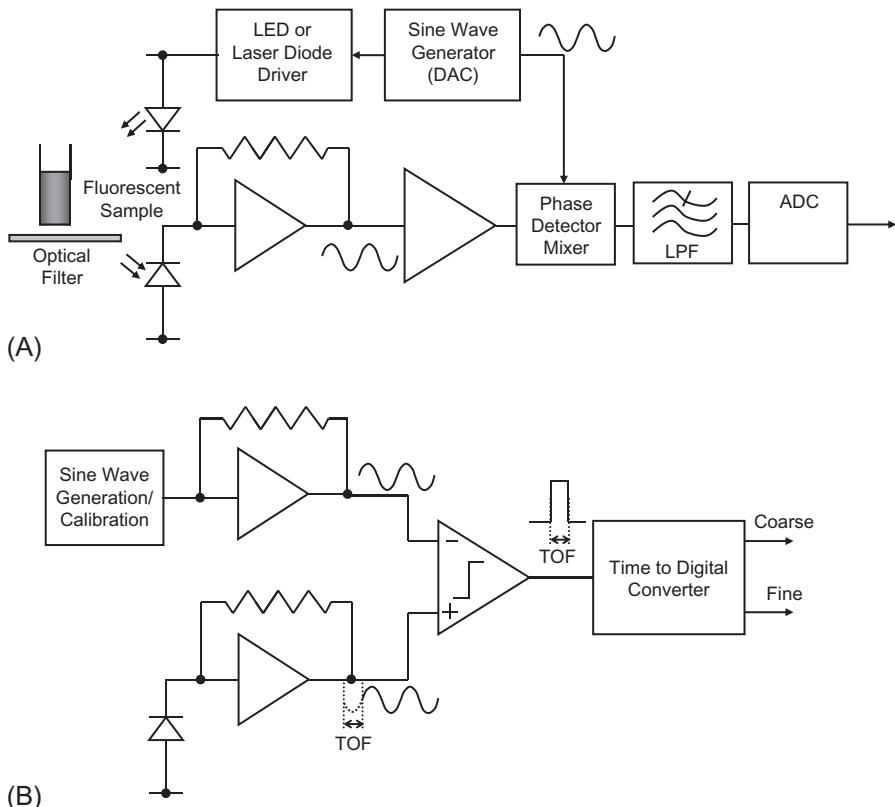
Active research on CMOS sensors for fluorescence lifetime started around 2005 spurred by the emergence of new detectors such as SPADs or the availability of low-noise pinned-photodiodes from consumer CMOS image sensors (Theuwissen, 2008). Prior development of very high-speed image sensors based on gated 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 few 100 ps 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 three-dimensional (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 toward 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 miniaturization afforded by CMOS sensors compared to GII and 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, b; 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, 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.

### 12.3.2 Frequency-domain lifetime pixels

The recovery of the lifetime from the phase shift ( $\varphi$ ) has been accomplished by techniques inspired by homodyne or direct conversion receivers (Razavi, 2007). A phase fluorometric oxygen sensor has been realized in 1.5- $\mu\text{m}$  CMOS (Chodavarapu et al., 2005) based on a large  $16 \times 16$  phototransistor array measuring  $720 \mu\text{m} \times 720 \mu\text{m}$  (Fig. 12.5A). 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 kHz range. This is acceptable for the few microsecond lifetime of the oxygen-sensitive Ruthenium complex  $[\text{Ru}(\text{dpp})_3]^{2+}$  doped xerogel-film. The phototransistor peak



**Fig. 12.5** Phase-fluorescent sensor: (A) phase demodulation pixel and (B) delay digitization pixel.

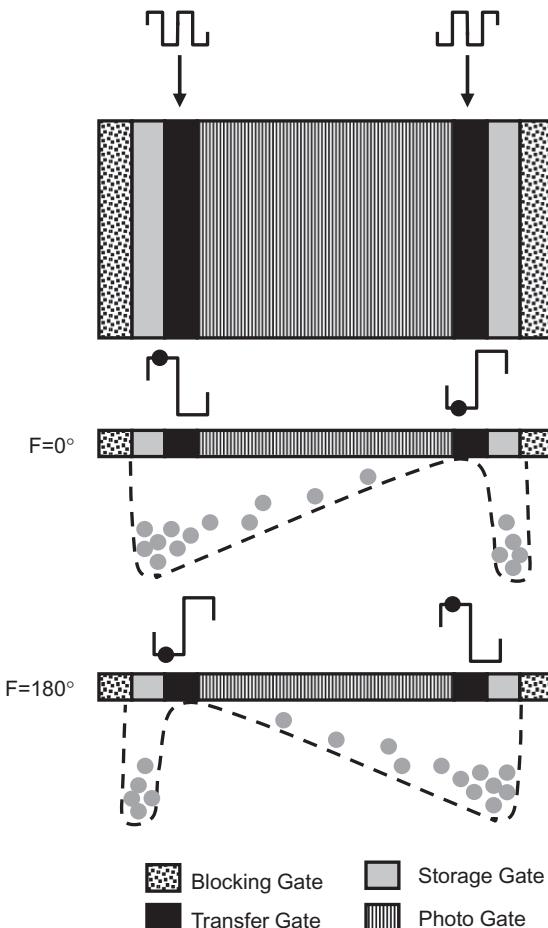
responsivity of 675 nm has been chosen close to the emission of the xerogel of 595 nm. This sensor is the first to integrate the entire receiver consisting of current-to-voltage converter, amplifier, band-pass filter, and phase detector.

A similar system has been realized in a modern 65-nm CMOS process with direct digital phase measurement (Guo and Sonkusale, 2011a, b). They replace the phase detector by a comparator and TDC, measuring phase directly as the time offset between the excitation signal and amplified, quantized fluorescence signal (Fig. 12.5B). The same team has reused this architecture in the form of a  $32 \times 32$  pixel image sensor for frequency-domain FLIM (Guo and Sonkusale, 2011a, b). Passive pixels are employed with a 50- $\mu\text{m}$  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 time-to-digital converter (TDC). The TDC has a 110-ps temporal resolution over a 414- $\mu\text{s}$  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 parallelized 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 developed for scientific FLIM. It comprises  $213 \times 212$  pixels at  $17\text{ }\mu\text{m}$  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 TOF image sensors which have been demonstrated in CMOS implementations (Lange and Seitz, 2001; Oggier et al., 2004; Kawahito, 2007) and demonstrated to be suitable for FLIM (Esposito, 2011). As shown in Fig. 12.6, fields under the transfer gates created by



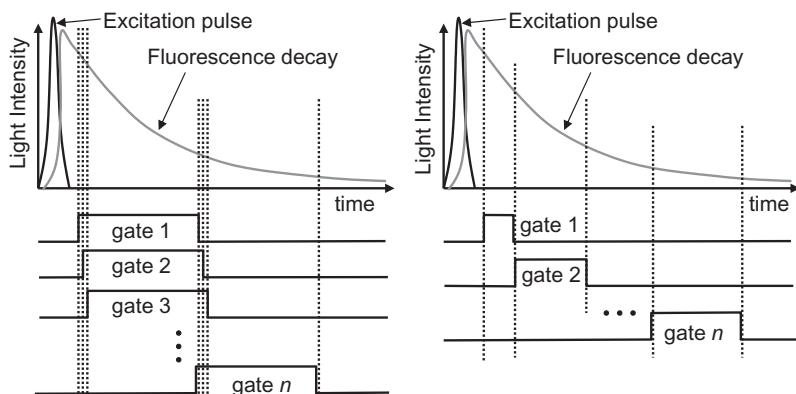
**Fig. 12.6** Modulated electron multiplied (MEM)-FLIM pixel and operating principle.

modulated voltages divert photogenerated carriers toward the STGs 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 STGs. The sensitivity of the 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.

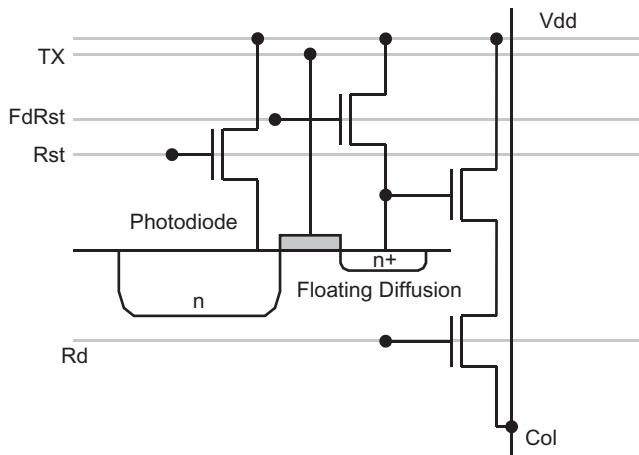
### 12.3.3 Time-gating pixels

Time gating is generally implemented using the two timing schemes shown in Fig. 12.7A. 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 (few 100 ps). 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. 12.7B. Here, multiple nonoverlapping time gates are used to sample the fluorescent decay. The most 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 equalize 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 generalized to multiple gates (Grauw and Gerritsen, 2001).

An early implementation of gated image sensors for FLIM employed large  $100\text{ }\mu\text{m} \times 100\text{ }\mu\text{m}$  photodiodes in a  $0.25\text{-}\mu\text{m}$  CMOS technology with a current mode



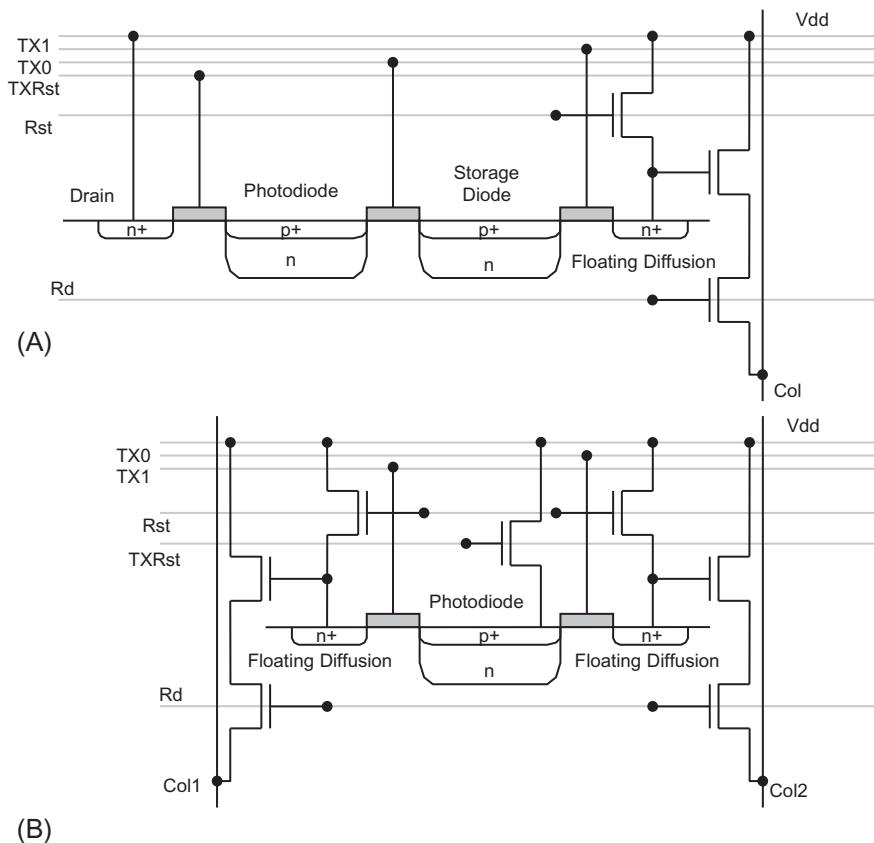
**Fig. 12.7** Gating schemes: (A) sliding window and (B) multiple gate.



**Fig. 12.8** Gated photodiode.

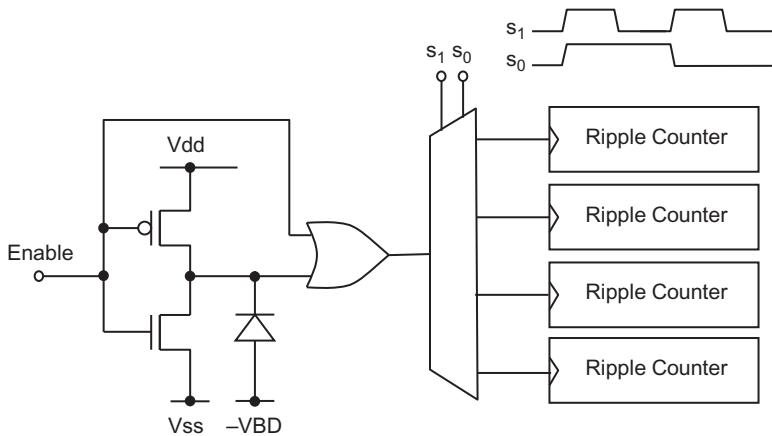
readout (Patounakis et al., 2006). The pixel shown in Fig. 12.8 samples the transient waveform induced by the photodiode through a gating signal TX onto a floating diffusion node. An innovative feature of this sensor was the improvement of signal to background ratio by draining carriers generated during 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 et al., 2011) schemes (Fig. 12.9). 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 CDS (Fig. 12.9A). The two-gate implementation (Fig. 12.9B) 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 and 7.5  $\mu\text{m}$  pitch and  $256 \times 256$  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.

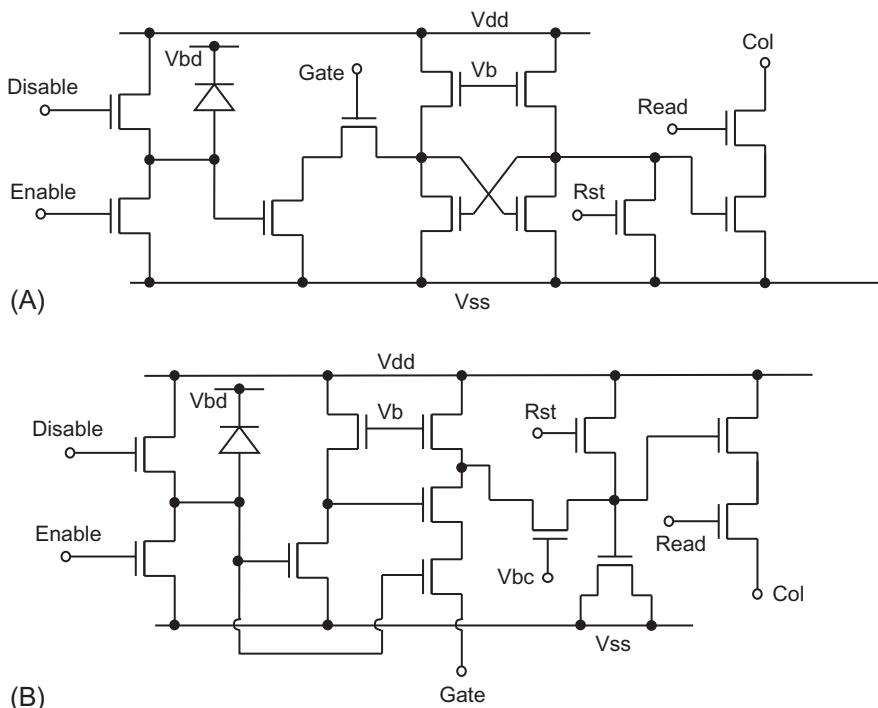


**Fig. 12.9** Gated pinned-photodiode implementations: (A) multi-gate and (B) sliding gate.

The SPADs provide both extremely high gain ( $>100$  k for few micron diameter detectors) and fast response speed ( $\sim 100$  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 realize this potential and proposed a number of gated SPAD pixel realizations (Schwartz et al., 2008a, b; Pancheri and Stoppa, 2008). A generic digital implementation is shown in Fig. 12.10. The Gate input to the circuit inhibits the SPAD by pulling the anode to ground while maintaining the output in a high state. Only if the SPAD triggers on a photon arrival when the quench 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 is 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.



**Fig. 12.10** Digital gating pixel of a single-photon avalanche diode.



**Fig. 12.11** Gated analog NMOS-only SPAD pixels: (A) SRAM pixel and (B) photon counting.

Two sensors composed of NMOS-only, time-gated SPAD pixels of 25  $\mu\text{m}$  pitch with fill factors of 4.5% and 20.8%, respectively, have been proposed recently in a 0.35- $\mu\text{m}$  HV CMOS (Maruyama and Charbon, 2011; Pancheri et al., 2011) (Fig. 12.11). 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 logic inversion. The pixel in Fig. 12.11A employs a single bit in-pixel 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 hybridization 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. 12.11B 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 NAND gate shortens the SPAD pulse to the enable 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. The analog photon counting and single bit approaches have been pursued with steadily increasing image resolution and fill factor while approaching practical pixel pitches for microscopy. Today SPAD array sizes of  $512 \times 512$  and pixel pitch of  $8 \mu\text{m}$  have been achieved with up to 50% fill factor through microlensing or efficient shared pixel layout schemes (Dutton et al., 2016; Perenzoni et al., 2016; Burri et al., 2014; Ulku et al., 2019). Sliding or multiple gates are applied to these imagers to generate fluorescence lifetime images in few second time spans. The integration of a process-insensitive single slope ADC by Perenzoni et al. (2016) is an attractive feature allowing greater dynamic range and faster acquisition of images. A good review of such analog pixels has been published recently (Perenzoni et al., 2016).

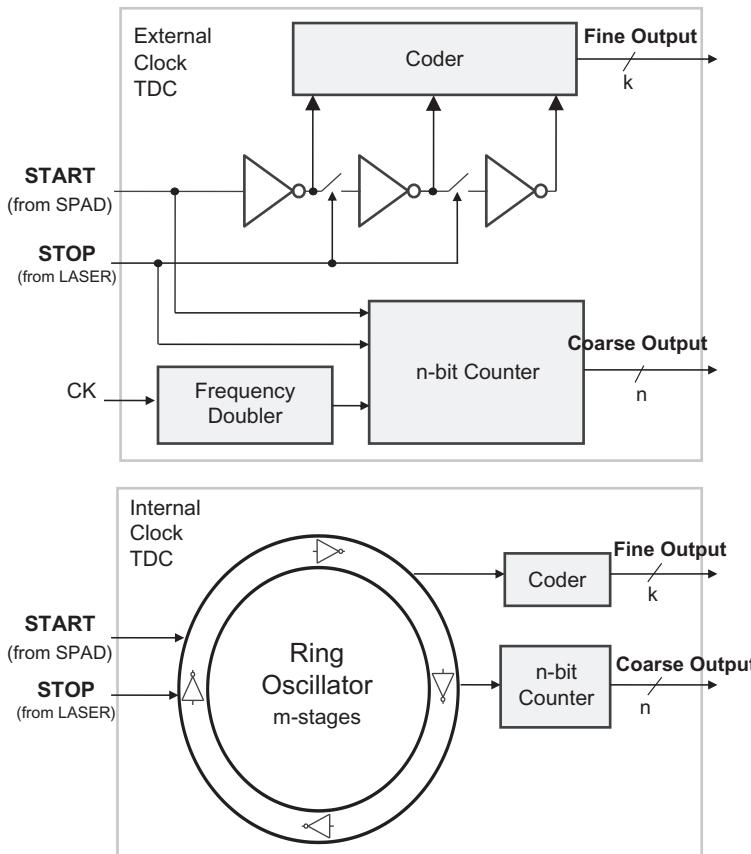
### 12.3.4 TCSPC pixels

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 synchronization signal. The first implementations of time-resolved SPAD imagers employed off-chip timing circuits or multiplexed column-parallel TDC (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  $50\text{-}\mu\text{m}$  pixel circuits were designed (Richardson et al., 2009a, b; Stoppa et al., 2009; Gersbach et al., 2012). The small pitch, high throughput and moderate time resolution of these pixels were 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., 2009a, b; 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. 12.12A). 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. 12.12B). The propagation



**Fig. 12.12** MegaFrame TDC pixel architectures: (A) external clock and (B) internal clock.

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 at all illumination levels. This architecture provides good global time resolution accuracy. Characterization resulting from a  $160 \times 120$  IC-TDC array has shown good TDC uniformity (Veerappan et al., 2011).

The IC-TDC architecture was pursued by a number of other researchers with improving fill factor and reduced power consumption (Vornicu et al., 2014; Gasparini et al., 2018). A hybrid IC/EC-TDC architecture allowing the ring oscillators to be injection locked for improved jitter and uniformity was recently proposed (Ximenes et al., 2017). A very comprehensive survey of ring oscillator-based TDCs was undertaken recently (Cheng et al., 2016).

TDC pixels suffer from high peak power consumption which can cause dynamic nonuniformities due to IR drops across the pixel matrix. This was addressed recently by the use of time to analog converter arrays with external analog to digital converters (Parmesan et al., 2015; Crotti et al., 2012).

## 12.4 FLIM system-on-chip

### 12.4.1 Introduction

FLIM, when combined with fluorescence resonance energy transfer (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 stage. The complexity, high cost, and highly specialized knowledge required for the operation/acquisition has limited the prevalence of FLIM instrumentation. There have been efforts to integrate 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; Mattioli Della Rocca et al., 2016; Poland et al., 2016). 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.

### 12.4.2 Wide-field FLIM systems

In conventional wide-field FLIM systems, photons are captured on an ICCD-based camera (Colyer et al., 2012; Hirvonen et al., 2016). 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. MCP-based CCDs require a high voltage, usually

600–900 V, applied to the MCP. Another limitation for MCP-TCSPC systems is that the photon count rate is limited by the dead time of the TCSPC module (50–200 ns), making the maximum achievable photon count rate around a few MHz. To enhance the photon count rate, multi-channel TCSPC modules or TDCs with an ultrashort dead time (Chen et al., 2017) can be integrated. EMCCDs contain several gain registers between the end of the shift register and the output amplifier (Hyncek, 2001). For EMCCDs, any dark current remaining will be multiplied up along the readout path. That is why EMCCDs must be cooled down to  $-100^{\circ}\text{C}$ , significantly increasing their cost ( $\sim\text{£}25\text{--}30 \text{ K}$ ). CCD-based systems mainly suffer from readout noise generated by the electronic readout circuitry, which includes charge transfer noise, amplifier noise, and quantization noise emerging from digitizing signals. In low light situations, readout noise (especially for EMCCDs) can deteriorate the 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, EMCCDs, the MCP and the electron multiplying (EM) register suffer from significant aging effects.

#### 12.4.3 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  $\mu\text{s}$  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 (Kumar et al., 2007), or multi-channel TCSPC systems (Rinnenthal et al., 2013) although instrumentally intensive, have been introduced. 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.

#### 12.4.4 CMOS miniaturized PMTs

Unlike the low optical gain CCD-based or CMOS imaging sensors (CIS), a SPAD detector (single-photon avalanche diode), 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 (1 MHz = 1 million photons per second) with a dark count much  $<100 \text{ Hz}$ , even without using a cooling system (Richardson et al., 2009a, b). With innovative structures and new manufacturing technologies, high-performance SPAD imagers have been rapidly reported (Al Abbas et al., 2017;

Lindner et al., 2017; Perenzoni et al., 2016; Gyongy et al., 2018; Webster et al., 2012). These recent SPAD imagers have much higher QE and greatly increased fill factor approaching within a factor of 3 the sensitivity of backside illuminated EMCCD/sCMOS image sensors (Veerappan et al., 2011). The SNR of SPADs is nearly shot-noise limited, an inherent property of light (Eq. 12.8). Thus, SPAD detectors have no readout noise with its dark current noise usually much lower than its photon count rate, so the SNR is only Poisson noise limited capable of operating at a high frame rate.

$$SNR_{SPAD} = \frac{G\eta\phi\tau}{\sqrt{(N_{shot})^2}} \quad (12.8)$$

With single-photon sensitivity, SPADs are highly suitable for photon-starved applications such as single-molecule detection. In essence, SPADs are miniaturized 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 12.1). 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. This is being addressed by on-chip compression schemes such as local histogramming (Erdogan et al., 2017) providing up to two orders of magnitude reduction in I/O rate. This sensor is being applied to confocal spectral FLIM for *in vivo* imaging and time-resolved fluorescence spectroscopy (TRFS) (Kufcsak et al., 2017).

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 CIS sensors. Timing jitter mainly depends on the time a photo-generated carrier requires from the absorption point to enter 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. Take a recently developed SPAD array as an example (Veerappan et al., 2011), the SPAD array has  $160 \times 128$  SPADs plus in-pixel 10-bit 55 ps TDCs that allow recording the time tag of each, individually detected photon. The dead time 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 have been rapidly reported. The data throughput in CMOS SPAD arrays is usually massive, challenging the limited bandwidth of readout circuitry. The bottleneck to real-time imaging in such systems shifts toward data readout and lifetime calculation. Lifetime imaging becomes very compute-intensive, and it is impossible to achieve real-time imaging using traditional curve-fitting algorithms or software tools. It is desirable to have on-chip or on-FPGA imaging processors compressing or preprocessing the data before sending all raw data into PCs (Li et al., 2011; Mattioli Della Rocca et al., 2016; Erdogan et al., 2017).

**Table 12.1** Comparison of photon-capture devices

	CCD	ICCD <sup>a</sup>	EMCCD	CMOS imaging sensor (CIS)	Multi-channel PMT	CMOS SPAD
QE <sup>b</sup>	50%–90%	10%–50%	50%–90%	~60%	20%–40%	30%–70%
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 <sup>c</sup>
Timing resolution	≥1 μs (gating)	2–10 ns (gating)	~μs (gating)		~100 ps	<100 ps
Frame rate (Hz; frames per second)	100	10 <sup>d</sup>	10–100	100	Confocal scanning; image size dependent	≥5000
Pixel Pitch <sup>e</sup>	7 μm	10 μm	8–15 μm	7 μm	Not applicable but some are several mm, pending application	5–30 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	512 × 512 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 scanning with both TCSPC and gating
Single-photon detection capability	No	No	No	No	Yes	Yes

<sup>a</sup>Based on Hamamatsu's GaAsP.

<sup>b</sup>QE: quantum efficiency.

<sup>c</sup>Current development costs. Market costs will be competitive.

<sup>d</sup>PI-MAX3 “near” video rate frame rate.

<sup>e</sup>Based on high-end devices.

### **12.4.5 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. Traditional frequency-domain experiments usually need to acquire tens of phase images in order to obtain accurate lifetime images. To improve the acquisition speed, Esposito et al. (2006) demonstrated an all-solid-state  $124 \times 160$  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  $40 \times 55 \mu\text{m}^2$  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 minimize the effects of readout noises (Guo and Sonkusale, 2011a, b) 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. Recently, Zhao et al. (2012) and Raspe et al. (2016) demonstrated very innovative all-solid-state cameras that can significantly increase the acquisition speed to avoid photobleaching.

### **12.4.6 Latest CMOS fluorescence lifetime imaging/sensing systems**

Scientific or laboratory confocal scanning FLIM systems usually contain a single or multi-anode PMTs 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 are commercially available to improve the imaging speed, but they are bulky and expensive with image scanning still required. In usual laboratory setups, gated micro-channel plates (MCP)-based CCDs are used with wide-field excitation to provide high-speed lifetime imaging. Such systems, however, are also expensive. The recent advance in CMOS technologies, on the other hand, can provide much cheaper solutions. Large scientific grade CIS arrays can be fabricated with amplifiers integrated in the same chip. With simple gating techniques, a  $256 \times 256$  image sensor for wide-field FLIM applications in a  $0.18\text{-}\mu\text{m}$  CMOS technology with a pinned-photodiode process option was developed by Kawahito's research group (Yoon et al., 2009), where an in-pixel charge transfer mechanism was applied to calculate the fluorescence lifetimes. On-chip amplifiers are included to preprocess the signals. If the charge transfer lifetimes can be made negligible as we expect from the devices shifting toward a more advanced processes,

the systems can be used to monitor lifetimes of nanoseconds. The group recently reported a larger  $512 \times 310$  image sensor (Li et al., 2016) and a  $128 \times 128$  image sensor with innovative four-gate control (Seo et al., 2017) fabricated in a  $0.11\text{-}\mu\text{m}$  CMOS image sensor. Huang et al. (2009) reported a FLIM system using CMOS APS arrays with a pixel pitch of  $50\text{ }\mu\text{m}$  and a fill factor of 23%. Another  $256 \times 256$  sensor array with a pitch of  $6.3\text{ }\mu\text{m}$  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, b; Gersbach et al., 2009; Webster et al., 2012). The integrated chip of low-noise SPAD plus TDC arrays can act as a miniaturized version of multichannel PMT plus TCSPC module (Richardson et al., 2009a, b; Veerappan et al., 2011; Field et al., 2014; Gyongy et al., 2018), and video-rate FLIM 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., 2008a, b; Gola et al., 2011; Pancheri and Stoppa, 2008; Pancheri et al., 2011; Li et al., 2012; Bronzi et al., 2014; Burri et al., 2014; Homulle et al., 2016; Lee et al., 2016; Perenzoni et al., 2016) 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; Mattioli Della Rocca et al., 2016).

#### **12.4.7 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; Seo et al., 2017) or in parallel gated counters (Grauw and Gerritsen, 2001) providing a real-time lifetime imaging. In the past, gating techniques were 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; Bronzi et al., 2014; Burri et al., 2014; Homulle et al., 2016; Lee et al., 2016; Perenzoni et al., 2016). Also, to study highly dynamic biological systems *in vivo* in real time such as following single-molecule behavior at replication forks requires fluorescent probes that photoswitch or photocconvert in combination with highly efficient detectors. Specifically, imaging integration times

need to be in the milliseconds or even an order of magnitude lower depending on 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 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 high-speed imaging, SPAD-based FLIM systems have been successfully demonstrated in low-cost CMOS imaging process. 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.

#### 12.4.7.1 *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 optimized 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. One of the advantages for using integrated CMOS image sensors is that correlated gating schemes can be applied to optimize the photon collection efficiency (Li et al., 2012). The photon counting rate of SPADs can be up to tens of MHz with a dark count  $<100$  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), Benetti et al. (2010), and Pancheri et al. (2013) used four or more gated counters in parallel for each binned pixel in their SPAD arrays to collect photons. They then applied the NLSM to solve fluorescence decays. In fact, end users can design nonoverlapping or overlapping gates in such systems and use fast direct algorithms proposed by Sharman and Periasamy (1999) to calculate lifetimes. 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 behaviors, 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  $128 \times 128$  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

**Table 12.2** Data throughput and hardware required for 128 x 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

20 photon events (5-bits) per pixel per frame. The data throughput is  $64 \times 128 \times 2 \times 5 \times 10^6 = 82$  Gbit/s (assuming two pads for a column). Digital I/O pad can readout  $64 \times 2 \times 5 \times 10^6 = 640$  Mbit/s in the latest CMOS technology, but high frame-rate readout does pose a significant challenge in data transfer. The data must be preprocessed 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 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  $128 \times 128 \times 2 \times 15 \times 10^3 = 480$  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 multicore PC to conduct  $128 \times 128$  divisions (18-bit) in 0.01 s. 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](#))

[Table 12.2](#).

#### 12.4.7.2 Non-gating embedded algorithms

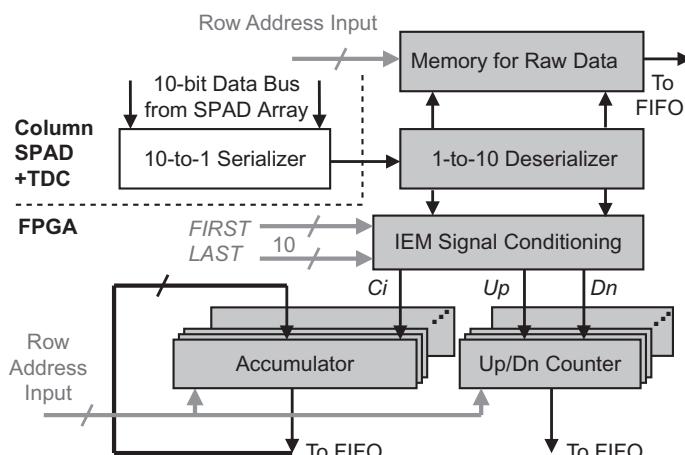
Gating systems usually contain a limited number of parallel gated counters due to hardware complexity. The main bottleneck for gating systems, however, is that when a full decay histogram is required for detailed analysis, it is necessary to generate shifting gating signals and acquire tens or even hundreds of snapshots. This process is therefore time consuming. [Grauw and Gerritsen \(2001\)](#) suggested that “near ideal” lifetime imaging is possible through optimally designed gating, but it is not easy to reach such optimization 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 w/ or w/o scanning for detailed observations. The latest CMOS SPAD array with in-pixel 10-bit 55 ps TDCs can offer such functionalities ([Richardson et al., 2009a, b; Veerappan et al., 2011](#);

([Gyongy et al., 2018](#)). SPAD arrays can contain in-pixel TDCs or TACs or column TDCs to generate time-resolved arrival time data and on-chip lifetimes ([Tyndall et al., 2012](#)) making them emulating PMT-based TCSPCs while providing extra parallelism. Consider a  $128 \times 128$  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; Arlt et al., 2013](#)), 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  $128/2 \times 10 \times 10^6 = 640$  Mbit/s per pad (64 pixels per pad). Similar to the argument above, the data must be preprocessed 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, 2009, 2010](#)), and the calculated lifetimes are.

$$\tau_{IEM} = \frac{h \cdot \sum_{j=0}^{M-1} (\bar{C}_j N_j)}{N_0 - N_{M-1}}, \quad (12.9)$$

where  $\bar{C} = [1/3, 4/3, 2/3, \dots, 4/3, 1/3]$  from the Simpson's integration rule,  $h$  is the TDC bin width, and  $N_j$  is the count number in the  $j$ th 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 Eq. (12.9), 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.

[Fig. 12.13](#) shows an on-FPGA IEM for a column of SPAD plus TDC pixels. Arrival time data are first transferred onto an FPGA through the serializer, and the deserializer



**Fig. 12.13** On-FPGA IEM for a column of CMOS SPAD plus TDC pixels.

can send the data to the FIFO directly or to the signal conditioning block where  $C_i$ ,  $Up$ , and  $Dn$  are generated to increment the accumulator and Up/Dn counter for the numerator and denominator of Eq. (12.9). 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 (Li et al., 2012). Assume that the fluorescence histogram is  $f(t) = \sum_{i=1}^k a_i e^{-t/\tau_i}$ , where  $a_i$  is the amplitude proportion and  $\tau_i$  is the lifetime. Its amplitude weighted and intensity weighted lifetimes are  $\tau_A = (\sum_{i=1}^k a_i \tau_i) / (\sum_{i=1}^k a_i)$  and  $\tau_I = (\sum_{i=1}^k a_i \tau_i^2) / (\sum_{i=1}^k a_i \tau_i)$ , respectively. The prior one is mostly used to calculate the FRET efficiency (Wu and Brand, 1994), whereas the later is used for the cases of dynamic quenching (Fišerová and Kubala, 2012; Sillen and Engelborghs, 1998). From the definition of the IEM (Li et al., 2008),

$$\tau_{IEM} = \frac{h \cdot \sum_{j=0}^{M-1} (\bar{C}_j N_j)}{N_0 - N_{M-1}} \sim \frac{\int_0^T f(t) dt}{f(0) - f(T)} = \frac{\sum_i a_i \tau_i (1 - e^{-T/\tau_i})}{\sum_i a_i (1 - e^{-T/\tau_i})} \sim \tau_A = \frac{\sum_i a_i \tau_i}{\sum_i a_i}, \quad (12.10)$$

where  $T$  is the measurement window.

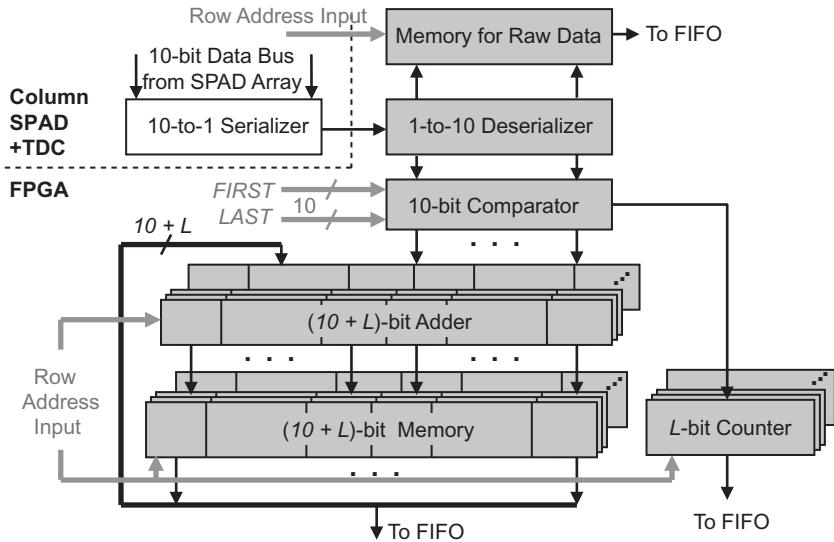
The lifetime calculated by the IEM is equal to  $\tau_A$  if  $T$  is much larger than the biggest lifetime component. Without resolving all  $a_j$  and  $\tau_j$  from the decay, IEM can directly calculate  $\tau_A$  and therefore the efficiency of the FRET.

Similar to the IEM, another lifetime embedded algorithm CMM (Li et al., 2010, 2011; Tyndall et al., 2012; Krstajić et al., 2014; Poland et al., 2016) can be implemented in our SPAD plus TDC systems. The lifetime is the average arrival time of all detected photons;

$$\tau_{CMM} \sim \frac{\int_0^T t f(t) dt}{\int_0^T f(t) dt} = \frac{\sum_i a_i \tau_i^2 (1 - e^{-T/\tau_i} - e^{-T/\tau_i} \cdot \frac{T}{\tau_i})}{\sum_i a_i \tau_i (1 - e^{-T/\tau_i})} \sim \left( \frac{\sum_{i=1}^{N_c} \bar{D}_i}{N_c} + \frac{1}{2} \right) h, \quad (12.11)$$

where  $\bar{D}_i$  is the 10-bit TDC output of the  $i$ th captured photon.

CMM can be used in confocal TCSPC systems considering nonideal instrument response (Won et al., 2011; Poland et al., 2016). On-FPGA CMM (Fig. 12.14) is slightly more complicated than 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  $32 \times 32$  CMOS SPAD camera and a  $32 \times 32$  silicon photon multiplier, respectively, to show video-rate lifetime imaging (Li et al., 2011) and high-speed sensing (Tyndall et al., 2012) or flow cytometry (Mattioli Della Rocca et al., 2016).



**Fig. 12.14** On-FPGA CMM for a column of CMOS SPAD plus TDC arrays.

For multi-exponential decays, the CMM generates an intensity weighted lifetime if the measurement window  $T$  is much larger than the biggest lifetime component. To fast resolve bi-exponential decays, an innovative hardware friendly algorithm with decent photon efficiency was also proposed (Li et al., 2015). This algorithm can be easily implemented to compress the image data.

## 12.5 Outlook

The cost, miniaturization, and robustness benefits of CMOS fluorescence lifetime sensors have been clearly demonstrated by various researchers in new applications such as contact imaging for microarrays, lab-on-chip, or portable sensing (Charbon, 2008; Palubiak and Deen, 2014; Lee et al., 2016; Bruchini et al., 2017). To replace PMT or GII in existing 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 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  $\mu\text{m}$  have been achieved. New 3D stacking technologies promise denser SPAD arrays to be introduced with more integrated functionalities (Al Abbas et al., 2017) such as on-chip histogramming (Erdogan et al., 2017)

and on-chip data processing for time-resolved imaging and spectroscopy. 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 autocorrelation. 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.

## References

- Al Abbas, T., et al., 2017. 8.25 $\mu$ m Pitch 66% fill factor global shared well SPAD image sensor in 40nm CMOS FSI technology. In: Proceedings of the International Image Sensors Workshop, Hiroshima, Japan P10.
- Arlt, J., Tyndall, D., Rae, B., Li, D., Richardson, J., Henderson, R., 2013. A study of pile-up in integrated time-correlated single photon counting systems. *Rev. Sci. Instrum.* 84 (10), 103105.
- Ballew, R.M., 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., 2015. Advanced Time-Correlated Single Photon Counting Techniques. Springer, Berlin. URL: <http://www.becker-hickl.com/literature.htm>. (Accessed 6 January 2019).
- Becker, W., 2012. Fluorescence lifetime imaging – techniques and applications. *J. Microsc.* 247, 119–136. <http://www.becker-hickl.com>. (Accessed 8 January 2012).
- Benetti, M., Iori, D., Panzeri, L., Borghetti, F., Pasquardini, L., Lunelli, L., Pederzolli, C., Gonzo, L., Dalla Betta, G.-F., Stoppa, D., 2010. Highly parallel SPAD detector for time-resolved lab-on-chip. *Proc. SPIE Int. Soc. Opt. Eng.* 7723, 77231Q.
- Bonjour, L.-E., Singh, A., Baechler, T., Kayal, M., 2011. Novel imaging method and optimized demodulation pixels for wide-field fluorescence lifetime imaging microscopy. *IEEE Sensors J.* 724–727.
- Bronzi, D., Villa, F., Tisa, S., Tosi, A., Zappa, F., Durini, D., Weyers, S., Brockherde, W., 2014. 100 000 frames/s 64x32 single-photon detector array for 2-D imaging and 3-D ranging. *IEEE J. Sel. Topics Quantum Electron.* 20(6)3804310.
- Bruchini, C., Homulle, H., Charbon, E., 2017. Ten years of biophotonics single-photon SPAD imager applications—retrospective and outlook. *Proc. SPIE* 10069, 100691S.
- Burri, S., Maruyama, Y., Michalet, X., Regazzoni, F., Bruschini, C., Charbon, E., 2014. Architecture and applications of a high resolution gated SPAD image sensor. *Opt. Express* 22 (14), 17573–17589.
- Chan, S.P., Fuller, Z.J., Demas, J.N., 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.* 41094010.
- Chen, H., Zhang, Y., Li, D.D.-U., 2017. Low nonlinearity, missing-code free time-to-digital converters based on 28nm FPGAs with embedded bin-width calibrations. *IEEE Trans. Instrum. Meas.* 66, 1921.
- Cheng, Z., Zheng, X., Deen, M.J., Peng, H., 2016. Recent developments and design challenges of high-performance ring oscillator CMOS time-to-digital converters. *IEEE Trans. Electron Devices* 63 (1), 235–251.

- Chodavarapu, V.P., Shubin, D.O., Bukowski, R.M., Titus, A.H., Cartwright, A.N., Bright, F.V., 2005. CMOS mixed-signal phase detector for integrated chemical sensor systems. In: IEEE Sensors.
- Christenson, M., Sternberg, S., 2004. Fluorescence lifetime imaging microscopy (flim) lets biologists study cellular processes. SPIE Oemagazine 28–30.
- Colyer, R.A., Siegmund, O.H.W., Tremsin, A.S., Vallerga, J.V., Weiss, S., Michalet, X., 2012. Phasor imaging with a widefield photon-counting detector. *J. Biomed. Opt.* 17 (1).016008 016008-1–016008-12.
- Crotti, M., Rech, I., Ghioni, M., 2012. Four channel, 40 ps resolution, fully integrated time-to-amplitude converter for time-resolved photon counting. *IEEE J. Solid State Circuits* 47 (3), 699–708.
- Cubeddu, R., Comelli, D., D'Andrea, C., Taroni, P., Valentini, G., 2002. Time-resolved fluorescence imaging in biology and medicine. *J. Phys. D Appl. Phys.* 35, 61–76.
- Durini, D., Spickermann, A., Mahdi, R., Brockherde, W., Vogt, H., Grabmaier, A., Hosticka, B.J., 2010. Lateral drift-field photodiode for low-noise, high-speed, large photoactive-area CMOS imaging applications. *Nucl. Inst. Methods Phys. Res. A* 624, 470–475.
- Dutton, N.A.W., et al., 2016. A SPAD-based QVGA image sensor for single-photon counting and quanta imaging. *IEEE Trans. Electron Devices* 63 (1), 189–196.
- 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., 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.
- Erdogan, A., Walker, R., Finlayson, N., Krstajic, N., Williams, G.O.S., Henderson, R., 2017. A 16.5 giga events/s 1024x8 SPAD line sensor with per-pixel zoomable 50ps-6.4ns/bin histogramming TDC. In: 2017 Symposium on VLSI Circuits, Kyoto, Japan, 5–8 June, pp. C292–C293.
- Esposito, A., 2011. Beyond range: innovating fluorescence microscopy. *J. Remote Sens.* 4, 111–112.
- Esposito, A., Gerritsen, H.C., Oggier, T., 2006. Innovating lifetime microscopy: a compact and simple tool for life sciences, screening, and diagnostics. *J. Biomed. Opt.* 11034016.
- 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., Takano, Y., 2003. An image sensor which captures 100 consecutive frames at 1,000,000 frame/s. *IEEE Trans. Electron Devices* 50, 144–151.
- Field, R.M., Realov, S., Shepard, K.L., 2014. A 100fps, time-correlated single-photon-counting-based fluorescence-lifetime imager in 130nm CMOS. *IEEE J. Solid State Circuits* 49 (4), 867–880.
- Fiserová, E., Kubala, M., 2012. Mean fluorescence lifetime and its error. *J. Lumin.* 132 (8), 2059–2064.
- Gadella, T.W., Jovin, T.M., Clegg, R.M., 1993. Fluorescence lifetime imaging microscopy (FLIM)-spatial-resolution of microstructures on the nanosecond time-scale. *Biophys. Chem.* 48 (2), 221–239.
- Gasparini, L., Zarghami, M., Xu, H., Parmesan, L., Garcia, M.M., Unternährer, M., Bessire, B., Stefanov, A., Stoppa, D., Perenzoni, M., 2018. A  $32 \times 32$ -pixel time-resolved single-photon image sensor with  $44.64\mu\text{m}$  pitch and 19.48% fill-factor with on-chip row/frame skipping features reaching 800kHz observation rate for quantum physics applications. In: IEEE International Solid-State Circuits Conference—(ISSCC), San Francisco, CA, pp. 98–100.
- Gersbach, M., Maruyama, Y., Trimananda, R., Fishburn, M.W., Stoppa, D., Richardson, J.A., Walker, R., Henderson, R., 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.
- Gersbach, M., Richardson, J., Mazaleyrat, E., Hardillier, S., Niclass, C., Henderson, R., Grant, L., Charbon, E., 2009. A low-noise single-photon detector implemented in a 130 nm CMOS imaging process. *Solid State Electron.* 53, 803–808.
- Gola, A., Pancheri, L., Piemonte, C., Stoppa, D., 2011. A SPAD-based hybrid system for time-gated fluorescence measurements. *Proc. SPIE.* 8033 803315-803315-4.
- Grauw, C.J., Gerritsen, H.C., 2001. Multiple time-gate module for fluorescence lifetime imaging. *Appl. Spectrosc.* 55, 670–678.
- Guo, J., Sonkusale, S., 2011a. A single chip fluorometer for fluorescence lifetime spectroscopy in 65nm CMOS. In: *IEEE Sensors*, pp. 766–769.
- Guo, J., Sonkusale, S., 2011b. A CMOS imager with digital phase readout for fluorescence lifetime imaging. In: *Proceedings of the 37th European Solid-State Circuits Conference*, pp. 115–118.
- Gyongy, I., et al., 2018. A  $256 \times 256$  100-kfps 61% fill-factor SPAD image sensor for time-resolved microscopy applications. *IEEE Trans. Electron Devices* 65 (2), 547–554.
- Hanson, K.M., Behne, M.J., Barry, N.P., Mauro, T.M., Gratton, E., 2002. Two-photon fluorescence imaging of the skin stratum corneum pH gradient. *Biophys. J.* 83, 1682–1690.
- Hirvonen, L.M., Becker, W., Milnes, J., Conneely, T., Smietana, S., Le Marois, A., Jagutzki, O., Suhling, K., 2016. Picosecond wide-field time-correlated single photon counting fluorescence microscopy with a delay line anode detector. *Appl. Phys. Lett.* 109, 071101 071101-1-071101-4.
- Homulle, H., Powolny, E., Stegehuis, P.L., Dijkstra, J., Li, D.D.-U., Homicsko, K., Rimoldi, D., Muchlethaler, K., Prior, J.O., Sinisi, R., Dubikovskaya, E., Charbon, E., Bruschini, C., 2016. Compact solid-state CMOS single-photon detector array for in vivo NIR fluorescence lifetime oncology measurements. *Biomed. Opt. Express* 7 (5), 1797–1814.
- Huang, T.C., et al., 2009. A  $0.18\mu\text{m}$  CMOS array sensor for integrated time-resolved fluorescence detection. *IEEE J. Solid-State Circuits* 44 (5), 1644–1654.
- Hynecek, J., 2001. Impactron-a new solid state image intensifier. *IEEE Trans. Electron Devices* 48 (10), 2238–2241.
- Kawahito, S., 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., Bright, F.V., 2009. Sensitivity-enhanced CMOS phase luminometry system using Xerogel-based sensors. *IEEE Trans. Biomed. Circuits Syst.* 3 (5), 304–311.
- Kleinfelder, S., Chen, Y., Kwaitkowski, K., Shah, A., 2004. High-speed CMOS image sensor circuits with in-situ frame storage. *IEEE Trans. Nucl. Sci.* 51, 1648–1656.
- Kleinfelder, S., Kwaitkowski, K., Shah, A., 2005. A solid state streak camera. *Proc. SPIE* 5580, 235–243.
- Krishnan, R., Saitoh, H., Terada, H., Centonze, V., Herman, B., 2003. Development of a multiphoton fluorescence lifetime imaging microscopy (flim) system using a streak camera. *Rev. Sci. Instrum.* 74 (5), 2714–2721.
- Krstajić, N., Walker, R., Levitt, J., Poland, S.P., Li, D., Ameer-Beg, S., Henderson, R., 2014. A  $256 \times 8$  SPAD line sensor for time resolved fluorescence and Raman sensing. In: *Proc. 40th IEEE European Solid State Circuits Conference (ESSCIRC)*pp. 143–146.
- Kufcsak, A., Erdogan, A., Walker, R., Ehrlich, K., Tanner, M., Megia-Fernandez, A., Scholefield, E., Emanuel, P., Dhaliwal, K., Bradley, M., Henderson, R.K., Krstajic, N., 2017. Time-resolved spectroscopy at 19,000 lines per second using a CMOS SPAD line array enables advanced biophotonics applications. *Opt. Express* 25, 11103.

- Kumar, S., Dunsby, C., De Beule, P.A.A., Owen, D.M., Anand, U., Lanigan, P.M.P., Benninger, R.K.P., Davis, D.M., Neil, M.A.A., Anand, P., Benham, C., Naylor, A., French, P.M.W., 2007. Multifocal multiphoton excitation and time correlated single photon counting detection for 3-D fluorescence lifetime imaging. *Opt. Express* 15, 12548–12561.
- Lakowicz, J.R., 2006. *Principles of Fluorescence Spectroscopy*, third ed. Springer, New York.
- Lange, R., Seitz, P., 2001. Solid-state time-of-flight range camera. *IEEE J. Quantum Electron.* 37 (3), 390–397.
- Lee, C., Johnson, B., Jung, T.S., Molnar, A., 2016. A 72x60 angle-sensitive SPAD imaging array for lens-less FLIM. *Sensors* 16 (9), 1422.
- Lenn, T., Leake, M.C., 2012. Experimental approaches for addressing fundamental biological questions in living, functioning cells with single molecule precision. *Open Biol.* 2, 120090.
- Li, D.-U., Ameer-Beg, S., Arlt, J., Tyndall, D., Walker, R., Matthews, D., Visitskul, V., Richardson, J., Henderson, R.K., 2012. Time-domain fluorescence lifetime imaging techniques suitable for solid-state imaging sensor arrays. *Sensors* 12 (5), 5650–5669.
- Li, D.-U., Arlt, J., Tyndall, D., Walker, R., Stopka, D., Charbon, E., 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. 096012-096012-12.
- Li, D.-U., Bonnist, E., Renshaw, D., 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., Rae, B., Andrews, R., Arlt, J., 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. 017006-017006-10.
- Li, Z., Seo, M.-W., Kagawa, K., Yasutomi, K., Kawahito, S., 2016. CMOS image sensor with lateral electric field modulation pixels for fluorescence lifetime imaging with sub-nanosecond time response. *Jpn. J. Appl. Phys.* 55, 04EM06.
- Li, D.-U., Walker, R., Richardson, J., Rae, B., Buts, A., Renshaw, D., Henderson, R., 2009. Hardware implementation and calibration of background noise for an integration-based fluorescence lifetime sensing algorithm. *J. Opt. Soc. Am. A* 26 (4), 804–814.
- Li, D.D., Yu, H., Chen, Y., 2015. Fast bi-exponential fluorescence lifetime imaging analysis methods. *Opt. Lett.* 40 (3), 336–339.
- Lindner, S., et al., 2017. Column-parallel dynamic TDC reallocation in SPAD sensor module fabricated in 180 nm CMOS for near infrared optical tomography. *Proc. Int. Image Sensor Workshop* 86–89.
- Lippitsch, M.E., Draxler, S., 1993. Luminescence decay time-based optical sensors: principles and problems. *Sensors Actuators B Chem.* 11, 97–101.
- Maruyama, Y., Charbon, E., 2011. An all-digital, time-gated 128x128 SPAD array for on-chip, filter-less fluorescence detection. In: *Proceedings of 16th International Solid-State Sensors, Actuators and Microsystems Conference*, pp. 1180–1183.
- Matthews, D., Summers, H., Njoh, K., Errington, R., Smith, P., Ameer-Beg, S., Vojnovic, B., 2005. Time-resolved fluorescence measurements using stroboscopic excitation. *Proc. SPIE Int. Soc. Opt. Eng.* 5864, MD3.
- Mattioli Della Rocca, F., Nedbal, J., Tyndall, D., Krstajic, N., Li, D.D.-U., Ameer-Beg, S.M., Henderson, R.K., 2016. Real-time fluorescence lifetime actuation for cell sorting using a CMOS SPAD silicon photomultiplier. *Opt. Lett.* 41 (4), 673–676.
- Michalet, X., Kapanidis, A., Laurence, T., Pinaud, F., Doose, S., Pflughoeft, M., Weiss, S., 2003. The power and prospects of fluorescence microscopies and spectroscopies. *Annu. Rev. Biophys. Biomol. Struct.* 32, 161–182.

- Mitchell, A.C., Wall, J.E., Murray, J.G., 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., Simoni, A., 2006. CMOS single-photon avalanche diode array for time-resolved fluorescence detection. In: Proceedings of 32nd European Solid-State Circuits Conference, pp. 564–567.
- Niclass, C., Favi, C., Kluter, T., Gersbach, M., Charbon, E., 2008. A 128x128 single-photon image sensor with column-level 10-bit time-to-digital converter array. *IEEE J. Solid State Circuits* 43, 2977–2989.
- Niclass, C., Rochas, A., Besse, P.A., 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.
- Oggier, T., Lehmann, M., Kaufmann, R., Schweizer, M., Richter, M., Metzler, P., Lang, G., Lustenberger, F., Blanc, N., 2004. An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger). *Proc. SPIE Int. Soc. Opt. Eng.* 5249, 534–545.
- Palubniak, D.P., Deen, M.J., 2014. CMOS SPADs: design issues and research challenges for detectors circuits and arrays. *IEEE J. Sel. Top. Quantum Electron.* 20 (6), 409–426.
- Pancheri, L., Massari, N., Borghetti, F., Stoppa, D., 2011. A 32x32 SPAD pixel array with nanosecond gating and analog readout. In: Proceedings of 2011 International Image Sensor Workshop, Hokkaido, Japan.
- Pancheri, L., Massari, N., Stoppa, D., 2013. SPAD image sensor with analog counting pixel for time-resolved fluorescence detection. *IEEE Trans. Electron Devices* 60 (10), 3442–3449.
- Pancheri, L., Stoppa, D., 2008. A SPAD-based pixel linear array for high-speed time-gated fluorescence lifetime imaging. In: Proceedings of IEEE European Solid-State Circuits Conference, pp. 428–431.
- Parmesan, L., Dutton, N.A.W., Calder, N.J., Grant, L.A., Henderson, R.K., 2015. A 256 x 256 SPAD array with in-pixel time to amplitude conversion for fluorescence lifetime imaging microscopy. In: Proceedings of the International Image Sensor Workshop, Vaals, The Netherlands.
- Patounakis, G., Shepard, K., 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., Bastiaens, P.I.H., 1999. Simultaneous detection of multiple green fluorescent proteins in live cells by fluorescence lifetime imaging microscopy. *Curr. Biol.* 9, 269–274.
- Perenzoni, M., Massari, N., Perenzoni, D., Gasparini, L., Stoppa, D., 2016. A 160x120 pixel analog-counting single-photon imager with time-gating and self-referenced column-parallel A/D conversion for fluorescence lifetime imaging. *IEEE J. Solid State Circuits* 51 (1), 155–167.
- Poland, S.P., Erdogan, A.T., Krstajic, N., Levitt, J., Devauges, V., Walker, R.J., Li, D.D.-U., Ameer-Beg, S.M., Henderson, R.K., 2016. New high-speed centre of mass method incorporating background subtraction for accurate determination of fluorescence lifetime. *Opt. Express* 24 (7), 6899–6915.
- Prasad, P.N., 2003. Introduction to Biophotonics. Wiley-Interscience, New York, United States.
- Rae, B.R., Yang, J.B., Mckendry, J., Gong, Z., Renshaw, D., Girkin, J.M., Gu, E., Dawson, M.D., Henderson, R.K., 2010. A vertically integrated CMOS microsystem for time-resolved fluorescence analysis. *IEEE Trans. Biomed. Circuits Syst.* 4, 437–444.
- Raspe, M., Kedziora, K.M., van den Broek, B., Zhao, Q., de Jong, S., Herz, J., Mastop, M., Goedhart, J., Gadella, T.W.J., Young, I.T., Jalink, K., 2016. SiFLIM: single-image

- frequency-domain FLIM provides fast and photon-efficient lifetime data. *Nat. Methods* 13, 501–504.
- Razavi, B., 2007. Design considerations for direct-conversion receivers. *IEEE Trans. Circuits Syst. II* 44, 428–435.
- Richardson, J.A., Grant, L.A., Henderson, R.K., 2009a. Low dark count single-photon avalanche diode structure compatible with standard nanometer scale CMOS technology. *IEEE Photon. Technol. Lett.* 21 (14), 1020–1022.
- Richardson, J., Walker, R., Grant, L., Stoppa, D., Borghetti, F., Charbon, E., Gersbach, M., Henderson, R.K., 2009b. A 32x32 50 ps resolution 10 bit time to digital converter array in 130 nm CMOS for time correlated imaging. In: Proceedings of IEEE Custom Integrated Circuits Conference, pp. 77–80.
- Rinnenthal, J.L., Börnchen, C., Radbruch, H., Andresen, V., Mossakowski, A., Siffrin, V., Seelemann, T., Specker, H., Moll, I., Herz, J., Hauser, A.E., Zipp, F., Behne, M.J., Niesner, R., 2013. Parallelized TCSPC for dynamic intravital fluorescence lifetime imaging: quantifying neuronal dysfunction in neuroinflammation. *PLoS One.* 8(4)e60100.
- Rochas, A., Gani, M., Furrer, B., Besse, P.-A., 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., Levine, Y.K., 1995. Quantitative pH imaging in cells using confocal fluorescence lifetime imaging microscopy. *Anal. Biochem.* 227, 302–308.
- Schwartz, D.E., Charbon, E., Shephard, K.L., 2008a. A single-photon avalanche diode array for fluorescence lifetime imaging microscopy. *IEEE J. Solid State Circuits* 43, 2546–2557.
- Schwartz, D.E., Gong, P., Shepard, K.L., 2008b. Time-resolved Förster-resonance-energy-transfer DNA assay on an active CMOS microarray. *Biosens. Bioelectron.* 24, 383–390.
- Seo, M.-W., Shirakawa, Y., Masuda, Y., Kawata, Y., Kagawa, K., Yasutomi, K., Kawahito, S., 2017. 4.8 A 0.44e<sup>-</sup> rms read-noise 32fps 0.5M pixel high-sensitivity RG-less-pixel CMOS image sensor using bootstrapping reset. In: IEEE International Conference of Solid-State Circuits (ISSCC) Dig. Tech. Papers, San Francisco, USA, 5–9 February, pp. 70–71.
- Sharman, K.K., 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., Engelborghs, Y., 1998. The correct use of average fluorescence parameters. *Photochem. Photobiol.* 67 (5), 475–486.
- Stoppa, D., Borghetti, F., Richardson, J., Walker, R., Grant, L., Henderson, R.K., Gersbach, M., Charbon, E., 2009. A 32x32-pixel array with in-pixel photon counting and arrival time measurement in the analog domain. In: Proceedings of European Solid-State Circuits Conference, pp. 204–207.
- Suhling, F., et al., 2015. Fluorescence lifetime imaging. In: *Handbook of Photonics for Biomedical Engineering*. Springer, Netherlands, pp. 353–405.
- Theuwissen, A.J.P., 2008. CMOS image sensors; State-of-the-art. *J. Solid-State Electron.* 52, 1401–1406.
- Theuwissen, A.J.P., Seitz, P., 2011. *Single-Photon Imaging*. Springer.
- Tubert, C., Simony, L., Roy, F., Tournier, A., Pinzelli, L., Magnan, P., 2009. High speed dual port pinned-photodiode for time-of-flight imaging. In: Proceedings of 2009 International Image Sensor Workshop, Bergen, Norway.
- Turconi, S., Bingham, R.P., Haups, U., Pope, A.J., 2001. Developments in fluorescence lifetime based analysis for ultra-HTS. *Drug Discov. Today* 6, 27–39.

- Tyndall, D., Rae, B., Li, D.-U., Richardson, J., Arlt, J., Henderson, R., 2012. A high-throughput time-resolved mini-silicon photomultiplier with embedded fluorescence lifetime estimation in  $0.13\mu\text{m}$  CMOS. *IEEE Trans. Biomed. Circuits Syst.* 6, 562–570.
- Ulku, A.C., et al., 2019. A  $512 \times 512$  SPAD image sensor with integrated gating for widefield FLIM. *IEEE J. Sel. Top. Quantum Electron.* 25 (1), 1–12.
- Veerappan, C., Richardson, J., Walker, R., Li, D.-U., Fishburn, M.W., Maruyama, Y., Stoppa, D., Borghetti, F., Gersbach, M., Henderson, R.K., Charbon, E., 2011. A  $160 \times 128$  single-photon image sensor with on-pixel 55ps 10b time-to-digital converter. In: *IEEE ISSCC Digest of Technical Papers*, pp. 312–314.
- Verveer, P.J., Squire, A., Bastiaens, P.I.H., 2000. Global analysis of fluorescence lifetime imaging microscopy data. *Biophys. J.* 78, 2127–2137.
- Vornicu, I., Carmona-Galan, R., Rodriguez-Vazquez, A., 2014. A CMOS  $0.18\mu\text{m}$   $64 \times 64$  single photon image sensor with in-pixel 11 b time-to-digital converter. In: *Proceedings of International Semiconductor Conference*, pp. 131–134.
- Wang, Y., Rae, B.R., Henderson, R.K., Gong, Z., McKendry, J., Gu, E., Dawson, M.D., Turnbull, G.A., Samuel, I.D.W., 2011. Ultra-portable explosive sensor based on a CMOS fluorescence lifetime analysis microsystem. *AIP Adv.* 1, 032115.
- Webster, E.A.G., Richardson, J.A., Grant, L.A., Renshaw, D., Henderson, R.K., 2012. A single-photon avalanche diode in 90-nm CMOS Imaging technology with 44% photon detection efficiency at 690 nm. *EEE Electron Device Lett.* 33 (5), 694–696.
- Won, Y.J., Moon, S., Yang, W., Kim, D., Han, W.-T., 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., Brand, L., 1994. Resonance energy transfer: methods and applications. *Anal. Biochem.* 218 (1), 1–13.
- Ximenes, A.R., Padmanabhan, P., Charbon, E., 2017. Mutually coupled ring oscillators for large array time-of-flight imagers. In: *Proceedings of the International Image Sensors Workshop*, Hiroshima, Japan. R25.
- Yoon, H.-J., Itoh, S., 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., Schelen, B., Schouten, R., van den Oever, R., Leenen, R., van Kuijk, H., Peters, I., Polderdijk, F., Bosiers, J., Raspe, M., Jalink, K., Geert Sander de Jong, J., van Geest, B., Stoop, K., Young, I.T., 2012. Modulated electron-multiplied fluorescence lifetime imaging microscope: all-solid-state camera for fluorescence lifetime imaging. *J. Biomed. Opt.* 17 (12), 126020.
- Zhong, W., Urayama, P., Mycek, M.-A., 2003. Imaging fluorescence lifetime modulation of a ruthenium-based dye in living cells: the potential for oxygen sensing. *J. Phys. D. Appl. Phys.* 36, 1689.

# Complementary metal-oxide-semiconductor (CMOS) X-ray sensors

13

A. Strum<sup>a</sup>, A. Fenigstein<sup>a</sup>

Revised by S. Rizzolo<sup>b</sup>

<sup>a</sup>TowerJazz, Newport Beach, CA, United States, <sup>b</sup>Institut Supérieur de l'Aéronautique et de l'Espace, Toulouse, France

## 13.1 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 consequence 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 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.

## 13.2 Intraoral and extraoral 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 by 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 (1000–10,000) 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 “cross talk”). 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 major drivers were:

1. much lower power consumption of CMOS compared to CCD;
2. possibility to integrate digital function directly on-chip in CMOS (such as analog-to-digital converter, ADC);
3. higher data rate and frame rate that allows higher resolutions; and
4. 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, about  $43 \times 30$  mm, and size 2 (smaller than size 1) about  $32 \times 24$  mm.

In the extraoral 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 intraoral 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 why 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  $5 \times 6$  in to allow the whole skull to be imaged with one shot.

### **13.3 Medical radiography, fluoroscopy, and mammography**

Medical radiography requires, by nature, large areas. The standard dimensions are  $10 \times 12$  in ( $25 \times 30$  cm) but can get as large as  $17 \times 17$  in ( $43 \times 43$  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 photosensitive 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 photosensitive 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 multiuse 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 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 (HDR), 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 5–8 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, that is, 60 frames per second (FPS) 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  $17 \times 17$  in 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  $\mu\text{m}$  pixels. For mammography, pixel sizes of 50–75  $\mu\text{m}$  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. However, recent advance were made in order to use these detector for mammography with encouraging results. Direct X-ray detection detectors, using amorphous selenium, were developed to fill the need of performing detector for large-area applications. 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.

### 13.4 CIS-based 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 ([Zentai, 2011](#)).

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

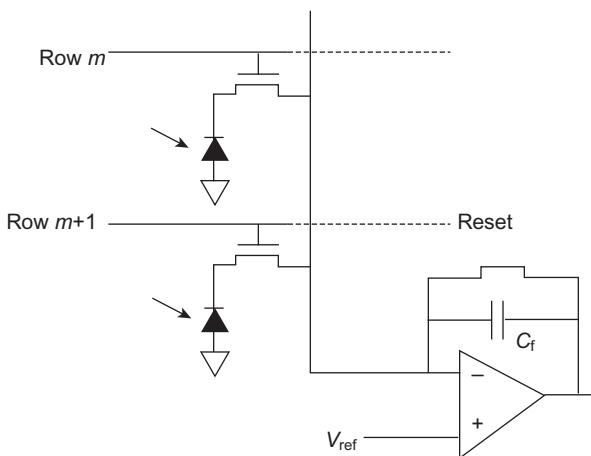
On the other hand, CIS technology is limited to wafer sizes (8 or 12 in. in diameter) while a:Si FPDs enable sizes up to  $43 \times 43$  cm panels quite easily. In addition, the cost per square centimeter is higher for CMOS technology. As described below, the area limitation can be solved by tiling 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 to 150  $\mu\text{m}$  compared to less than 10  $\mu\text{m}$  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.

## 13.5 Pixel design considerations for CMOS-based FPDs

### 13.5.1 1T pixel architecture on a:Si TFT FPD

The simplest architecture of a pixel requires a photodiode and one transistor (see Fig. 13.1). 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 photoelectrons 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 read-out circuitry where it is sampled. Double sampling is typically used; namely another

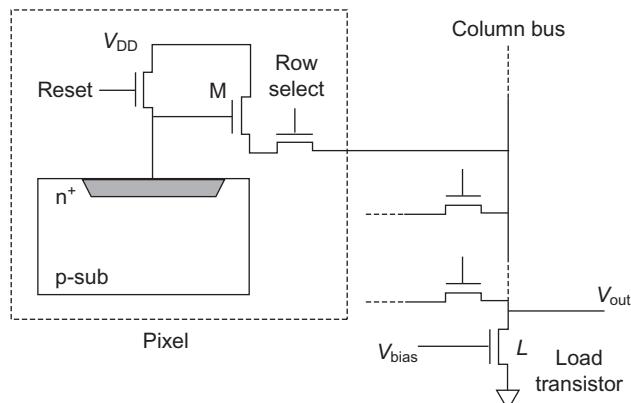


**Fig. 13.1** 1T pixel architecture. Each pixel contains a photodiode and one transistor. The amplifier is located at the bottom of the column.

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.

### 13.5.2 Three transistors (3T) pixel architecture

3T pixels are simply what are known as “active pixels.” The scheme of such pixels is shown in Fig. 13.2. 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 photoelectrons 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 readout column circuit unity gain amplifier realized by a transistor connected as “source follower” (SF) through the row-select transistor switch 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 noncorrelating 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 following paragraph.



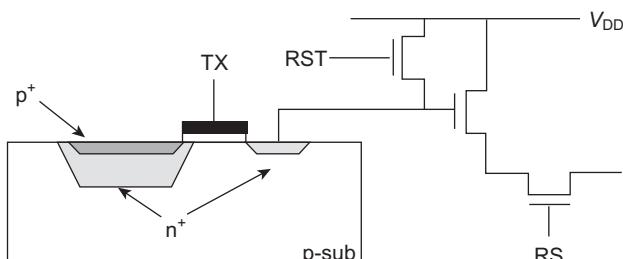
**Fig. 13.2** 3T pixel architecture. Transistor “M” is connected as a “source follower” (SF) and serves as a unity gain amplifier.

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, for an intraoral application, CISs with pixels which usually range between 15 and 20  $\mu\text{m}$ , 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 is about 70  $\mu\text{V}$ , 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 CIS for intraoral dental applications, 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.

### 13.5.3 Partially pinned and fully pinned photodiode: 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. 13.3.

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 photoelectrons, 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 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 since the pinning layer “encapsulates” the photodiode, avoiding the contact between the depleted region and the oxides. Then, the photodiode is protected from carrier generation in surfaces and interfaces. A well-optimized



**Fig. 13.3** The 4T pixel architecture.

4T pixel can achieve a noise floor of a few electrons. It is worth mentioning that the charge transfer between the collection region (the pinned photodiode) and the sense node, if not optimized can lead to CIS performance degradation such as image lag which, if important, can deteriorate the image.

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 therefore is not required for such applications. 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 photodiode” (PPPD) can have the required impact without the penalty of the longer transfer time. This particular photodiode retains all the advantages of standard 3T pixels such as high fill factor, dynamic range thanks to the elimination of the transfer between the photodiode and the sense node. On the other hand, the presence of the pinning layer (even if partial) allow maintaining the dark current reduction achieved by the abovementioned “encapsulation” of the n-region of the photodiode by a grounded p-region, which 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 ([Lahav et al., 2011](#)).

### **13.5.4 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 following 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 doses, 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:

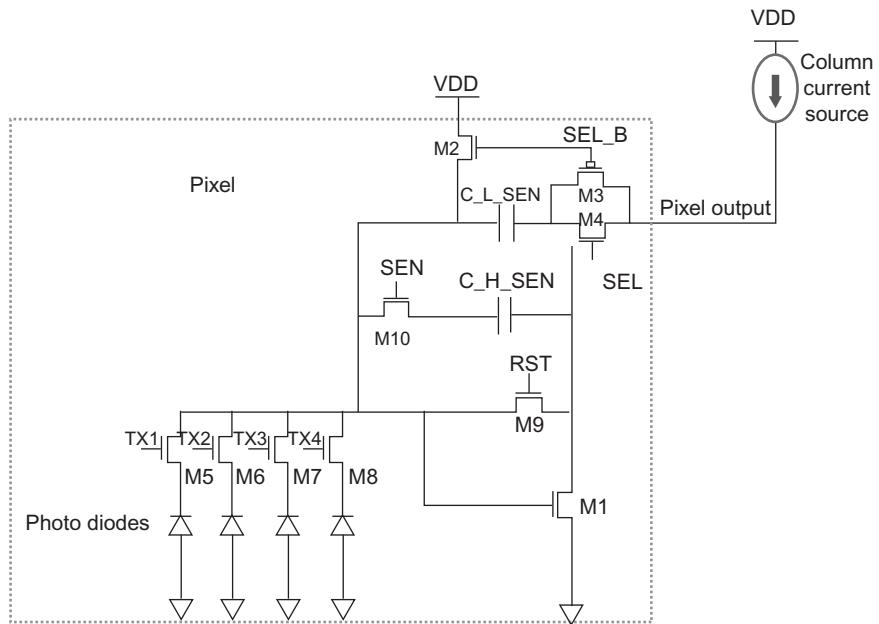
1. By employing “analog” binning in the array: it improves the SNR by eliminating the noise coming from the signal sampling operation.
2. One simultaneous sampling of several photodiodes together saves the need for multiple sampling followed by digital summing up “off chip.”

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 following section.

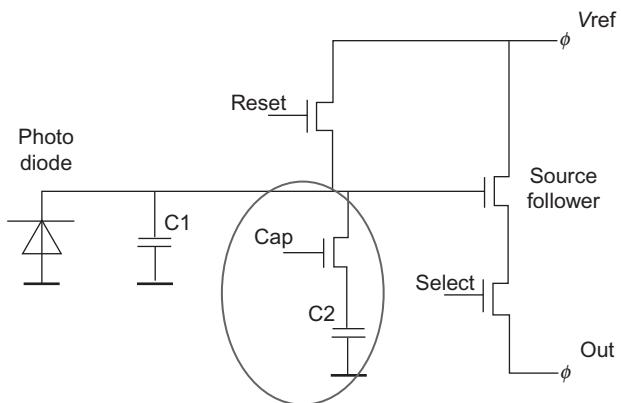
### 13.5.5 Increasing functionality of CIS pixels

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. 13.4. This pixel transfers the charge from the photodiode (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 nonlinear 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 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 (Lahav et al., 2011). 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  $\mu\text{V}/\text{e}$  units) is inversely proportional to the capacitance of the diode (3T pixel) or feedback capacitor (charge amplifier pixel). The schemes of dual gain in a regular 3T pixel (presented by Dalsa) (Bosiers et al., 2012) are shown in Fig. 13.5 and for a charge amplifier pixel (in a C&T product) shown in Fig. 13.4 (Fenigstein et al., 2013).

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



**Fig. 13.4** 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.



**Fig. 13.5** Dalsa's dual gain 3T pixel. Switch cap is used to add/remove the extra capacitance  $C_2$  to the “floating diffusion” capacitance  $C_1$ .

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 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 Ref. ([Fenigstein et al., 2013](#)) and the overall sensor performance is listed below.

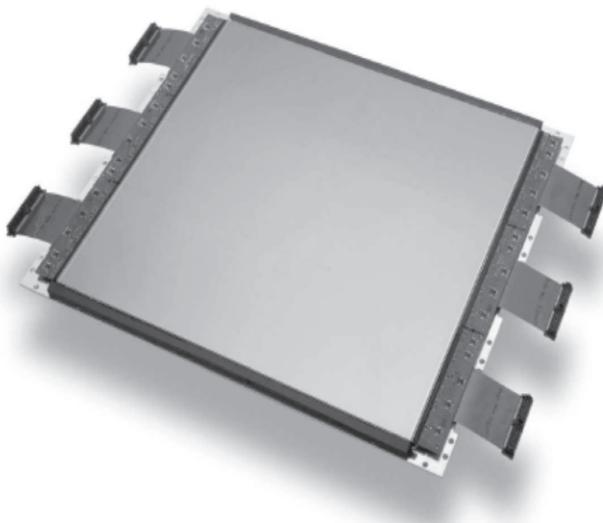
Another example of an even more advanced pixel functionality is the “photon counting” pixel, proposed by Caeleste ([Dierickx et al., 2011](#)). The number of simultaneously impinging visible photons gives information about the energy of the X-ray photon which enables what is called sometimes “color” 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.

### **13.5.6 On-chip ADC**

The use 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.

### **13.5.7 Yield and cost considerations**

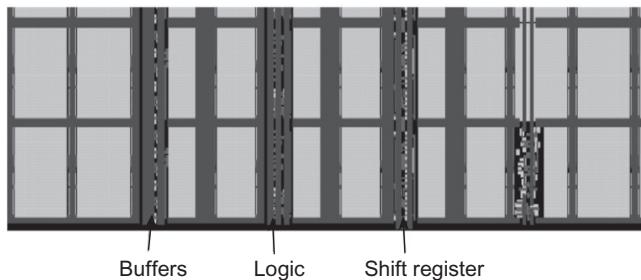
X-ray sensors dimensions are much larger than common VLSI chips. The number of dies on wafers varies between a few tens (intraoral dental) to one die per wafer (medical). This places a huge challenge on yield expectations. In fact, using any standard wafer yield model will result in a yield close to 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).



**Fig. 13.6** 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).

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. 13.6.

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 driver’s direction. This becomes even more complicated due to the fact that such large sensors are made using



**Fig. 13.7** Array arrangement for three-side buttable sensor with drivers and other circuitry located between the columns.

photolithography with “stitching” which requires building the array from repeating parts. The way to deal with this is either by pushing circuitry inside the array (Bosiers et al., 2012) as shown in Fig. 13.7, or by a smart stitching scheme (Sarig et al., 2010).

## 13.6 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 sections these figures of merit are discussed along with their typical numbers and the trade-offs between them.

### 13.6.1 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  $\sim 3$  is considered “visible,” namely an image can be seen on the background of the noise if the image signal level is at least about  $3 \times$  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 noise floor of the pixel and the analog chain. Assuming that noise sources are independent, the total SNR for a specific illumination level (or dose)

will be calculated by  $SNR = V_{sig} / \sqrt{\sum V_n^2}$  where  $V_1, V_2, \dots, V_N$  are noise sources which can be signal dependent like the shot noise or signal independent like the analog chain noise floor.

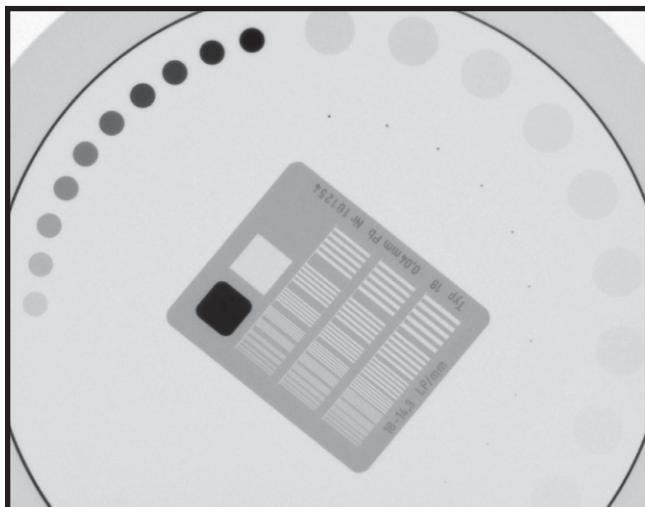
### 13.6.2 Resolution, MTF, and DQE

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 on pixel dimensions but also on the amount of cross talk 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. 13.8.

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

$$MTF = \sin\left(\frac{\pi f}{2f_n}\right) / \frac{\pi f}{2f_n}$$

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.



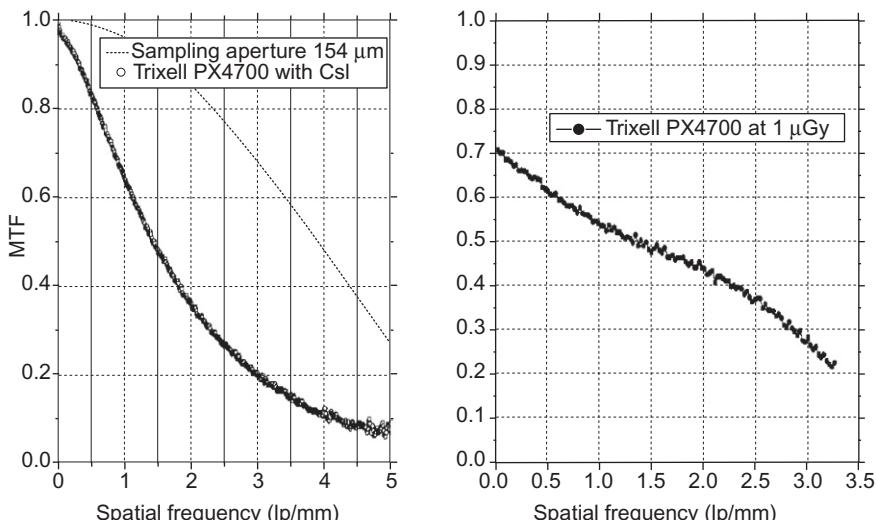
**Fig. 13.8** X-ray phantom for MTF evaluation.

The MTF drops further when there is cross talk between pixels. This cross talk can be due to several mechanisms: optical cross talk (light getting to the “wrong” pixel), diffusion cross talk (photoelectrons diffused to an adjacent pixel), and electrical cross talk (interference between output signals). Due to the rather large X-ray pixel geometry, diffusion cross talk usually can be minimized and for a well-designed X-ray sensor, electrical cross talk is avoided as well. Thus, for an indirect X-ray sensor, the major cross talk contribution is optical cross talk usually due to the scintillator. An example of MTF taken for Trixell’s 154  $\mu\text{m}$  pixel FPD is shown in the left curve in Fig. 13.9.

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 ([Rose, 1948](#)), takes both SNR and resolution into the same framework. The DQE is defined by

$$\text{DQE} \equiv \frac{\text{SNR}_{\text{out}}^2}{\text{SNR}_{\text{in}}^2}$$

for a perfect sensor  $\text{DQE}=1$ . 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 same Trixell’s sensor is shown above. As stated before, the SNR for low doses is dominated by the pixel and analog



**Fig. 13.9** An example of typical MTF and DQE curves for a medical X-ray sensor (Trixell’s PX4700 flat panel sensor).

chain noise floor. 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 noise floor 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.

### 13.6.3 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:

$$\begin{aligned} 20 \cdot \log_{10} \left( \frac{\text{Maxdose}}{\text{NED}} \right) &= 20 \cdot \log_{10} \left( \frac{\text{Voltage Swing}}{\text{Voltage - noise floor}} \right) \\ &= 20 \cdot \log_{10} \left( \frac{\text{Full Well}}{\text{noise floor - in - electrons}} \right) \end{aligned}$$

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 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 nonlinearity, to others that need linearity of better than 0.1% (as in CT applications).

## 13.7 X-ray sensors: Types and requirements

Different X-ray sensors and their specifications are listed in [Table 13.1](#). In [Tables 13.2 and 13.3](#) there is a list of several sensors of various vendors with their main parameters and features. This list is changing fast due to growing interest in

**Table 13.1** Different X-ray applications and the typical requirements for each application.

Application	Size (cm <sup>2</sup> )	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	—

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 13.4](#). The table shows the sensitivity and noise parameter of the sensor in its HDR mode (lsb stands for least significant bit).

The dark current distribution is shown in [Fig. 13.10](#). This parameter is usually measured by the dark current per unit area (pA/cm<sup>2</sup>). The actual dark current per pixel is reported in (e<sup>-</sup>/s). The MTF and the DQE curves of the Thales HD sensor are shown in [Figs. 13.11 and 13.12](#), respectively. The SNR versus radiation dose is presented in [Fig. 13.13](#). 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.

## 13.8 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 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 ([Kasap and Rowlands, 2002](#)). 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. 13.14](#).

**Table 13.2** Examples of different X-ray sensors from different vendors and their main parameters (part 1).

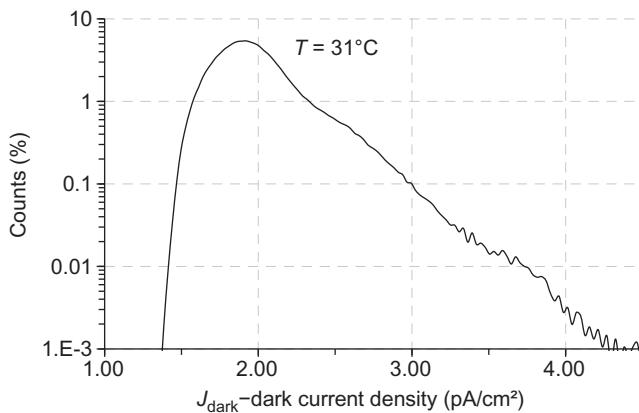
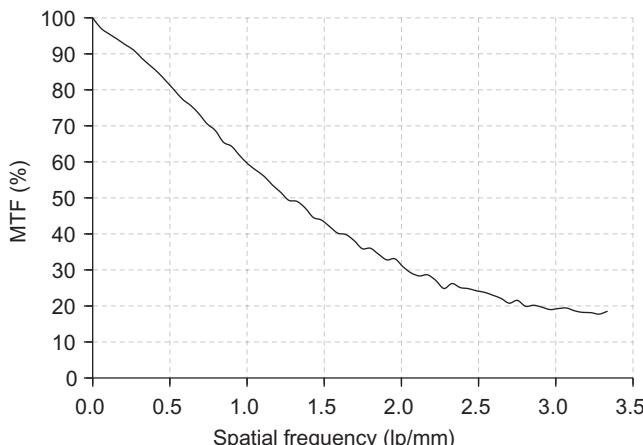
<b>Manufacturer</b>	<b>Reference</b>	<b>Wafer size</b>	<b>Image size</b>	<b>Pixel size (um)</b>	<b>Resolution (Mp)</b>	<b>FF</b>	<b>FPS</b>	<b>Pixel type</b>	<b>Full well</b>
Hamamatsu	Mori et al. (2002)	8	12 × 12	50 × 50	5.8	79	2	–	2.2 M
Hamamatsu	Fujita et al. (2004)	12	22 × 18	50 × 50	15.5	76	1	–	
Vatech	Heo et al. (2010)	8	12 × 14	200 × 200	0.5	67	30	4T	10 M
Vatech	Keoa et al. (2011)	12	23 × 17	70 × 70	8.5	75	3	3T	17 M
Rutherford	Turchetta et al. (2011)	8	14 × 12	50 × 50	–	–	–	–	–
Canon	Yamashita et al. (2011)	12	20 × 20	160 × 160	1.6	–	100	Smart	77 k
U. Lincoln	Esposito et al. (2011)	8	13 × 13	50 × 50	6.5	70	30	3T	280 k
T-DALSA 1	Korthout et al. (2009)	8	14 × 8	33.5 × 33.5	10	84	8.7	3T	650 k
T-DALSA 2	Teledynedalsa (n. d.)	8	13 × 13	100 × 100	1.7	85	45	3T	500 k
Thales	Fenigstein et al. (2013)	8	12 × 15	150 × 150	0.75	80	30	Smart	5.2 M
Dexela	Konstantinidis et al. (2012)	–	11.5 × 14.5	75 × 75	–	–	26	3T	1.5M

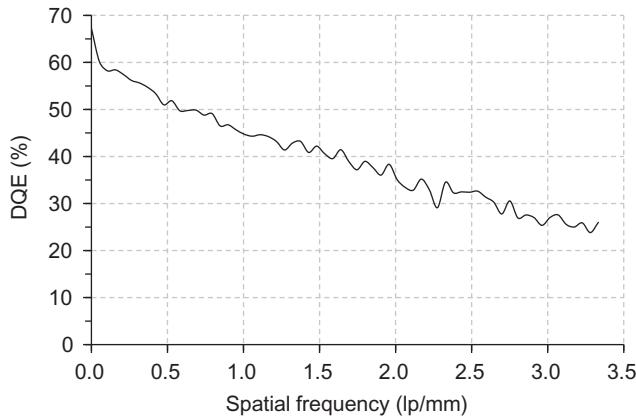
**Table 13.3** Examples of different X-ray sensors from different vendors and their main parameters (part 2).

Manufacturer	Reference	DR	QE (%)	Noise floor	Output	Comments
Hamamatsu	Mori et al. (2002)	–	–	1100e <sup>–</sup>	Analog	–
Hamamatsu	Fujita et al. (2004)	–	–	1100e <sup>–</sup>	Analog	–
Vatech	Heo et al. (2010)	84 dB	55	910e <sup>–</sup>	Analog	–
Vatech	Keoa et al. (2011)	84 dB	45	910e <sup>–</sup>	Analog	–
Rutherford	Turchetta et al. (2011)	–	–	30e <sup>–</sup>	Analog	Three-side buttable
Canon	Yamashita et al. (2011)	75 dB	–	13e <sup>–</sup>	Analog	In pixel circuitry PGA, S/H
U. Lincoln	Esposito et al. (2011)	65 dB	45	140e <sup>–</sup>	Analog	2 × 2 charge binning two-side buttable
T-DALSA 1	Korthout et al. (2009)	71 dB	50	175e <sup>–</sup>	Analog	Dose sensing binning three-side buttable
T-DALSA 2	Teledynedalsa (n.d.)	70 dB	50	135e <sup>–</sup>	Analog	Dose sensing binning three-side buttable
Thales	Turchetta et al. (2011)	96 dB	65	87e <sup>–</sup>	Digital	2 × 2 binning three-side buttable, real HDR
Dexela	Konstantinidis et al. (2012)	70 dB	–	170e <sup>–</sup>	Digital	Multi-resolution readout, high and low full well capacity switching

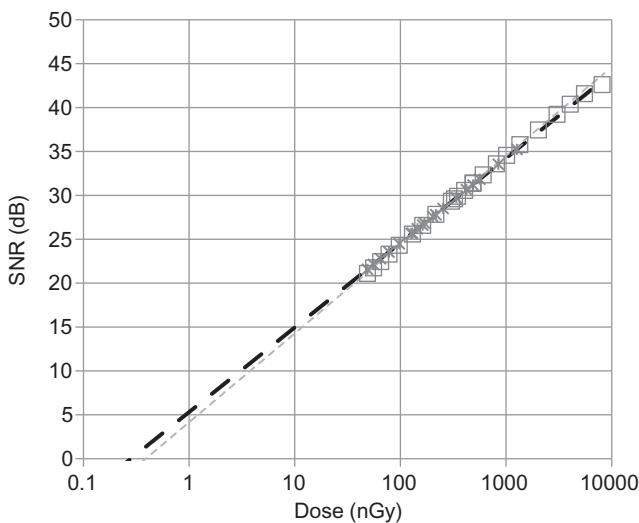
**Table 13.4** Parameters of SiX HD sensor of Thales.

Sensitivity	<b>8.47 lsb/nGy</b>
QE × fill factor	65%
CG (high/low)	2.32/0.29 $\mu$ V/e
Noise	2.19 lsb, 147 $e^-$
NED	0.16 nGy
Maximum Dose	11 $\mu$ Gy
MTF (Nyquist)	15%
DQE(0)	71%
Average dark current	1.9 pA/cm <sup>2</sup> , 2800 $e^-/s$ per pixel
Lag	0.12%

**Fig. 13.10** Dark current distribution for Thales sensor.**Fig. 13.11** MTF curve for Thales SiX HD sensor, pixel size is 150  $\mu\text{m}$ .

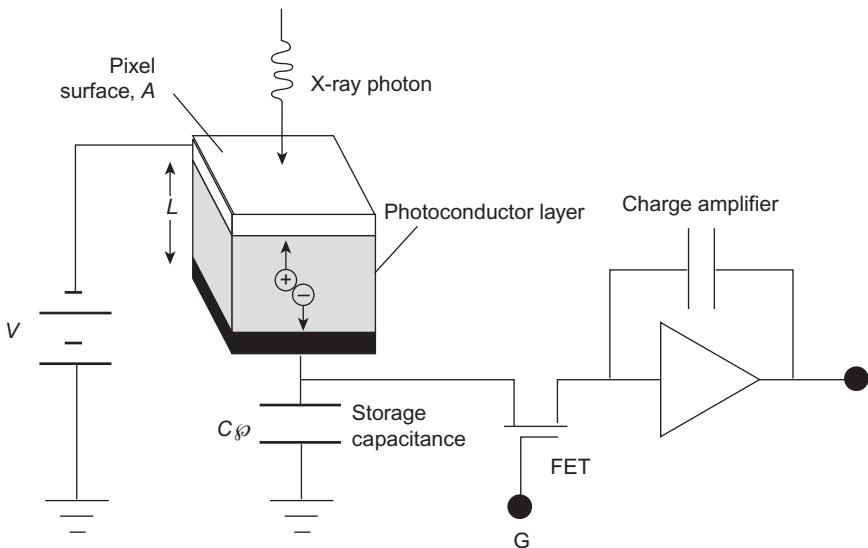


**Fig. 13.12** DQE curve of Thales SiX HD sensor.



**Fig. 13.13** SNR in dB vs radiation dose.

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 as an X-ray photoconductor, but the one which is already commercially available is based on amorphous selenium ( $\alpha$ -Se). The operation of such a photoconductor requires rather high voltages; a few kVolts in the case of  $\alpha$ -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.



**Fig. 13.14** 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.

The main noise source in such a pixel is the dark current of the photoconductor. This current is in the range of  $\text{nA}/\text{cm}^2$  for most of the candidate materials excluding  $\alpha\text{-Se}$  which is in the range of tens of  $\text{nA}/\text{cm}^2$ . The most significant advantage of direct sensors over indirect sensors is related to MTF. The scintillator 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  $\text{CsI(Tl)}$  improves dramatically the MTF by wave guiding the visible light into the underlying pixel (Nikl, 2006). But, even this improved MTF is still inferior to the geometrical ideal MTF as can be seen in Fig. 13.11 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 (Samei and Flynn, 2003). 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.

## 13.9 Conclusion and future trends

CIS sensors for X-ray applications are 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, noise floor, 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 HDR, 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 yielding, 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.

## References

- Bosiers, J., Korthout, L., Peters, I., 2012. Medical X-ray imaging using wafer-scale CMOS imagers. In: *6th Fraunhofer IMS Workshop on CMOS Imaging*, pp. 12–13.
- Dierickx, B., Dupont, B., Defernez, A., Ahmed, N., 2011. Color X-ray photon counting image sensing. In: *IISW 2011*.
- Esposito, M., Anaxagoras, T., Fant, A., Wells, K., Konstantinidis, A., Osmond, J.P.F., Evans, P.M., Speller, R.D., Allinson, N.M., 2011. DynAMITe: a wafer scale sensor for biomedical applications. *J. Instrum.* 6 (12), C12064.
- A. Fenigstein, R. Reshef, S. Alfassi, G. Yehudian, 2013. Single-Exposure High Dynamic Range CMOS Image Sensorpixel With Internal Charge Amplifier, Google Patents.
- Fujita, K., Mori, H., Kyuushima, R., Honda, M., Yamamoto, K., Sensor, A.C.I., 2004. High resolution large formatted CMOS flat panel sensors for X-ray. In: *IEEE Nuclear Science Symposium Conference Record*, pp. 2114–2118.
- Heo, S.K., Park, S.K., Hwang, S.H., Im, D.A., Kosonen, J., Kim, T.W., Yun, S., Kim, H.K., 2010. Development of a large-area CMOS-based detector for real-time X-ray imaging. In: *Proc. SPIE Medical Imaging*, p. 76223.
- Kasap, S.O., Rowlands, J.A., 2002. Direct-conversion flat-panel X-ray image sensors for digital radiography. *Proc. IEEE* 90 (4), 591–604.
- Keoa, T.W.S., Kosonena, J., Hwanga, S.H., Kima, H.K.K., Yunb, S., 2011. 12-inch-wafer-scale CMOS active-pixel sensor for digital mammography(empty).pdf. In: *Physics of Medical Imaging, Proc. of SPIE*, p. 7961.
- Konstantinidis, A.C., Szafraniec, M.B., Speller, R.D., Olivo, A., 2012. The Dexela 2923 CMOS X-ray detector: a flat panel detector based on CMOS active pixel sensors for medical imaging applications. *Nucl. Instrum. Methods Phys. Res. A* 689, 12–21.
- Korthout, L., Verbugt, D., Timpert, J., Mierop, A., de Haan, W., Maes, W., de Meulmeester, J., Muhammad, W., Dillen, B., Stoldt, H., et al., 2009. A wafer-scale CMOS APS imager for

- medical X-ray applications. In: *Proc. of 2009 International Image Sensor Workshop*, vol. 70, pp. 22–28.
- Lahav, A., Leitner, T., Reshef, R., Fenigstein, A., 2011. Enhanced X-RAY CMOS sensor panel for radio and fluoro application using a low noise charge amplifier pixel with a partially pinned PD. In: IISW 2011.
- Mori, H., Kyuushima, R., Fujita, K., Honda, M., 2002. High resolution and high sensitivity CMOS panel sensors for X-ray imaging. In: 2001 IEEE Nuclear Science Symposium Conference Record (Cat. No.01CH37310). vol. 1, pp. 29–33.
- Nikl, M., 2006. Scintillation detectors for X-rays. *Meas. Sci. Technol.* 17 (4), R37.
- Rose, A., 1948. The sensitivity performance of the human eye on an absolute scale. *JOSA* 38 (2), 196–208.
- Samei, E., Flynn, M.J., 2003. An experimental comparison of detector performance for direct and indirect digital radiography systems. *Med. Phys.* 30, 608.
- E. Sarig, R. Reshef, S. Alfassi, D. Cohen, 2010. Horizontal Row Drivers for CMOS Image Sensor With Tiling on Three Edges, Google Patents.
- Teledynedalsa, n.d. <http://www.teledynedalsa.com/ls/products/cameradetail.aspx?partNumber=Xineos-1313-EO-CL>.
- Turchetta, R., Guerrini, N., Sedgwick, I., 2011. Large area CMOS image sensors. In: *Image Sensors Europe*.
- Yamashita, Y., Takahashi, H., Kikuchi, S., Ota, K., Fujita, M., Hirayama, S., Kanou, T., Hashimoto, S., Momma, G., Inoue, S., 2011. 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). IEEE International, pp. 408–410.
- Zentai, G., 2011. Comparison of CMOS and a-Si flat panel imagers for X-ray imaging. In: IEEE International Conference on Imaging Systems and Techniques (IST), pp. 194–200.

# Complementary metal-oxide-semiconductor (CMOS) and charge coupled device (CCD) image sensors in high-definition TV imaging

14

P. Centen

Grass Valley, Nederland BV, Breda, The Netherlands

## 14.1 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 three-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  $\frac{3}{3}$ -in. 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 three-imager  $\frac{3}{3}$ -in. 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).

☆ This chapter is a reprint of the chapter originally published in the first edition of "High Performance Silicon Imaging: Fundamentals and Applications of CMOS and CCD Sensors."

## 14.2 Broadcast camera performance

In assessing broadcast cameras there are >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, for example, 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 (CrCbs), 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 five 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 modulation transfer function (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 nonlinear. 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 et al., 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 7 years. On average service parts need to be available for 5–7 years. As a consequence, the same imagers must be available for >10–14 years. Many of the imager specifications relate to nonvisibility 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 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 on 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 nonuniformity (PRNU). Spatially there are column artifacts, row artifacts and large area artifacts.

Another offset error is the dark current induced FPN. 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 nonuniformity, is the difference between pixels and should be <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.

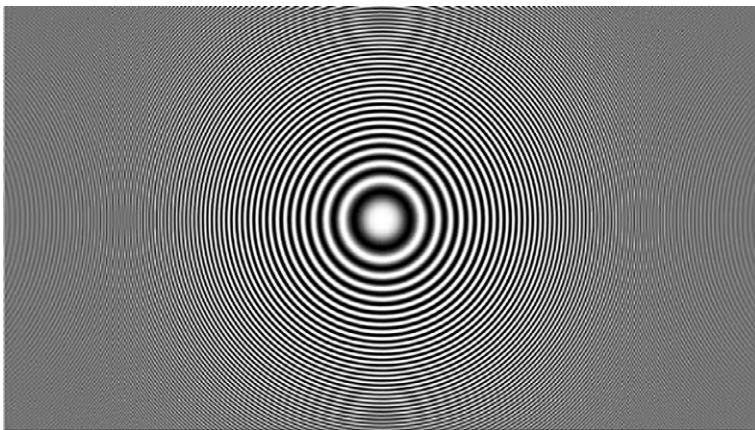
A broadcast video camera has a 0 dB setpoint and around that point the gain can be changed from, for example, -6 dB to +12 dB and beyond. The 0 dB is defined through the sensitivity and the noise of the camera. In SDTV the noise in 0 dB is around 60 dB in luminance ( $Y$ ) in a 5 MHz bandwidth and was 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  $f/1.4$  upward. 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.

## 14.3 Modulation transfer function, aliasing and resolution

The 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



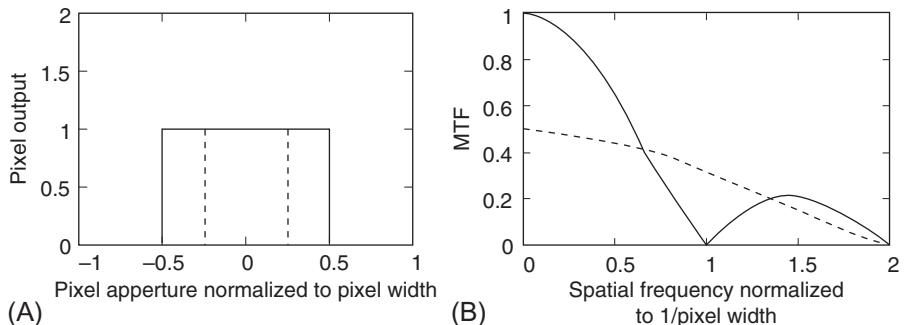
**Fig. 14.1** Zone chart; a two-dimensional spatial frequency sweep.

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 sub-harmonics. A practical method to judge MTF and aliasing of a camera system is the ZONE-chart (Drewery, 1979). The ZONE-chart (Fig. 14.1) is a two-dimensional (2D) 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 OLP 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 et al., 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 MTF of a pixel is a  $\sin(x)/x$  function with zero transfer slightly above the inverse of the pixel width. Fig. 14.2 depicts the effect pixel aperture (100% and 50%) has on MTF. In the monochrome case and with  $\mu$ -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 Bayered 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.



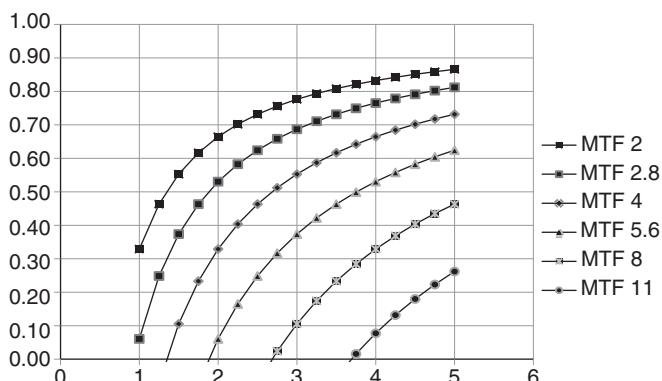
**Fig. 14.2** Solid trace: (A) a pixel aperture of 100% and (B) the MTF of the pixel. Dotted trace idem for aperture 50%.

The MTF of broadcast lenses is very well described with the diffraction limit equation

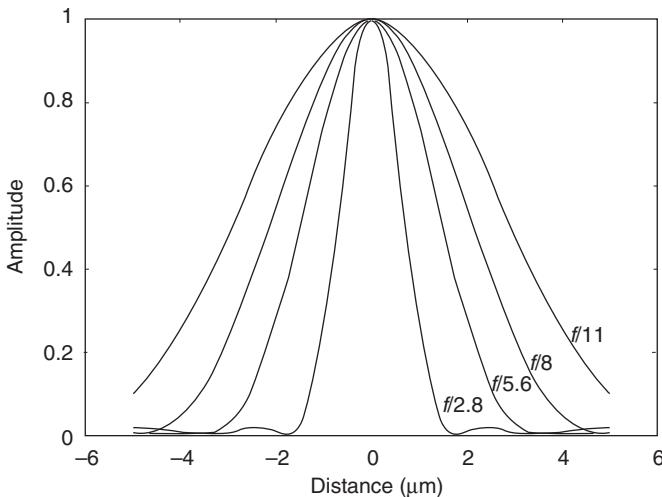
$$\text{MTF}(f_s) := 1 - 1.22 \cdot F \cdot \lambda \cdot f_s \quad (14.1)$$

where  $F$  is the  $f$ -number of the lens,  $\lambda$  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  $\mu$ -lens shift to be applied to the imager.

In Fig. 14.3 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  $5\text{ }\mu\text{m}$  pixel still has a MTF of slightly more than 25% and at  $f/5.6$  even a  $2.5\text{ }\mu\text{m}$  pixel can resolve with 25% at Nyquist. In Fig. 14.4 the pixel profile (Airy disc) at different  $f$ -numbers is given. The widening of the pixel profile for the higher  $f$ -numbers is obvious.

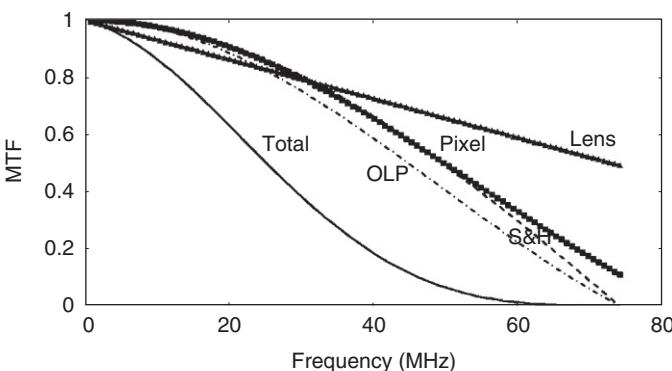


**Fig. 14.3** Y-axis: MTF at Nyquist ( $=1/(2 \times \text{pixel\_pitch})$ ), X-axis: the pixel pitch in  $\mu\text{m}$ . Curves have  $f$ -number as parameter.  $\lambda = 550\text{ nm}$ .



**Fig. 14.4** Airy disk for several  $f$ -numbers at 520 nm.

The MTF of a camera system is determined by the optics: lens, OLP filter, the pixel aperture, and the electrical processing: zero order hold, bandwidth limiting, low pass filtering, aperture correction. The camera 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. 14.5 the MTF of the individual elements in a camera chain is plotted. The horizontal axis is in MHz and the system clock is 74.25 MHz with Nyquist at 37.125 MHz. In a  $\frac{3}{3}$ -in. imager (5  $\mu\text{m}$  pixel) the 74.25 MHz frequency corresponds to 200 lp/mm and Nyquist is at 100 lp/mm. The OLP filter has a cosine shape, the pixel a  $\sin(x)/x$  just as the sample and hold.



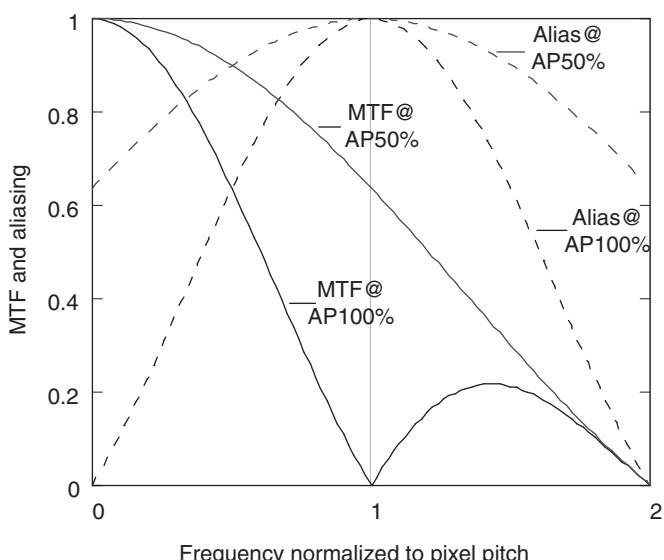
**Fig. 14.5** MTF curves of the electrical and optical elements in the camera. Lens at  $f/4$  and 520 nm.

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.

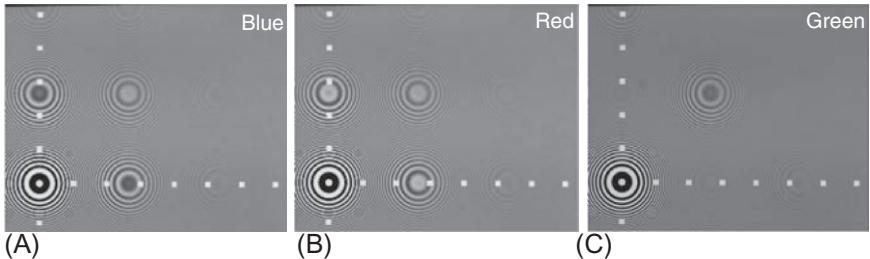
## 14.4 Aliasing and OLP filtering

The imager is a 2D spatial sampling device. In general the scene that is viewed does not comply to the Nyquist criteria: that is, 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  $\sin(x)/x$  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. [Fig. 14.6](#)



**Fig. 14.6** 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 is the spatial frequency normalized to the inverse of the pixel pitch (sample frequency). Hence Nyquist is at  $f_x = 0.5$ . AP = pixel aperture.



**Fig. 14.7** 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.

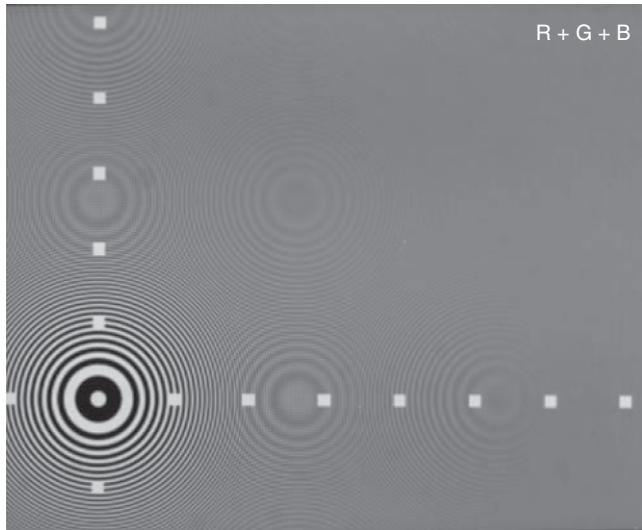
shows the MTF (solid) curve and its first alias (dotted) curve for a pixel with a 100% aperture (notch at  $f_s = 1$ ) and one with a 50% aperture (notch at  $f_s = 2$ ). The alias for the pixel with 100% aperture goes to zero for values near 0 Hz and the pixel with 50% aperture has 64% alias at 0 Hz.

With a birefringent crystal one can engineer an OLP filter. The MTF of the OLP 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 OLP filter has the same curve for the three colors.

In a layered, 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 it is different because of the quincunx pattern (Fig. 14.7). 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. 14.7). 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 canceled and is at a same position as the red and blue (Figs. 14.7 and 14.8).

In Fig. 14.7 the red, green and blue output of a layered imager, when exited with a zone chart (Fig. 14.1), 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.

Fig. 14.7 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. Fig. 14.7A and B show the expected aliases on a rectangular lattice. The green aliasing



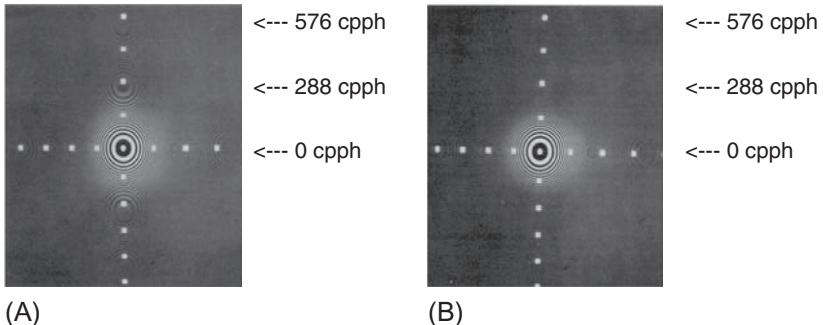
**Fig. 14.8** Two-dimensional frequency sweep. The RGB outputs of an imager with Bayer pattern (Fig. 14.7) are added to obtain the luminance signal. The aliasing in horizontal and vertical direction occur at half the monochrome lattice resolution and at the monochrome lattice resolution.

pattern is in quincunx (Fig. 14.7C). 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.

Fig. 14.8 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 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 OLP 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).



**Fig. 14.9** The effect of low pass filtering: (A) without optical low pass filtering and (B) with optical low pass filtering.

In cinematography there is a multitude of sampling structures found: rhombic pixels (Toyama et al., 2011, F65), diagonally striped color filters (Panavision) 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 (Panavision) 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. 14.9. 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  $\mu\text{m}$  pixel. Assume a diffraction limited lens at  $f/4$ , and an 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:

lens at  $f/4 = 1 - 1.22 * 4 * (520 * 10^{-9}) * (73 * 10^3) = 82\% @ 520 \text{ nm}$

$$OLP = \cos\left(\frac{\pi}{2} * \frac{73}{200}\right) = 84\%$$

$$\text{pixel} = \frac{\sin\left(\pi^* \frac{73}{200}\right)}{\pi^* \frac{73}{200}} = 80\%$$

$$S\&H = \frac{\sin\left(\pi^* \frac{27}{74.25}\right)}{\pi^* \frac{27}{74.25}} = 80\%$$

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 OLP 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:  $1920 \times 1080$  ([Thorpe, 2013](#)). This is why the three-imager HDTV video camera performed so well in a cinematographically environment.

## 14.5 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 transfer curve, 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.

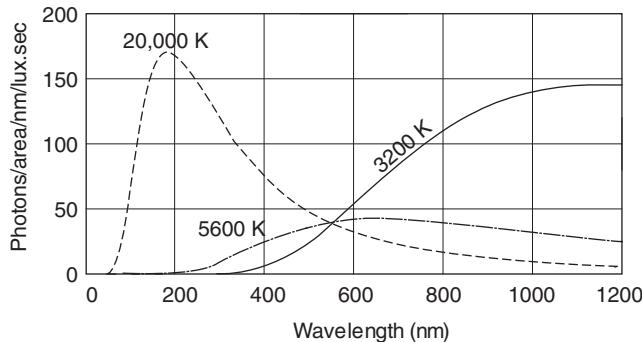
Also mechanical stability and flatness of the die in the package is important. In a three-imager camera, imagers are aligned to the color splitter within sub-pixel accuracy ( $\mu\text{m}$  range). It must stay within sub-pixel accuracy during its life time of some 5–7 years. In three-imager cameras 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  $\frac{2}{3}\text{-in}$ . On one hand the  $\frac{2}{3}\text{-in}$ . 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, back focus, and MTF.

When people say the imager is  $\frac{2}{3}\text{-in}$ . they do not mean  $\frac{2}{3} \times 25.4 = 16.9\text{ mm}$  but 11.0 mm, 1-in. is not 25.4 mm but 16.0 mm. It is a metric from the past, stemming from the tubes ([Table 14.1](#)). There are two flavors of  $\frac{2}{3}\text{-in}$ : in broadcast the image diagonal is 11.0 mm and in industrial vision it is 12.0 mm. Either one uses

**Table 14.1** Image diagonals in broadcast.

Format	Image diagonal
1"	16.0 mm
2/3"	11.0 mm
1/2"	8.0 mm



**Fig. 14.10** Photon flux of a blackbody radiator at 3200, 5600, and 20,000 K. Horizontal axis has the optical wavelength in nm. Vertical axis has the number of photons per area in  $\mu\text{m}^2$  per nm optical wavelength per lux.second.

16 mm/in or 18 mm/in for the conversion from inch to mm. The  $\frac{2}{3}$ -in. 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  $\frac{2}{3}$ - to  $\frac{1}{2}$ -in. the pixel area shrinks by a factor two. The  $\frac{2}{3}$ -in. imager has a 4 times smaller pixel area. When the  $\frac{2}{3}$ -in. imager would have  $f/11$  for a given exposure then the  $\frac{1}{2}$ -in. imager would need  $f/8$  to obtain that same exposure and the  $\frac{2}{3}$ -in. imager  $f/5.6$ . A 4 k,  $\frac{4}{3}$ -in. single-imager camera, and a three-imager  $\frac{2}{3}$ -in. 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; Krymski et al., 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. Fig. 14.10 shows the photonflux of a blackbody radiator at 3200, 5600, and 20,000 K. The unit of the vertical axis is in the number of photons per (pixel) area in  $\mu\text{m}^2$ , per nm of wavelength and per lux.second. 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 (400–480 nm). The blue channel receives almost 3 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 for blue is 0.0722 (Eq. 14.6) 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

$$\text{Photons } G(T) := \frac{A_{\text{cell}} \cdot T_{\text{int}}}{\eta \cdot h \cdot c} \cdot \frac{E_v \cdot \rho}{4 \cdot F^2} \cdot \frac{\int_{400}^{750} P(\lambda, T) \cdot \tau_{\text{IR}}(\lambda) \cdot \tau_{\text{lens}}(\lambda) \cdot \tau_{\text{KssG}}(\lambda) \cdot \lambda d\lambda}{\int_{400}^{750} V(\lambda) \cdot P(\lambda, T) d\lambda} \quad (14.2)$$

where  $P$  is the blackbody radiator with power density in  $\text{W/m}^3$ , the transmission curves  $\tau_{\text{IR}}$  for the IR filter,  $\tau_{\text{lens}}$  the transmission curve for the lens, and  $\tau_{\text{KssG}}$  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.  $\eta$  is a constant with value 683 lm/W,  $c$  the speed of light and Planck's constant  $h$ .  $A_{\text{cell}}$  is the image cell area which is  $25 \mu\text{m}^2$  for a HDTV pixel in a 3/5-in. imager and  $T_{\text{int}}$  the integration time of the pixel of. The scene illumination is  $E_v$  often equal to 2000 lux,  $\rho$  the reflection equal to 89.9% for the standardized measurement condition and finally  $F$  the  $f$ /number of the lens, for example,  $f/11$ .

Dividing Eq. (14.2) by Eq. (14.3) one arrives at the photon density in #photons/ $\mu\text{m}^2/\text{lux}\cdot\text{second}$  for a given color channel.

$$\xi := A_{\text{cell}} \cdot T_{\text{int}} \cdot \frac{E_v \cdot \rho}{4 \cdot F^2} \quad (14.3)$$

In Table 14.2 the photon densities in RGB and monochrome are tabulated for a few blackbody radiator temperatures. To calculate the number of photons in the RGB 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 relatively constant. Assuming an overall conversion efficiency ( $QE_{\text{eff}}$ ) of 60%,

$$QE_{\text{eff}} = QE^* \text{Aperture}^*(1 - \text{refelction})^* \mu_{\text{lens}} \text{gain} \quad (14.4)$$

then the charge accumulated in a  $5 \mu\text{m}$  green pixel under nominal illumination conditions of 2000 lux,  $f/11$ , 89.9%, 3200 K is  $G = 1818$  electrons. With an overexposure

**Table 14.2** Photon density as a function of black body radiator temperature.

Photon density	Number of photons/ $\mu\text{m}^2/\text{lux}\cdot\text{second}$			
	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
20,000	1218	2446	3277	8786

margin of 500% the maximum charge handling capacity,  $Q_{\max}$ , must be larger than:  $Q_{\max} > = 9.1 \text{ kel}$ . The number of signal electrons in [Table 14.3](#) are given for the RGB channel at three master gain 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 100,000 lux or more. At these lux levels the photon flux increases up to  $10^9 \text{ photons}/\mu\text{m}^2/\text{s}$ . This amounts to 16 f-stops more photons being generated than under nominal conditions.

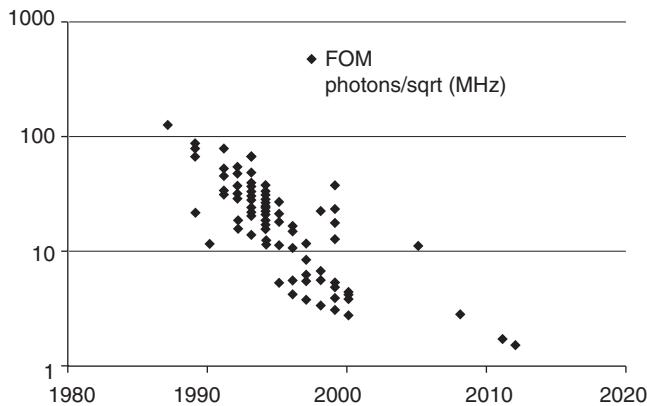
In [Fig. 14.11](#) the read-noise per unit of bandwidth divided by the quantum efficiency is plotted for broadcast cameras. One can calculate this figure of merit (FOM) from Eq. (14.5) by using the fact that the signal-to-noise in green is about 1.5 dB less than in Y. All the other numbers can be found in any brochure of a broadcast camera.

$$\text{FOM} = \frac{A_{\text{cell}} * T_{\text{int}} * E_v * \rho * 2.25}{4 * F^2 * \sqrt{BW} * 10^{\frac{|dBY| - 1.5}{20}}} \quad (14.5)$$

FOM in photons/(unit bandwidth),  $A_{\text{cell}}$  is pixel area in  $\mu\text{m}^2$ ,  $T_{\text{int}}$  is integration time in ms,  $E_v$  is scene illumination in lux,  $\rho$  is scene reflection, for broadcast standardized to 89.9%,  $F$  is aperture of the lens,  $BW$ : videobandwidth in MHz,  $dBY$  is the signal

**Table 14.3** Number of photons and electrons in a HDTV pixel of 5  $\mu\text{m}$  as a function of master-gain for 100% video-signal at 2000 lux, 89.9%, f/11, 3200 K, QE = 60% and exposure of 1/60 s

	Red	Green	Blue
0 dB	3206 photons 1924 electrons	3029 photons 1818 electrons	1114 photons 668 electrons
+6 dB	962 electrons	909 electrons	334 electrons
+12 dB	481 electrons	454 electrons	167 electrons



**Fig. 14.11** Noise and sensitivity improvements in broadcast cameras.

to noise in luminance channel in dB, and  $2.25$  is a number with dimension in  $\text{photons}/\mu\text{m}^2/\text{ms}/\text{lumx}$ .

Inspection of Fig. 14.11 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.

## 14.6 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 274 M standard describes a  $1920 \times 1080$  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 296 M describes a  $1280 \times 720$  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.

## 14.7 CCD and 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  $\frac{2}{3}$ -in. 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  $1920 \times 1080$ . 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 ( $2048 \times 1080$ ). 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  $\mu\text{s}$  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 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

**Table 14.4** Possible scanning formats at constant image diagonal, with 4320 vertical gates.

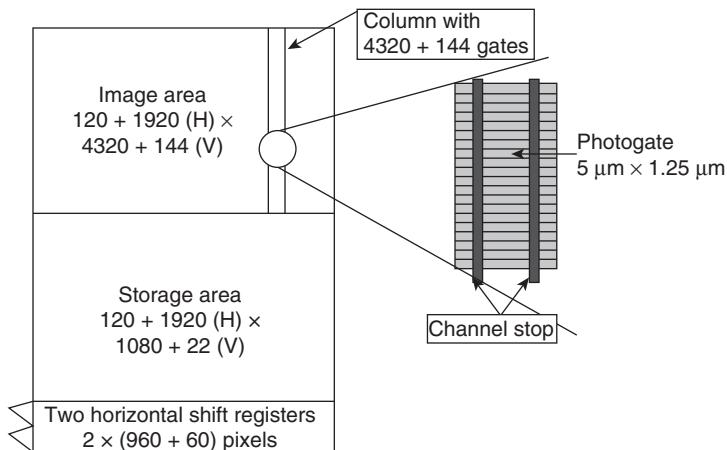
<b>16:9 scanning formats</b>	<b>Number of gates/pixel</b>	<b>In 12-phase clocking schema</b>
1080p	4 (=4320/1080)	4-phase clocking 1 pixel
720p	6 (=4320/720)	6-phase clocking 1 pixel
1080i	8 (=4320/(1080/2))	4-phase clocking 2 sub-pixels
480p	9 (=4320/480)	3-phase clocking 3 sub-pixels
576i	15 (=4320/(576/2))	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

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-in. 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/s. The imager reported in Honda et al. (2005) supports up to 720p96 at 118 Mpixel/s.

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 14.4) 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 Moeland et al. (1998).

Fig. 14.12 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.



**Fig. 14.12** Schematic representation of the HD-DPM imager.

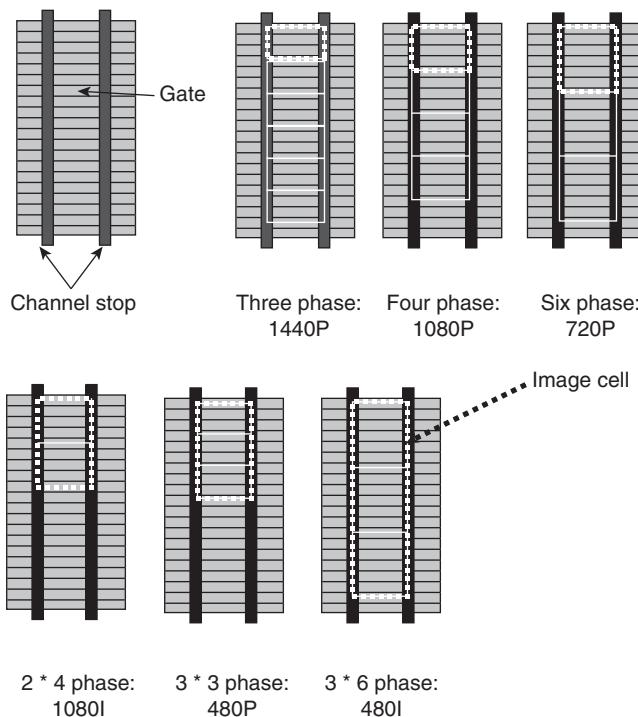
The storage area has a fixed number of storage sites and includes the over scan and black reference pixels  $2040 \text{ (H)} \times 1102 \text{ (V)}$ , and has a four-phase interconnect scheme. During the horizontal line blanking a (charge) TV line from the storage area is distributed over the two readout registers with the four-phases driven independently.

During the active line time, when horizontal transport is at  $2 \times 111.375 \text{ MHz}$ , the horizontal register is driven in quasi-two phase mode. 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. [Fig. 14.13](#) denotes the possible vertical scanning formats in 16:9 aspect ratios by binning several patterns.

The first column in [Table 14.4](#) 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. 14.13](#)). Note that in 1080i the image area is scanned in 1080p and only in the horizontal registers ([Fig. 14.12](#)) the interlaced image is generated by adding two TV lines of a 1080p image together. The active image width (H) is  $1920 \times 5 \mu\text{m}$  and the active image height (V) is  $4320 \times 1.25 \mu\text{m}$ . This results in an image diagonal of 11.0 mm for all scanning formats as required for a  $\frac{3}{4}\text{-in.}$  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  $\frac{3}{4}\text{-in.}$  HDTV CMOS imager, with  $5 \mu\text{m}$  pixels one needs a feature size of  $0.25 \mu\text{m}$  or less. [Loose et al. \(2001\)](#) published the first  $\frac{3}{4}\text{-in.}$  HDTV CMOS imager with soft-reset 3T-pixels in  $0.25 \mu\text{m}$  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

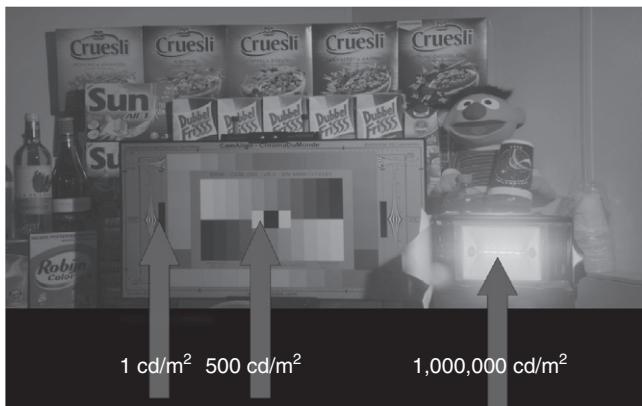


**Fig. 14.13** Various binning patterns.

patterns, like changing the state of the gates in each 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 feed trough 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. 14.14). Another attractive feature is the multiple nondestructive 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  $\frac{2}{3}$ -in. HDTV imagers became 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; Meynarts et al., 2011). Grass Valley was the first to successfully introduce a  $\frac{2}{3}$ -in. three-imager broadcast camera for use in live productions (LDX) with a global shutter CMOS imager (Centen et al., 2013).



**Fig. 14.14** What a CMOS pixel can do with highlights (Jeroen Rotte, Grass Valley, Nederland BV).

## 14.8 Signal-to-noise ratio

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

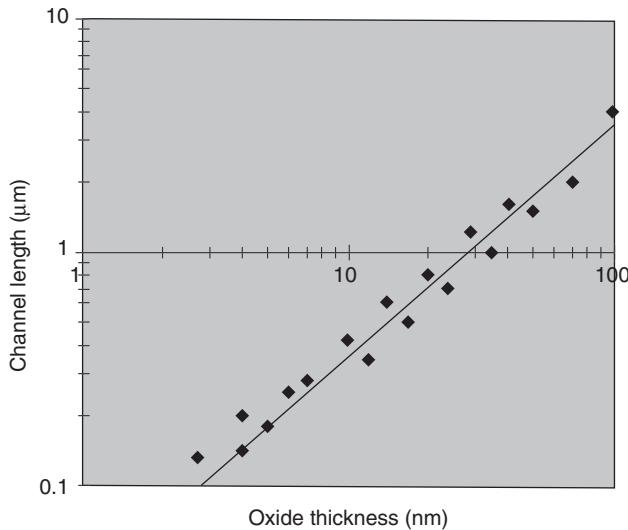
$$Y = 0.2126*R + 0.7152*G + 0.0722*B \quad (14.6)$$

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

$$SNY = SNG + 1.5dB \quad (14.7)$$

In SDTV the signal-to-noise in  $Y$  is stabilized at 60–62 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



**Fig. 14.15** Minimum channel length versus oxide thickness.

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. 14.15). The straight line is given by  $L_{\min} = 0.035 * d_{\text{ox}}$  ( $L$  in  $\mu\text{m}$  and  $d$  in nm).

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

$$\text{NED} = \frac{4kT}{g_m} * \left( \frac{C_{\text{tot}}}{q} \right)^2 \quad (14.8)$$

and the bandwidth during charge sensing as

$$F_{3\text{dB}} \equiv \frac{g_m}{2 * \pi * C_{\text{load}}} \frac{C_{\text{fd}}}{C_{\text{tot}}} \quad (14.9)$$

with  $g_m$  the transconductance of the source follower,  $C_{\text{tot}}$  the sum of all the capacitance connected to the floating diffusion, and  $C_{\text{fd}}$  known as the detection node capacitance. The capacitance at the source, loading the source follower, is  $C_{\text{load}}$ . Note that the 3 dB bandwidth of the source follower stage is reduced during charge sensing with the ratio  $C_{\text{fd}}/C_{\text{tot}}$ , 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  $C_{\text{tot}} = \frac{1}{3} \times C_g + C_{\text{fd}}$ , with  $C_g = WLC_{\text{ox}}$  where the source follower has gate capacitance  $C_g$ , channel length  $L$ , and channel width  $W$ . The capacitance per unit area is  $C_{\text{ox}}$ .

A first-order approximation for the transconductance  $g_m$ , using for the drain-source current  $I_{\text{ds}}$  the product of channel width and a maximum current density  $I_{\text{ds}} = W^*J_x$ , with maximum current density at  $J_x = 10 \text{ A/m}$ , and simplifying, is

$$g_m = \sqrt{2 * \frac{W}{L} * \frac{\epsilon}{d_{\text{ox}}} * u_n * I_{\text{ds}}} \approx \frac{C_g}{L} * \sqrt{\frac{d_{\text{ox}}}{L}} \approx \frac{C_g}{d_{\text{ox}}} \quad (14.10)$$

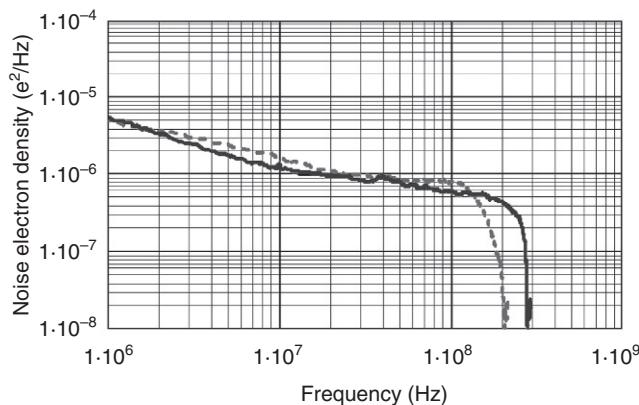
A minimum for the NED is at  $C_{\text{tot}} = 2 \times C_{\text{fd}}$ , or equivalently the gate capacitance of the source follower should be  $C_g = 1.5 \times C_{\text{fd}}$ .

Often the floating diffusion ( $C_{\text{fd}}$ ) is determined by other design parameters that relate to the CCD channel and the CCD pixel. The total capacitance of the detection node  $C_{\text{tot}}$  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  $C_{\text{tot}}/C_{\text{fd}}$  is 2 and the bandwidth ( $F_{3\text{dB}}$ ) increases with increased transconductance of the source follower. Halving oxide thickness  $d_{\text{ox}}$  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).

Fig. 14.16 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  $\text{NED} = 0.75e^2/\text{MHz}$  at 37 MHz and at 112 MHz,  $\text{NED} = 0.59e^2/\text{MHz}$ . After CDS the new amplifier delivered a noise level of 8e in 30 MHz bandwidth.

In Table 14.5 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 4.6–2 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  $< 1 \text{ MHz}$ .

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



**Fig. 14.16** Static noise measurement of the on-chip amplifier. Dotted curve prior amplifier; solid curve the new amplifier.

is on a  $\mu\text{s}$  time scale, while the CCD pixels are on an  $\text{ns}$  time scale and hence the noise is generated in a much larger bandwidth.

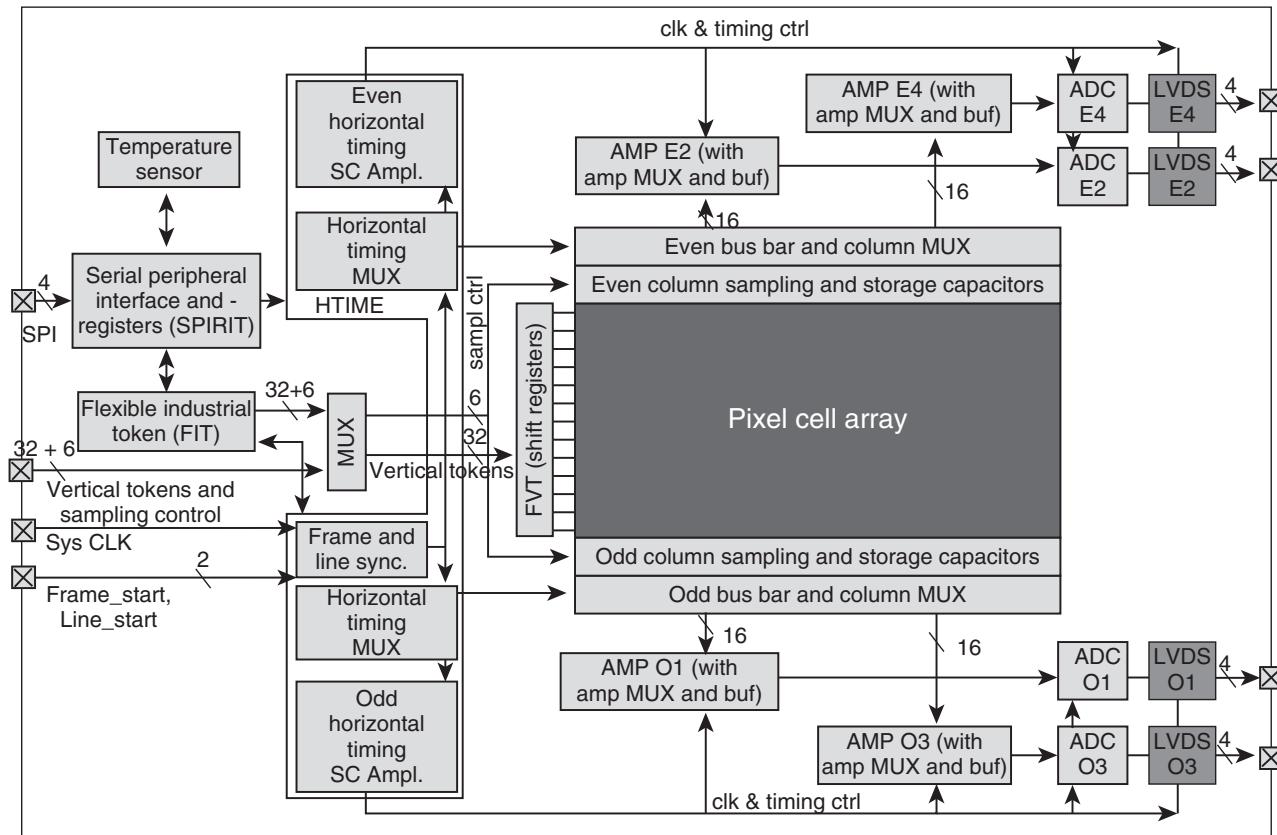
Given a CCD imager designed for single ( $1 \times$ ) and triple speed ( $3 \times$ ) operation. Compared with single-speed operation, the sensitivity in triple-speed is reduced due to the shortening of the integration (light exposure) time per frame, additionally affected by the increase in the equivalent noise level originated for the same frame, and both due to the increase in the video bandwidth from 30 to 90 dB. The SNR in  $3 \times$  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 by the kTC-noise at the column (with  $C$  the column capacitor). The kTC is the same in  $1 \times$  and in  $3 \times$  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  $3 \times 2.2$  Mpixel CMOS imagers have been used in  $\frac{3}{5}$ -in. 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 5 T pixel.

In Fig. 14.17 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  $1920 (\text{H}) \times 1080 (\text{V})$ . 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.

**Table 14.5** State-of-the-art amplifier noise performance.

Type	Ref	Reset frequency	Amplifier type	Conversion gain	Noise after CDS/ $\sqrt{\text{reset frequency}}$
FF- CCD CMOS	Burke et al. (1997)	100 kHz	Source follower	20 $\mu\text{V/e}$	6.3 e/ $\sqrt{\text{MHz}}$
	Krymski et al. (2003)	50 kHz	Source follower +gain	60 $\mu\text{V/e}$	6.3 e/ $\sqrt{\text{MHz}}$
FF- CCD CMOS	Draijer et al. (2005)	25 MHz	Source follower	40 $\mu\text{V/e}$	2.8 e/ $\sqrt{\text{MHz}}$
	Kozlowski et al. (2005)	104 kHz	Source follower +gain	—	46 e/ $\sqrt{\text{MHz}}$
CMOS	Yoshihara et al. (2006)	156 kHz	Source follower +gain	40 $\mu\text{V/e}$	17.7 e/ $\sqrt{\text{MHz}}$
	Takahashi et al. (2007)	156 kHz	Source follower +gain	75 $\mu\text{V/e}$	11.6 e/ $\sqrt{\text{MHz}}$
CMOS	Cho et al. (2007)	625 kHz	Source follower +gain	101 $\mu\text{V/e}$	10.4 e/ $\sqrt{\text{MHz}}$
FT- CCD	This amplifier	111 MHz	Source follower	18 $\mu\text{V/e}$	1.3 e/ $\sqrt{\text{MHz}}$

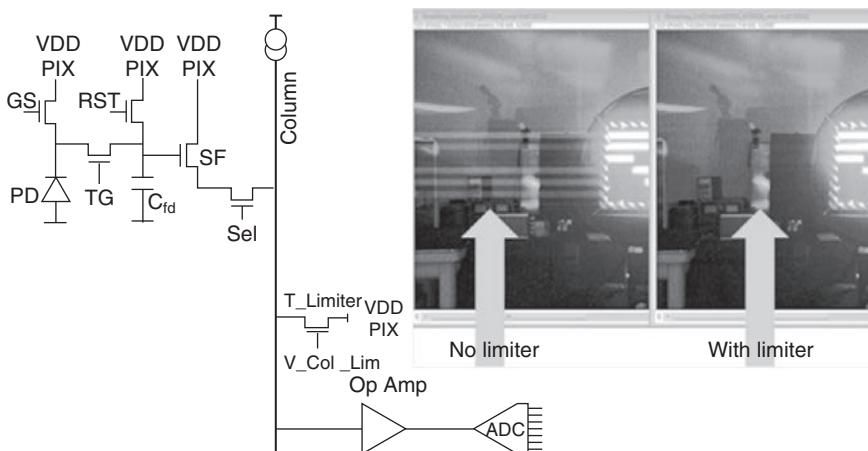


**Fig. 14.17** Block diagram CMOS imager (Steffen Lehr, Viimagic).

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)  $\times$  1080 (V) at 120 frames/s and the kTC noise reduced output is in 1920 (H)  $\times$  1080 (V) at 60 frames/s. The 4e read-noise is achieved at a moderate analog gain of  $2\times$ . 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 ([Shimamoto et al., 2012](#); [Solhusvik et al., 2011](#)) reducing the maximum output swing that can be reached. The imager depicted in [Fig. 14.17](#) is fabricated in 0.18  $\mu\text{m}$  1P4M technology, with light shield, and can support 1920 (H)  $\times$  1080 (V) up to 240 frames/s, which is 120 frames/s after DDS.

The 5 T-pixels ([Fig. 14.18](#)) 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.

After sampling at the column capacitor, the signal is multiplexed onto one of the 16 switch-capacitor-amplifiers (SCAs). There are  $4 \times 16$  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 signaling (LVDS) transmission.



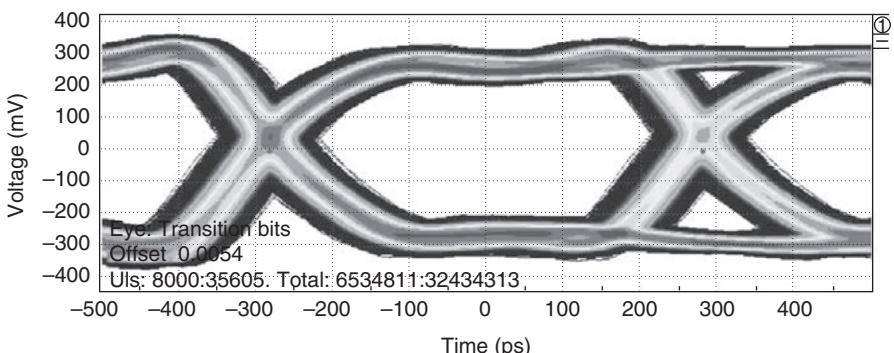
**Fig. 14.18** 5T-pixel with limiter transistor; T\_Limiter, at the column to reduce streaking (Jeroen Rotte, Grass Valley, Nederland BV).

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  $1920 \times 1080\text{Q}480$  which is  $1920 \times 1080\text{Q}240$  after DDS, doubling the frame rate, without the need for a rhombic-shaped pixel (Toyama et al., 2011).

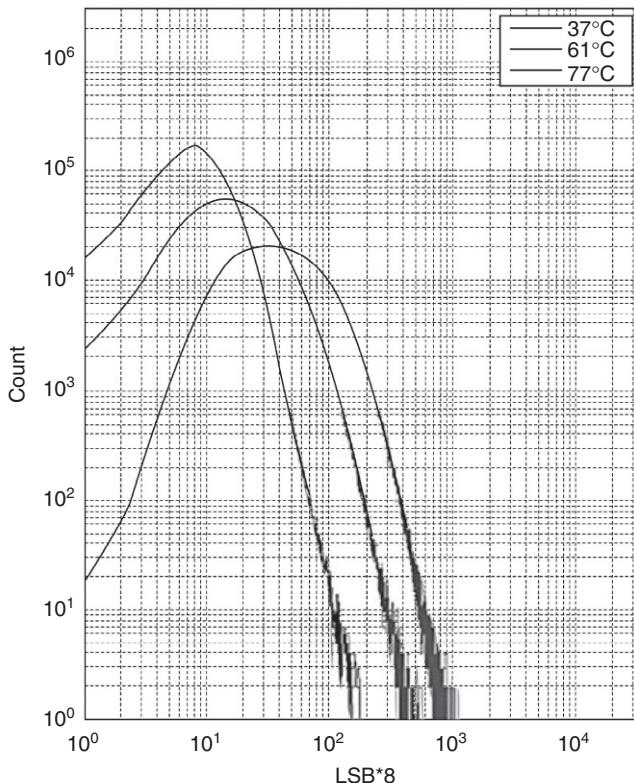
To reduce image artifacts (Purcell et al., 2009) at highlights, each column has a n-type metal-oxide-semiconductor (NMOS) transistor, with an adjustable gate voltage. It reduces the voltage swing at the input of the analog gain stages and the input of the ADC, preventing overload. The additional (limiter) transistor is shown in Fig. 14.18 in relation to the 5 T-pixel. Each column has a transistor, T\_Limiter, that limits the negative going signal excursions through an adjustable voltage at the gate ( $V_{\text{Col\_Limi}}$ ). The source of T\_Limiter is connected to the column and the drain to  $VDD_{\text{PIX}} = 3.3$  V. 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 known 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. 14.19 an eye diagram is measured of an LVDS lane at 1.782 Gbps. The maximum supported data rate of the  $4 \times 4$  lanes is 28.5 Gbps.

A dark current histogram of the floating diffusion at  $37^{\circ}\text{C}$ ,  $61^{\circ}\text{C}$ , and  $77^{\circ}\text{C}$  is shown in Fig. 14.20. The median value for the floating diffusion dark current is 114 e/s at  $37^{\circ}\text{C}$ , 282 e/s at  $61^{\circ}\text{C}$ , and 720 e/s at  $77^{\circ}\text{C}$ , a doubling of the dark current of about  $11^{\circ}\text{C}$ . The dark current of the photodiode doubles approximately with 6C: 2.5 e/s at  $37^{\circ}\text{C}$ , 38 e/s at  $61^{\circ}\text{C}$ , and 230 e/s at  $77^{\circ}\text{C}$ . 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. 14.21) from an early test device.



**Fig. 14.19** Eye diagram of an LVDS lane at 1.782 Gbps (Ruud van Ree, Grass Valley, Nederland BV).



**Fig. 14.20** Histogram of the floating diffusion dark current at 37°C, 61°C, and 77°C (Jeroen Rotte, Grass Valley, Nederland BV).

The scanning mode was  $1920 \times 1080$ p60 at an integration time of 1 ms. The ventilator has no rolling shutter distortion, only motion blur. The chip specification, measured at  $1920 \times 1080$ p60 after DDS (imager running at  $1920 \times 1080$ p120), is summarized in Table 14.6. The pixel fill factor without  $\mu$ -lens is 44 and the conversion gain 95  $\mu$ V/e. The read-noise after DDS is 4e at 27°C with an analog gain of  $2 \times$ . Power dissipation for the imager running at  $1920 \times 1080$ p120 is 1.1 W. The size of the active image array is  $10.34 \times 5.5$  mm. Chip size is  $203 \text{ mm}^2$ , mounted in a customer defined ceramic package  $\mu$ PGA-185. With this CMOS imager the transition from CCD to CMOS was a fact.

## 14.9 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!



**Fig. 14.21** Image after DDS at  $1920 \times 1080\text{p}60$  in global shutter mode at 1 ms integration time (Jeroen Rotte, Grass Valley, Nederland BV).

**Table 14.6** Some performance parameters of Xensium-FT.

Process	CMOS 0.18 $\mu\text{m}$ 1P4M with light shield
Supply voltages	3.3 V and 1.8 V
Chip size	203 $\text{mm}^2$
Number of light sensitive pixels and overscan pixels	$1920 + 152 (\text{H}) \times 1080 + 24 (\text{V})$
Pixel size	$5 \mu\text{m} \times 5 \mu\text{m}$
Transistors per pixel	5 T
Pixel fill factor	44% w/o micro-lens
Analog to digital conversion LVDS	4 ADCs 4*4 lanes
Maximum data rate per LVDS lane	1.782 Gbps/lane 28.5 Gbps all lanes
Maximum frame rate	120 frames/s with DDS 240 frames/s with DDS quincunx
PRNU	<1%
Conversion gain FD	90–100 $\mu\text{V/e}$
Idark FD @ 61°C	282 e/s; 0.18 $\text{nA}/\text{cm}^2$
Idark PD@ 61°C	38 e/s; 24 $\text{pA}/\text{cm}^2$
Read noise global shutter mode, after DDS @ 27°C	4 e at gain = 2 $\times$
Sensitivity in green (495 nm–573 nm)	2170 el.lux-s/ $\text{um}^2$
Qmax	>15 kel
Power dissipation	1.6 W @ 120 frames/s, DDS 1.1 W @ 60 frames/s, DDS

Therefore, the number of bits needed to quantize the number of electrons (Fig. 14.22) in the pixel is

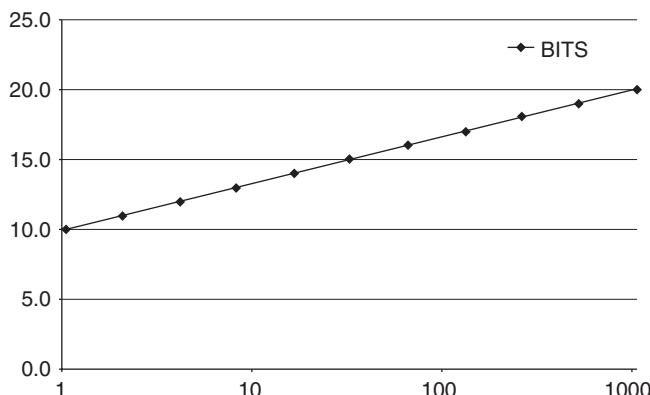
$$N = \frac{\log(Q_{\max})}{\log(2)} \quad (14.11)$$

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 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 4 k scanning! But with that same token a full HDTV three-imager camera could be regarded as a 6 k 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  $4 \text{ k} \times 2 \text{ k}$  each and



**Fig. 14.22** 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 14.7** 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 4 k)	Square pixels
UHDTV 1	3840 × 2160	7680 × 4320 (aka 8 k)	Square pixels
UHDTV 2	7680 × 4320	15,360 × 8640	Square pixels

the number of pixels is 17.67 Mpix. The approach of NHK is as it should be: in 4 k one needs in a three-imager camera a total of  $3 \times 4\text{ k} \times 2\text{ k} = 24$  Mpixels which for a single imager Bayer camera should  $4\text{ k} \times 4\text{ k} \times 2\text{ k} = 32$  Mpixel (Kitamura, 2011; Takayanagi et al., 2009) and for 8 k it is  $3 \times 32$  Mpixel (Shimamoto et al., 2012) or  $4\text{ k} \times 32\text{ k} = 128$  Mpixel in Bayer. In Table 14.7 a resolution comparison is made between single imager and three-imager cameras.

## 14.10 Three-dimensional and ultrahigh-definition (UHD) television

### 14.10.1 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.

### 14.10.2 UHD television (UHDTV) level 1 and level 2

UHDTV levels 1 and level 2 are the new topics. Level 1 is about 4 k resolution (in RGB) and level 2 is 8 k (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/s), 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, that is, 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 RGB pixels, must be limited. The cross talk determines the reduction of the color space. 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 4 k RGB means literally that in the single imager case it has to be  $4\text{ k} \times 4\text{ k} \times 2\text{ k} = 32\text{ Mpixels}$  (photo-sites) and 8 k RGB will need 4 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.

## 14.11 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.

## 14.12 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](http://www.smpte.org)). Another, European based body is the International Telecommunications Union (ITU—[www.itu.int](http://www.itu.int)) whose broadcasters are represented by the European Broadcasters Union ([www.ebu.ch](http://www.ebu.ch)).

### *Test charts*

The following two companies deliver test charts for cameras: DSC-Labs ([dsclabs.com](http://dsclabs.com)) and Esser ([www.image-engineering.de](http://www.image-engineering.de)).

### *Books*

[Poynton \(2003\)](#) discusses many video aspects of cameras and [Theuwissen \(1995\)](#) the imagers.

### Standardization

ITU-R BT.709. Parameter values for the HDTV standards for production and international program exchange.

SMPTE 274 M standard describes a  $1920 \times 1080$  matrix of light sensitive pixels.

SMPTE 296 M describes a  $1280 \times 720$  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, “Ultrahigh definition television—image parameter values for program production,” 2009.

SMPTE ST 2048-1 2011 “ $2048 \times 1080$  and  $4096 \times 2160$  digital cinematography production image formats.”

ITU-R Recommendation BT2020, “Parameter values for ultra-high definition television systems for production and international program exchange,” International Telecommunications Union, Geneva, 2009.

## References

- Asano, T., et al., 2011. A  $\frac{2}{3}$ -type 2.3-Mega Pixel IT-CCD for HD 1080p60. In: 2011 International Image Sensor Workshop, Hokkaido.
- Blackwell, H.R., 1946. Contrast thresholds of the human eye. *J. Opt. Soc. Am.* 36 (11), 624–643.
- Blankevoort, J., et al., 1994. A high performance, full-bandwidth HDTV camera applying the first 2.2 Milion pixel frame transfer CCD sensor. *SMPTE J.* 103, 319–329.
- Bosiers, J.T., et al., 2002. Frame transfer CCDs for digital still cameras: concept, design, and evaluation. *IEEE Trans. Electron Devices* 49 (3), 377–386.
- Bosiers, J.T., et al., 2006. Technical challenges and recent progress in CCD imagers. *Nucl. Inst. Methods Phys. Res. A* 565, 148–156.
- Burke, B.E., et al., 1997. Soft X-ray CCD imagers for AXAF. *IEEE Trans. Electron Devices* 44 (10), 1633–1642.
- Centen, P., 1991. CCD on-chip amplifiers: noise performance versus MOS transistor dimensions. *IEEE Trans. Electron Devices ED-38* (5), 1206–1216.
- Centen, P., et al., 1995. Dynamic pixel management for second generation aspect ratio switching. In: Montreux International Television Symposium Record, 8–13 June/ pp. 199–211.
- Centen, P., et al., 2001. A multi-format HDTV camera head. *SMPTE J.* 110 (8), 510–516.
- Centen, P., et al., 2007. A  $\frac{2}{3}$ -inch CMOS image sensor for HDTV applications with multiple high-DR modes and flexible scanning. In: ISSCC Dig. Tech. Papersvols. 512–513.
- Centen, P., et al., 2009. A  $\frac{2}{3}$ -inch low noise HDTV FT CCD-imager for 1080i180, 1080p90 and 720p120 scanning at constant image diagonal. In: International Imager Workshop (IISW2009), Bergen, Norway.
- Centen, P., et al., 2013. A 4e-noise  $\frac{2}{3}$ -inch global shutter  $1920 \times 1080$ P120 CMOS-imager. In: International Image Sensor Workshop (IISW2013), Snow-Bird.
- Chen, Y., et al., 2012. A 0.7erms-temporal-readout-Noise CMOS image sensor for low-light-level imaging. In: 2012 IEEE International Solid-State Circuits Conference, pp. 384–385.
- Cho, K.-B., et al., 2007. A inch 8.1Mpixel CMOS image sensor for digital cameras. In: IEEE International Solid-State Circuits Conference Dig. Tech. Papers, pp. 508–509.
- Draijer, C., et al., 2005. Third generation large area professional DSC CCD process and design technology. In: International Image Sensor Workshop, Karuizawa Japan. vol. 64.

- Drewery, J.O., 1979. The zone plate as a television test pattern. *SMPTE J.* 88, 763–770.
- Fasoli, L., Sampietro, M., 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, K., et al., 1992. A 2/3-inch FIT 2M pixel FIT-CCD HDTV image sensor. In: Proceedings of the IEEE International Solid-State Circuits Conference, pp. 170–171.
- Honda, T., et al., 2005. Development of 2/3-type 1-mega pixel progressive scan CCD for HDTV capable of high frame rate of 96fps. In: Proc. of the IEEE AIS Workshop, Kuruizawa, pp. 177–180.
- Hynecek, J., 1984. Design and performance of a low-noise charge detection amplifier for VPCCD devices. *IEEE Trans. Electron Devices* ED-31 (12), 1713–1719.
- Janesick, J.R., Klaassen, K.P., Elliott, T., 1987. Charge-coupled-device charge-collection efficiency and the photon transfer technique. *Opt. Eng.* 26 (10), 972–980.
- Kitamura, K., 2011. A 33 Mpixel, 120 fps CMOS image sensor for UDTV application with two-stage column-parallel cyclic ADCs. In: International Image Sensor Workshop (IISW2011), Hokkaido, Japan.
- Kozlowski, L., et al., 2005. A progressive 1920 × 1080 imaging system-on-chip for HDTV cameras. In: IEEE International Digest of Technical Papers. Solid-State Circuits Dig. Tech. Papers, pp. 358–359.
- Krymski, A., et al., 2003. A 2 e noise 1.3 Megapixel CMOS sensor. In: Proceedings of the IEEE AIS Workshop.
- Lang, H., 2001. Colour and Its Reproduction, Part 1: Colorimetry. Muster-Schmidt Verlag, Mainz, Germany.
- Loose, M., et al., 2001. 2/3-inch CMOS imaging sensor for high definition television. In: IEEE Workshop on Charge-coupled Devices and Advanced Image Sensors, 7–9 June.
- Martinez-Verdu, F., et al., 2002. Calculation of the color matching functions of digital cameras from their complete spectral sensitivities. *J. Imag. Sci. Technol.* 46, 15–58.
- Meynants, G., et al., 2011. Backside illuminated global shutter CMOS image sensors. In: International Image Sensor Workshop (IISW2011), June, Hokkaido, Japan.
- Moeland, M., et al., 1998. A high speed sports action camera system. In: Technical Papers International Broadcast Convention, Amsterdam. vols. 11–15, pp. 86–92.
- Morimoto, M., et al., 1994. A 2/3-inch 2 M-pixel IT-CCD image sensor with individual p-wells for separate V-CCD and H-CCD formation. In: IEEE International Solid-State Circuits Conference. Dig. Tech. Papers, pp. 222–223.
- Nitta, Y., et al., 2006. High-speed digital double sampling with analog CDS on column parallel ADC architecture for low-noise active pixel sensor. In: ISSCC. Dig. Tech. Papers, vols. 500–501.
- Peek, H.L., 1999. A low dark current double membrane poly-Si FT-technology for 2/3 inch 6M pixel CCD imagers. In: International Electron Devices Meeting. Technical Digest, pp. 871–874.
- Poynton, C., 2003. Digital Video and HDTV Algorithms and Interfaces. Morgan Kaufmann, San Francisco.
- Purcell, M., et al., 2009. Image artifacts caused by pixel bias cells in CMOS imagers targeted for mobile applications. In: International Image Sensor Workshop, Norway. (Invited Paper 73).
- Sakakibara, M., et al., 2012. An 83dB-dynamic-range single-exposure global-shutter CMOS imager sensor with in-pixel dual storage. In: International Solid-State Circuits Conference Dig. Tech. Papers, pp. 380–381.
- Salmon, O.E., 2012. The impact of LED lighting on colour rendition in TV and other cameras. In: Image Sensors, London.

- Shimamoto, H., et al., 2012. 120 Hz-frame-rate SUPER Hi-VISION capture and display devices. In: SMPTE Conference, Hollywood.
- Solhusvik, J., et al., 2011. A 1.2 MP  $\frac{1}{3}$ -inch global shutter CMOS image sensor with pixel-wise automatic gain selection. In: International Image Sensor Workshop, Japan, (Paper 53).
- Spitzer, R., et al., 1996. Design and implementation of a 3-CCD state of the art, 750-line HDTV progressive scan broadcast camera. In: Proc. NAB, Las Vegas, pp. 171–178.
- Stoldt, H., et al., 1996. CCD imagers for broadcast applications. In: Technical Digest IEDM96, San Francisco, 8–11 December, pp. 899–902.
- Takahashi, H., et al., 2007. A low-noise CMOS image sensor for full HD camcorders. In: ISSCC Dig. Tech. Papers, vols. 510–511.
- Takayanagi, I., et al., 2009. Progress in 1.25-inch digital-output CMOS image sensor developments for UDTV application. In: International Image Sensor Workshop, Norway.
- Teranishi, N., et al., 1982. No image lag photodiode structure in the interline CCD image sensor. In: Technical Digest IEDM, 13–15 Dec, pp. 324–327.
- Theuwissen, A.J., 1995. Solid-State Imaging With Charge-Coupled Devices. Kluwer Academic Publishers, Dordrecht.
- Theuwissen, A.J., et al., 1991. A 2.2 Mpixel FT-CCD imager, according to the Eureka HDTV-standard. In: IEDM Tech. Dig, pp. 167–170.
- Thorpe, L.J., 2000a. Technical aspects of the new world of multiformat DTV embodying progressive, interlaced, and segmented video frame format. SMPTE J., 713–719.
- Thorpe, L.J., 2000b. Acquisition in 24p and 35mm film: a technical enquiry. In: 142nd SMPTE Conference, Pasadena.
- Thorpe, L.J., 2013. New optical and image sensor technologies for high-resolution digital cinematography. SMPTE Mot. Imag. J. 122, 33–41.
- Thorpe, L.J., et al., 1988. New advances in CCD imaging. SMPTE J. 97, 378–387.
- Toyama, T., et al., 2011. A 17.7Mpixel 120fps CMOS image sensor with 34.8Gb/s readout. In: ISSCC Dig. Tech. Papers, Feb, pp. 420–421.
- Van Rooy, J., et al., 2002. The VIPER filmstream camera: a marriage between film and video. In: Proceedings of the Broadcast Asia Technical Conference and Exhibition, Singapore, 18–22 June, pp. 1–10.
- Wang, X., et al., 2010. A 2.2M CMOS image sensor for high speed machine vision applications. Proc. SPIE. 7536.
- Wong, H.-S., 1996. Technology and device scaling considerations for CMOS imagers. IEEE Trans. Electron Devices ED-43 (12), 2131–2142.
- Yoshihara, S., et al., 2006. A 6.4MPixel 60frames/s CMOS image sensor with seamless mode change. In: ISSCC Dig. Tech. Papers, pp. 492–493.

This page intentionally left blank

# High-performance silicon imagers, back illumination using delta and superlattice doping, and their applications in astrophysics, medicine, and other fields

15

S. Nikzad, M.E. Hoenk

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, United States

## 15.1 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 extrasolar planets, and new windows into our solar system and the universe have been made possible using high-performance silicon imagers in Earth-orbiting telescopes such as the Hubble, Galaxy Evolution Explorer (GALEX), and Kepler Space Telescopes. In addition, instruments in missions visiting or orbiting our solar system planets such as Cassini (mission to Saturn), Galileo (mission to Jupiter), Mars Exploration Rover (MER), 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. Their presence is so commonplace that we routinely record and transmit images and video around the world. Digital imagers capture fleeting precious moments, instantly record data such as faces of people we have just met, business cards, prices of items, coupons, boarding passes, recipes, and more broadly, generally aid 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 cameras in consumer applications, medicine, and scientific instruments 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.

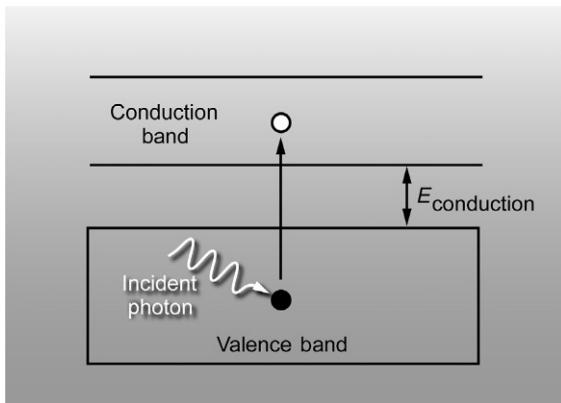
Starting from sand, the microelectronics industry manufactures nearly perfect single-crystalline silicon wafers and transforms them into silicon microprocessors, digital imagers, and high-speed digital video recording detectors. Whereas other semiconductor materials offer higher mobility (e.g., Ge, GaAs), direct bandgap (e.g., GaAs, InAs, InSb), and heterojunctions (e.g., GaAs/Al<sub>1-x</sub>GaxAs), the successes of the microelectronics and digital imaging industries are due to the abundance and the low cost of silicon starting material, the exceptional quality of crystalline silicon, and the reliability, stability, and repeatability of silicon surfaces and interfaces passivated with silicon oxide.

Silicon oxide has proven to be an indispensable 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<sub>2</sub> revolution and, because of their unique properties, created their own revolution in widespread imaging device applications.”

In this chapter, we discuss principles of operation in solid-state imaging followed by a short description of various electronic readout schemes and structures used today in imaging devices. We then focus on scientific imagers, including photon counting detectors and the role of back illumination in the design of high-performance silicon imagers, in particular focusing on delta and superlattice doping of back-illuminated detectors. In the remainder of the chapter, we briefly discuss applications in astronomy and planetary science, ending with a brief note on biomedical applications and an example in commercial, machine vision.

## 15.2 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. 15.1). When light enters a semiconductor detector, some fraction of incident photons is absorbed. If the photon energy is equal or greater than the semiconductor bandgap, the absorption of a photon creates an electron-hole pair in the semiconductor. Electron-hole pairs are then separated by the electric field in the detector. Depending on the architecture or readout scheme of the imager design, the resulting current or accumulated charge is measured by readout circuits that are either attached to or built into the detector. Silicon detectors are typically fabricated in high purity 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 occur in this epilayer. Therefore, the quality of this epilayer



**Fig. 15.1** Principle of photon absorption and carrier generation in a semiconductor.

(defect density, doping level, and uniformity) is 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 uniformity of the response are among the most important parameters of imager performance.

QE is the measure of detector's sensitivity to incident photons. In practice, detectors measure electrons, so the QE is often defined as the number of detected electrons divided by the number of incident photons at a given wavelength. As explained below, this definition leads to the counterintuitive result that QE can be greater than unity for high-energy photons. 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 QE:

$$\text{QE} = \left( \frac{\text{absorbed photons}}{\text{incident photons}} \right) \left( \frac{\text{photons generating e-h pairs}}{\text{absorbed photons}} \right) \left( \frac{\text{photoelectrons generated}}{\text{photons generating e-h pairs}} \right) \left( \frac{\text{collect eelectrons}}{\text{photoelectrons generated}} \right)$$

or

$$\text{QE (measured)} = T \times P \times \text{QY} \times \text{QE}_{\text{internal}}$$

where  $T$  is fraction of incident photons that are absorbed in the detector. For wavelengths shorter than about 600 nm,  $T$  is mostly affected by reflectance from the silicon surface, and can be approximated by the transmittance of incident photons into the silicon (semiconductor).  $T$  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 most cases,  $T$  is approximated as  $1 - R$  where  $R$  is the surface reflectivity, which is a function of photon energy. For wavelengths longer than about 600 nm,  $T$  must also account for losses due to incomplete absorption due to the finite thickness of the silicon detector.

$P$  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; QY (quantum yield) is the average number of electron-hole pairs generated for each absorbed photon. (QY) is can be defined as

$$QY = E_{\text{photon}}/E_{\text{e-h}}$$

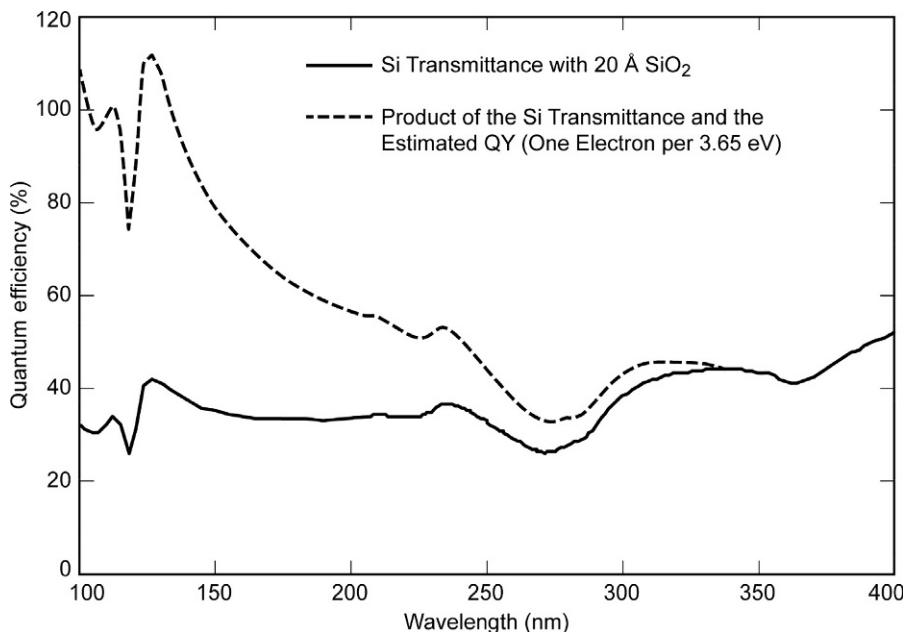
where  $E_{\text{photon}}$  is the energy of the incident photon and  $E_{\text{e-h}}$  is the average energy required by the electron-hole (e-h) pair. QY is wavelength dependent as well as material dependent, but a commonly used (and *very* rough) estimate of QY in silicon is  $E_{\text{ph}}/3.65 \text{ eV}$ , corresponding to the production of one electron-hole pair (on average) for each 3.65 eV of photon energy (Wilkinson et al., 1982). It should be emphasized that this is an oversimplification, especially in the ultraviolet where QY is a function of the photon energy and only a few electrons are produced in a stochastic process. Furthermore, QY has been shown in silicon to deviate from this approximation for lower energy (sub keV) electrons as well as charged and neutral particles (Funsten et al., 2004; Nikzad et al., 1999, 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; Kuschnerus et al., 1998).

The final term, internal quantum efficiency (QE<sub>internal</sub>), is the collection efficiency of the photoelectrons. For  $P=1$  and for a QY of unity (i.e., one photoelectron is produced for every absorbed photon), 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 the reflection from the bare silicon surface. The internal QE can be affected by trapping or recombination of the photoelectrons in the vicinity of the illuminated silicon surface. These effects depend strongly on processes used for surface passivation and antireflection (AR) coating of back-illuminated detectors.

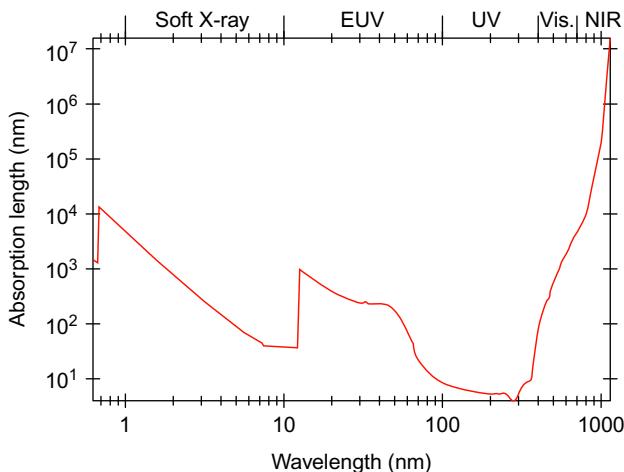
Fig. 15.2 shows a calculation of the above effects for a back-illuminated silicon detector with a 20-Å silicon oxide. The solid line is the silicon transmittance, including reflection and absorption in the oxide, and corresponds to the reflection-limited QE (or 100% internal QE) for a QY of unity. The dashed line is the silicon transmittance (with oxide) including the effect of QY. 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 it is crucial to measure the QY in order to determine external QE, defined as the ratio of detected to incident photons (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 range. As we will discuss later in this chapter, silicon detectors also can be modified for high sensitivity in the UV, far UV, and even soft X-rays through back illumination.

At the other end of the spectrum, the spectral response of silicon imagers typically decreases rapidly in the near-infrared (NIR) range, dropping to a few percent around 1000 nm. This is again a matter of photon absorption, as these longer wavelength photons are only weakly absorbed in silicon because of its indirect bandgap. Fig. 15.3 shows the absorption length of photons in silicon as a function of wavelength.



**Fig. 15.2** 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).



**Fig. 15.3** Photon absorption length in silicon as a function of photon wavelength. In the entire UV and blue section of spectrum, the absorption length is below 10 nm whereas at longer wavelengths, the absorption length is 10's or 100's of micron into silicon.

To allow efficient detection of  $\sim 1 \mu\text{m}$  light, a thickness of at least  $100 \mu\text{m}$  is required. Holland and team have developed fully depleted charged-coupled devices (CCDs) using ultrahigh purity silicon several hundreds of microns thick (Holland et al., 2003). Note that silicon's indirect bandgap, and the consequent long absorption length of visible and NIR 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 ultrahigh purity 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), that is, delta doping, these devices have exhibited high QE over a wide spectral range (near UV to near IR) (Blacksberg et al., 2005, 2008). An added advantage of the p-channel design used by Holland and team is higher radiation tolerance, which is attributed to the absence of phosphorus-vacancy traps in irradiated p-channel CCDs (Bebek, 2002).

## 15.3 Scientific imaging detectors

Detector requirements in scientific applications are defined by the end user (observer) based on the scientific objectives, instrument design, and environmental considerations. These requirements push the limits of detector technologies in noise, photon detection probability, fill factor, QE, and tolerance to radiation. Sensitivity over a broad spectral range is required in some applications, while in other applications high sensitivity is required in a narrow spectral band together with exceptionally low out-of-band sensitivity.

Although no single detector can achieve a perfect combination of the attributes described above, 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.

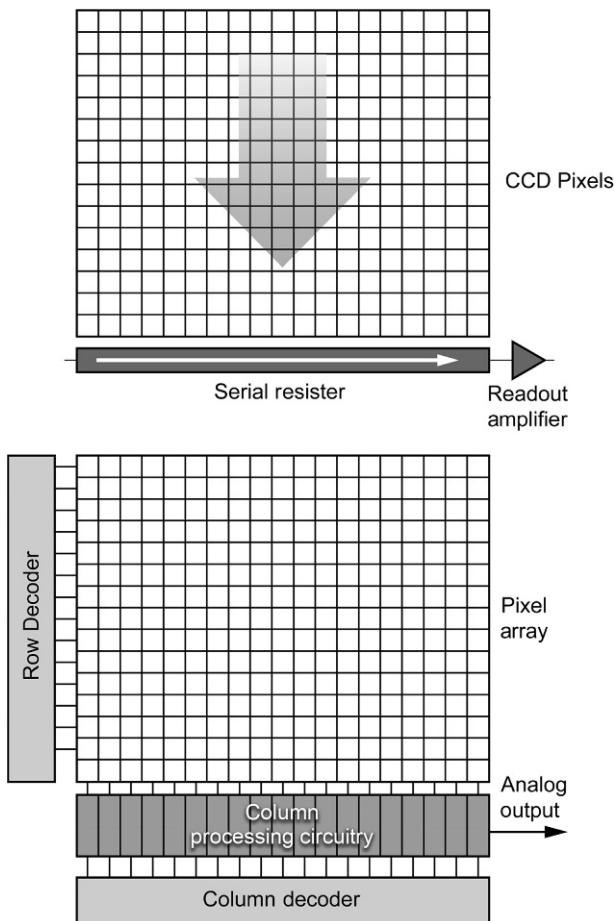
To achieve the highest performance in silicon imaging arrays, these detectors need to be back illuminated. Silicon imaging detectors are fabricated with VLSI-fabricated pixels on one side of the silicon wafer. This VLSI-fabricated pixel side is conventionally called the front surface and when illuminated at this surface the device is referred to as front illuminated. Materials used for pixel fabrication cause scattering and absorption of incident light and therefore, affect the detector QE and spectral sensitivity. By flipping the detector over and illuminating the “back 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. For this reason, back-illuminated detectors have become the norm for high-performance silicon detectors, and they are increasingly used in commercial applications as back illumination is essential for three-dimensional (3D) integration.

Back illumination allows expansion of spectral range sensitivity beyond visible and into UV, 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 for indirect detection of extrasolar planets. Over a decade later, it has become clear that for other silicon imagers, for example, complementary metal-oxide-semiconductor (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 electronic circuitry, is packed into each pixel. This leaves only a small area of the pixel devoted for photon sensing. By using back illumination, the part of the pixel's area that is allocated 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.

## 15.4 Readout structures

As described above, absorption of photons in a semiconductor leads to the 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 describe the fundamentals of design and operation of CCDs and CMOS arrays (e.g., 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. [Fig. 15.4](#) 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 detectors 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.



**Fig. 15.4** The conceptual structure of CCD (*top*) with serial readout and single output amplifier and CMOS imagers with parallel readout scheme (*below*).

## 15.5 Photon counting detectors

In many scientific observations such as biomedical application, UV astrophysics, and exo-solar planet studies using a coronagraph or starshade (Nikzad, 2017; Jewell et al., 2018; Martin et al., 2017) photon counting is essential. In order to count photons, signal amplification or gain is usually employed to increase the signal-to-noise ratio of the detector. Detectors that possess gain are divided into two general categories: image tube-based technologies and all-solid-state imagers with avalanche gain. More recently, a new class of silicon detectors, the Quanta Image Sensor (QIS), has achieved single-photon sensitivity without gain (Ma and Fossum, 2015).

### 15.5.1 Electron-bombarded arrays

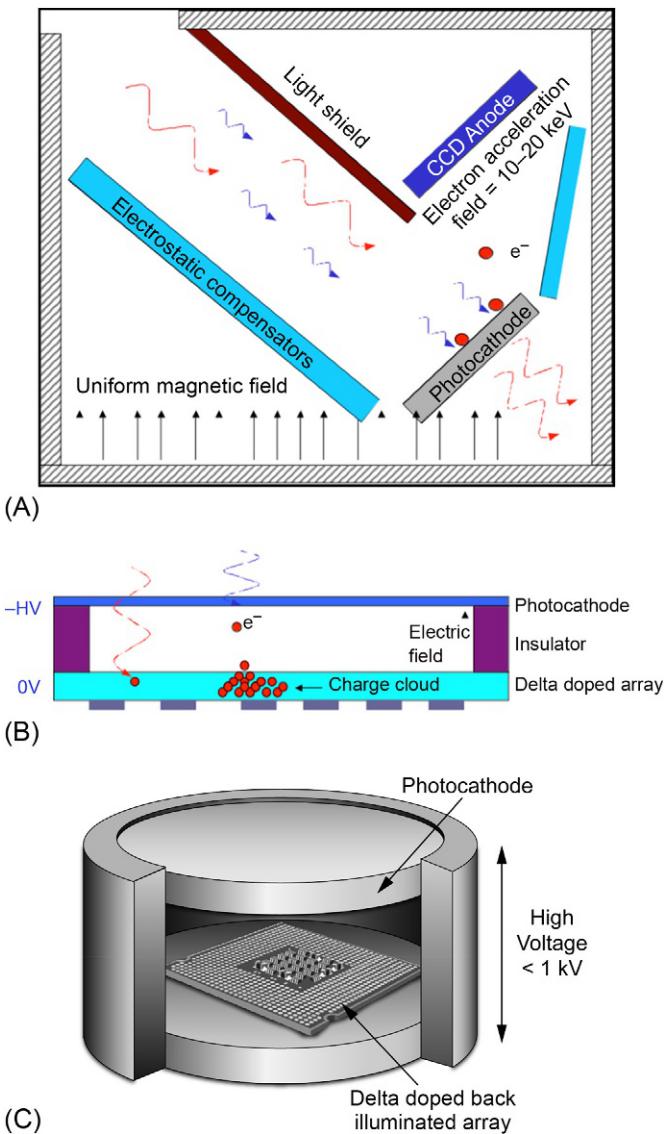
In image-tube-based detectors such as photomultiplier tubes, microchannel plates (MCPs), and electron-bombarded (EB) arrays, photons are absorbed in a photocathode; photoelectrons emitted from the photocathode are then accelerated into vacuum under an electric field of typically several thousand volts. These photoelectrons undergo amplification by multiplication in the image tube and are then detected by various schemes that are described below. In almost all of these schemes, there is still a role for a silicon digital imagers, as the final signal is often measured either by directly detecting electrons using a silicon detector array as in EB arrays, or by using a silicon detector array to detect photons that have been converted from electrons in a phosphor screen—the so-called image-intensified CCD or CMOS. We focus our discussion in this section only to EB arrays such as electron-bombarded charge-coupled device (EBCCD) and EBCMOS.

In EB array structure, photoelectrons are accelerated to energies of several keV in order to allow detection of electrons by the back-illuminated CCD or CMOS detector. The acceleration voltage required in EB design depends on the sensitivity of the back-illuminated array to low-energy electrons. A magnetic field is used to bend and focus the accelerated photoelectrons onto a non-line-of-sight back-illuminated CCD or back-illuminated CMOS imaging sensor (Fig. 15.5A). The non-line-of-sight positioning of the back-illuminated array is to prevent the detection of those unwanted photons that pass through the photocathode. Compact EBCCD and EBCMOS designs without a magnet in a line of sight configuration using low acceleration voltage have been devised by using a delta-doped CCD. Rejecting out-of-band (e.g., visible) photons in these compact designs can be obtained using thresholding for the background signal produced by visible photons (Morrissey et al., 2006). An example of an EB array in proximity focus or line of site configuration is shown in Fig. 15.5B and C.

Photocathodes determine the efficiency of the entire detector as well as the reliability and manufacturability of the device. Photocathodes typically require the deposition of an ultrathin layer of cesium on the surface in order to allow photoelectrons to escape from the photocathode surface into vacuum with relatively high efficiency. This process of cesiation takes advantage of Cesium's low work function to create a surface dipole, which lowers the vacuum energy level below the edge of the conduction band [also known as negative electron affinity (NEA)]. Because cesiated surfaces are highly reactive, the photocathode must be protected from air exposure throughout the lifetime of the detector. Using Cs, therefore, adds complexity and reliability issues to this class of devices. In order to improve stability and reliability, it is therefore desirable to achieve NEA without cesiation. Work is underway to use surface band-structure engineering in gallium nitride and its alloys (GaN and  $\text{Al}_{1-x}\text{Ga}_x\text{N}$ ) in order to fabricate stable, high-efficiency photocathodes for UV applications (Tripathi et al., 2010a, 2010b; Marini et al., 2018).

### 15.5.2 Solid-state detectors with gain

To increase the S/N ratio in a semiconductor imager, efforts are made to both decrease the noise and amplify the signal. Gain on the order of several tens of thousands can be achieved in semiconductor detectors by inducing avalanche multiplication. Avalanche



**Fig. 15.5** (A) Schematic representation 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. 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.

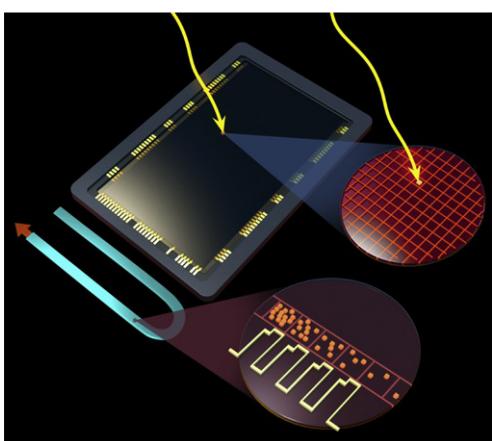
photodiode (APD) arrays are variations of PIN diode arrays that contain a region characterized by a sufficiently high electric field to induce electron multiplication. The first APDs were fabricated in silicon; but because the major motivation for requiring the high S/N ratio also required sensitivity in 1  $\mu\text{m}$  and larger wavelengths, silicon was abandoned in favor of semiconductors with a smaller bandgap energy (such as InGaAs and mercury cadmium telluride). Since that time, silicon APD popularity has been renewed, partially due to the advances in detector design and fabrication technologies and partially due to the need for high S/N ratio imagers in the visible and UV spectral range. Here, we briefly review the broad classifications of silicon imagers that possess gain along with their specific advantages and applications.

### 15.5.2.1 CCDs as integrating detectors with gain: electron multiplying CCDs

In the last two decades, a breakthrough in CCD detectors has made it possible to achieve photon counting performance in electron multiplying CCDs (EMCCDs) (Jerram et al., 2001; Hynecik, 2001). Fig. 15.6 schematically shows the EMCCD structure. Gain is achieved in EMCCDs by adding an additional gain register to otherwise conventional CCDs. Within this gain register, high voltage (in the range of 20–50 V) is applied at each stage of electron transfer. Whereas the avalanche gain is small for a single stage, hundreds of such transfers can achieve cumulative gains of 1000 or more while maintaining low noise at the readout amplifier. The resulting improvement in the S/N ratio enables photon counting sensitivity and unique capabilities in the detection of faint signals.

### 15.5.2.2 Avalanche photodiodes

APD operations can be divided into several regimes depending on the magnitude of the reverse bias voltage. At low applied voltages, a small photoresponse will be detected. As the reverse bias voltage is increased, an output current is detected that



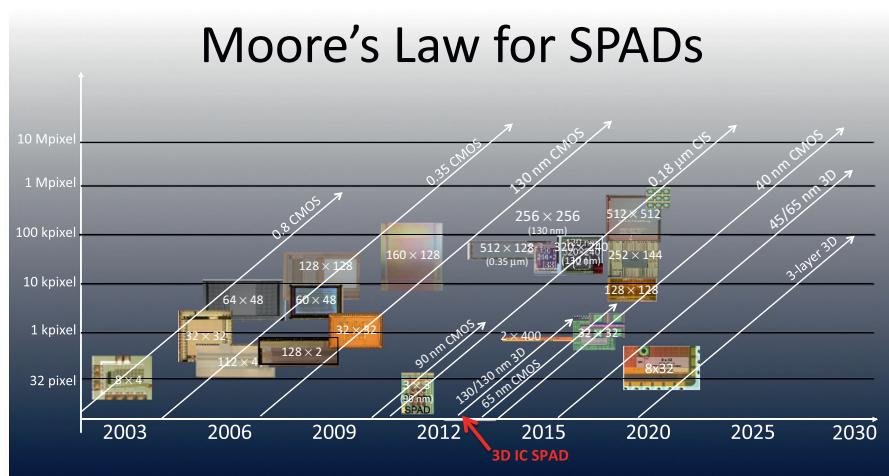
**Fig. 15.6** Electron multiplying charge coupled device (EMCCD) concept shown here in artist conception. The multiplication or gain register added as an additional serial register, creates impact ionization probability in each stage of the transfer producing gain at the end prior to readout through the output amplifier.

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 regime, 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 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 due to high gain fluctuations with temperature (Farr, 2009).

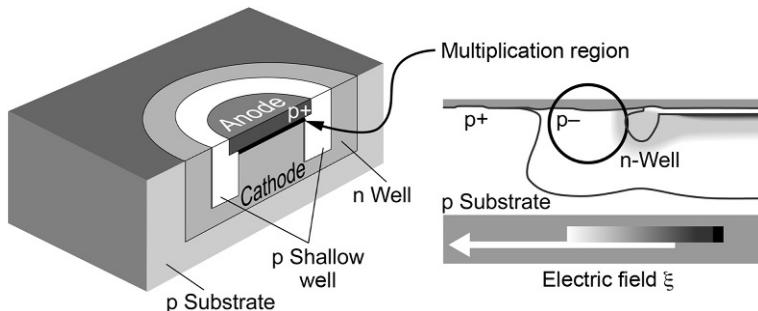
### 15.5.2.3 Single-photon avalanche diodes

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 (SPAD) has progressed as shown in Fig. 15.7 (Charbon, E., 2019. Private Communication).

As described above, the reverse-biased pn junction operating in the Geiger mode produces a pulse in response to single photons. Fig. 15.8 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 timing resolutions of 100 ps per pixel achieved on relatively large arrays, SPADs are especially useful for time-resolved measurements (Charbon, 2014).



**Fig. 15.7** Evolution of SPADs.  
Courtesy of Professor Edoardo Charbon.



**Fig. 15.8** Schematic representation 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 pixel area being devoted to circuitry. As with all other silicon detectors, back-illuminated SPADs require a surface processing method such as delta doping (described below) for passivation of defects at the Si-SiO<sub>2</sub> interface. This could significantly improve QE while potentially improving the noise performance. Dark counts may be further improved by optimizing doping in the multiplication regions to reduce tunneling and improving the guard rings used in SPADs for mitigation of premature edge breakdown (Charbon, 2011).

## 15.6 High-performance imaging through back illumination

In the context of our discussion here, high-performance imaging detectors need to meet stringent requirements on spectral range [e.g., UV, visible, NIR in silicon], sensitivity, fill factor, stability, uniformity, noise, and 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). By illuminating from the back surface, one avoids the absorbing layers of metals and polysilicon that are necessary for the pixel structure. This allows more efficient photon absorption and subsequently higher QE and potentially allows sensitivity to shorter wavelength photons that are absorbed within a few nanometers of the silicon surface. Furthermore, as stated before, CMOS imaging arrays and SPADs suffer from low fill factor, because a significant fraction of the pixel area is devoted to electronic signal processing and readout. By illuminating from the back surface, it is possible to approach 100% fill factor and consequently higher photon detection efficiency.

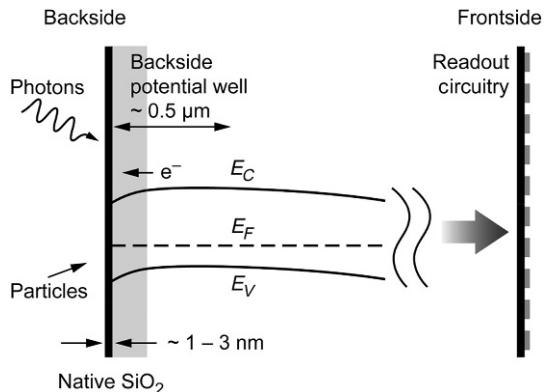
### 15.6.1 Back-illuminated devices

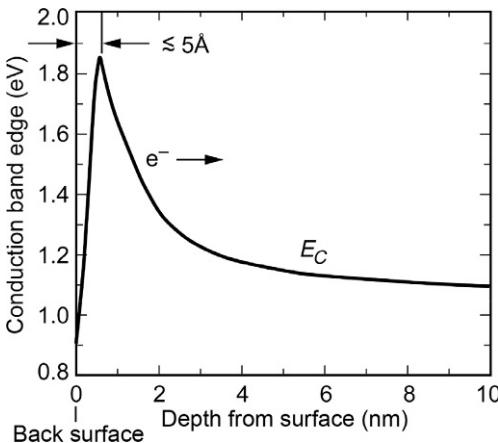
For back illumination operation, the silicon substrate must be removed in order to expose the epitaxial layer to light. Thinning or substrate removal produces a fresh silicon surface with many dangling bonds that react with air to form a thin silicon oxide. This new surface and its native oxide have an interesting electronic structure that is determined by the quality of the oxide as well as the silicon crystal orientation. The silicon-silicon oxide interface has a high density of states in the silicon bandgap. These states trap majority carriers, thereby depleting the surface and creating a backside potential well that degrades QE by trapping photogenerated minority carriers. The surface depletion region forms a backside potential well, which can trap photoelectrons. Various techniques for passivating the surface and removing this trapping potential are discussed here, including band-structure engineering using MBE for an ideal solution to this problem.

It is a well-documented phenomenon that surface states trap charge that unfavorably affects the near-surface band structure of silicon devices (Bardeen, 1947; Grunthaner et al., 1988; Hoenk et al., 1992; Nikzad et al., 1994a, 1994b, 2012). In back-illuminated detectors, surface charging causes unfavorable band bending that can lead to low, unstable QE, and high dark current, as schematically illustrated in Fig. 15.9. Since the invention of back-illuminated CCDs, several solutions to this problem have been devised, included the UV flood, the Schottky barrier “flash gate” using platinum or iridium, chemisorption charging, ion implantation, and MBE-processed band-structure engineering (Fig. 15.10). These back-surface passivation processes fit in three general categories:

1. Chemical methods of back-surface charging, which offset interface trapped charge with charge outside the silicon. Examples include UV flood, platinum flash gate, and chemisorption charging (Williams et al., 1997; Lesser and Iyer, 1998).
2. Surface doping methods, which offset interface-trapped charge with ionized dopant atoms near the silicon-silicon oxide interface. Examples include ion implantation (Janesick et al., 1989) and band-structure engineering using MBE (Hoenk et al., 1992). Surface doping using

**Fig. 15.9** Schematic representation illustration of cross section of silicon imager. The detail of the frontside readout circuitry is not depicted to illustrate the independence of the back-surface band bending and charge trapping issue from the readout method.





**Fig. 15.10** Schematic representation of the edge of the conduction band of a silicon imager after modification of the back-surface band structure 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.

chemical vapor deposition of pure boron has shown some success in producing front-illuminated detectors with high sensitivity in the UV (Nanver et al., 2012).

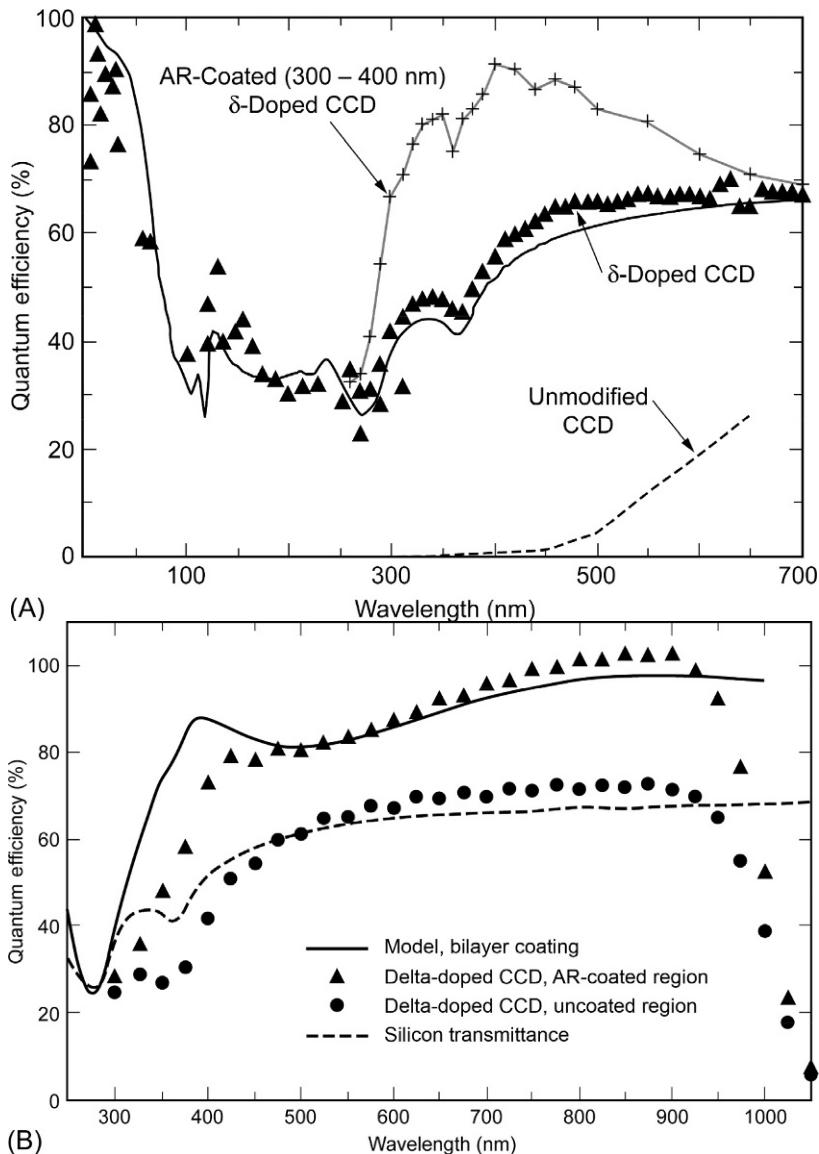
3. Field-effect techniques that offset interface-trapped charge by applying an electrical bias to a thin metal electrode deposited on the detector surface. Examples include the biased platinum flash gate developed at JPL for WF/PC 2 CCDs (Trauger, 1990 and Janesick, 2001).

In all these cases, charge is placed near the surface in order to favorably affect the near-surface electric field. In the first category, charge is chemically induced by using electronegative material on the surface. In the second category, charge is 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 a high temperature annealing step. Finally, the third category applies a bias on a metal contact to directly manipulate the near-surface electric field using the applied voltage.

Each of these techniques provides improvement with some degree of success. Surface doping techniques have the advantage of improved stability, as charge is directly incorporated in the crystal lattice to counteract the effect of the charge trapped in the interface states of silicon and its native oxide.

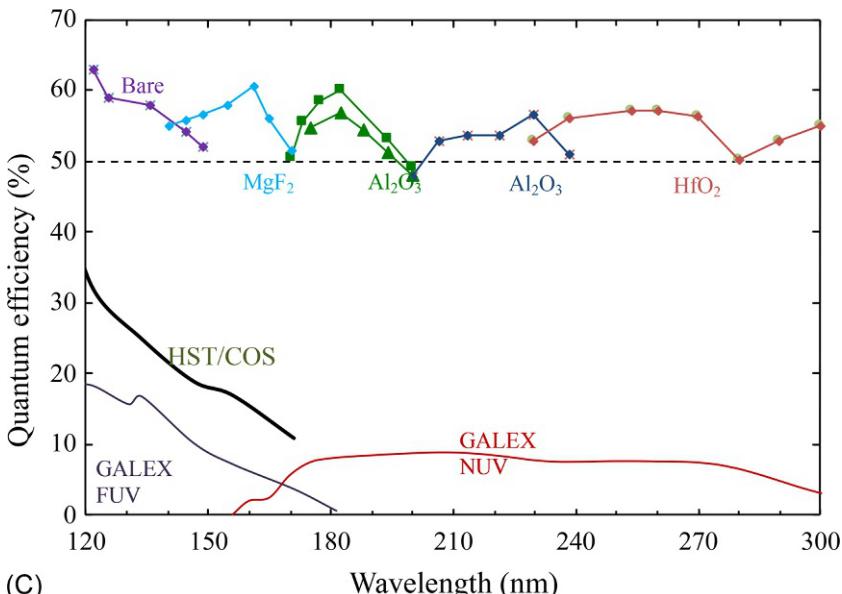
### 15.6.2 *Delta doping and superlattice doping as ideal solution for back-illuminated devices*

Delta doping is a technology invented at JPL for producing high-performance back-illuminated silicon imagers (Hoek et al., 1992). Delta-doped devices with nearly 100% internal QE (Fig. 15.11A) have been demonstrated on various platforms,



**Fig. 15.11** (A) Typical 100% internal quantum efficiency of delta-doped arrays, shown here for CCDs from extreme ultraviolet (UV) to the visible part of the spectrum. The CCDs are thinned to a  $\sim 10\text{-}\mu\text{m}$  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 AR coatings. The thickness of these devices allows efficient detection of higher efficiency near 1000 nm.

(Continued)



**Fig. 15.11—cont'd** (C) Quantum efficiency in the far UV and near UV achieved with atomic layer deposition for AR coatings and molecular beam epitaxy for band structure engineering of silicon-silicon oxide interface.

After (with added HST/COS data) Nikzad, S., Hoenk, M., Greer, F., Jacquot, B., Monacos, S., Jones, T., Blacksberg, J., Hamden, E., Schiminovich, D., Martin, C., 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. *Appl. Opt.* 51, 356–369.

designs, and formats. These include conventional CCDs (Hoenk et al., 1992; Nikzad et al., 1994a, 1994b), 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). Invented in 1992 at JPL (Hoenk et al., 1992), the development of this technology was motivated by the instability and low QE observed in the back-illuminated CCDs in the Hubble Space Telescope's (HST) WF/PC 1. 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). Superlattice doping is an extension of delta doping which incorporates multiple delta layers on the backsurface offering even higher stability of the response as well as high QE. Delta doping and superlattice doping (Hoenk et al., 2014) are two-dimensional (2D)-doping techniques achieved by low-temperature MBE.

In the 2D-doping processes, the temperature never rises above 450°C and therefore, fully fabricated imaging detectors can be processed. An ultrathin, high-quality single-crystal silicon layer is grown on the thinned and atomically clean surface of a 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).

### 15.6.3 Reducing photon losses with advanced AR coatings

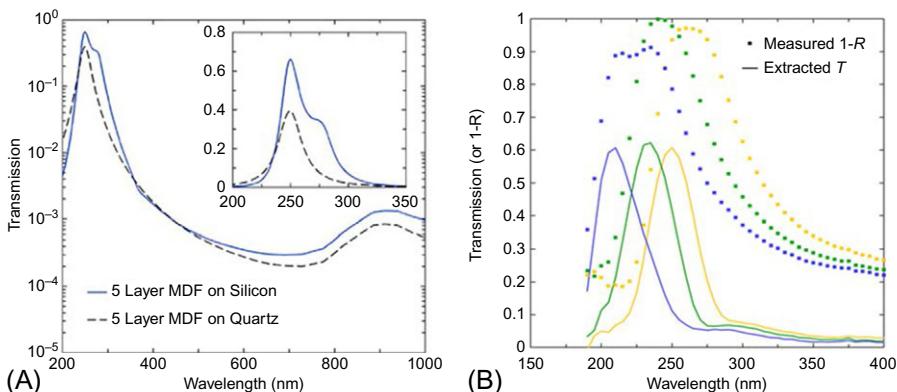
In the previous section, we dealt with collection of all the photoelectrons that are produced by the absorbed photons in the detector leading to 100% internal QE or reflection-limited response. Achieving high external QE is further possible by reducing reflection losses through depositing AR coatings. 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 dielectric materials or a combination of high and low-index dielectric materials on the 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, sharply defined single or multilayer AR coatings (Hamden et al., 2011; Nikzad et al., 2012; Nikzad, 2017). These films, in combination with delta doping and superlattice doping, have shown record high QE (>50%) in far and near UV (Fig. 15.11C).

### 15.6.4 Solar blind UV silicon detectors: A breakthrough

Silicon arrays have made great strides in UV and visible systems. They have high efficiency, are linear, uniform, can photon count, and can be fabricated with small pixels and high resolution. The only remaining area where image-tube devices historically hold advantage over silicon detectors is UV imaging and spectroscopy, where signal-to-noise considerations often require the suppression of high background signals in the visible and near infrared regions of the spectrum. In the past several years, a breakthrough in silicon detector development has made it possible to achieve high UV QE while suppressing the unwanted visible photons. In order to achieve stable UV response together with high red-leak suppression, metal dielectric filters can be directly integrated on back-illuminated and 2D-doped silicon detectors. Metal dielectric filters have long been used with transparent substrates for stand-alone filters and now that their utility has been extended to silicon substrates for use as detector-integrated filters with great success.

Stacks of Al/Al<sub>2</sub>O<sub>3</sub> as well as Al/AlF<sub>3</sub> and Al/MgF<sub>2</sub> metal dielectric filters have been implemented on APDs and EMCCDs showing excellent suppression of out of band light as UV bandpass filters. Thus, it is possible to achieve high in-band QE with out-of-band suppression on the order of 10<sup>4</sup> (Hennessy et al., 2015, 2017a, b; Hennessy and Nikzad, 2018; Nikzad et al., 2016; Nikzad, 2017). Compared to in-line transparent filters prepared on transparent substrates, on-chip filters allow for better index matching and higher in-band throughput (Fig. 15.12).

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 UV applications. While MCP detectors work well, they have limits, and for decades scientists have attempted the development of 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 are not limited by count rates as



**Fig. 15.12** (A) Calculated performance of a five-layer MDF on silicon or quartz substrates. In band QE > 60%, and out-of-band rejection approaches  $10^4$ . (B) Prototype MDFs on Si using aluminum fluoride layers as an alternative to aluminum oxide with measured reflectance and projected transmittance based on the ellipsometric characterization (not shown) for three designs optimized for 200–250 nm.

MCP detectors are locally and globally. This is now possible with silicon detectors. Advances in silicon imaging to achieve high efficiency and photon counting such as delta-doped EMCCDs as seen in Fig. 15.11C or superlattice doping of SPADs, QIS, or other silicon imagers enable miniaturization, simplification, and reduced power consumption while improving the performance of UV/VUV instruments.

### 15.6.5 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 the applications of back-illuminated imagers in the UV to NIR spectral range. It should be noted that the band-structure engineering that removes the traps for photoelectrons, allowing higher-energy photons in UV and extreme UV 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 EB arrays with lower operating voltages.

### 15.6.6 High-throughput delta and superlattice doping of CMOS, CCD, and other silicon detectors

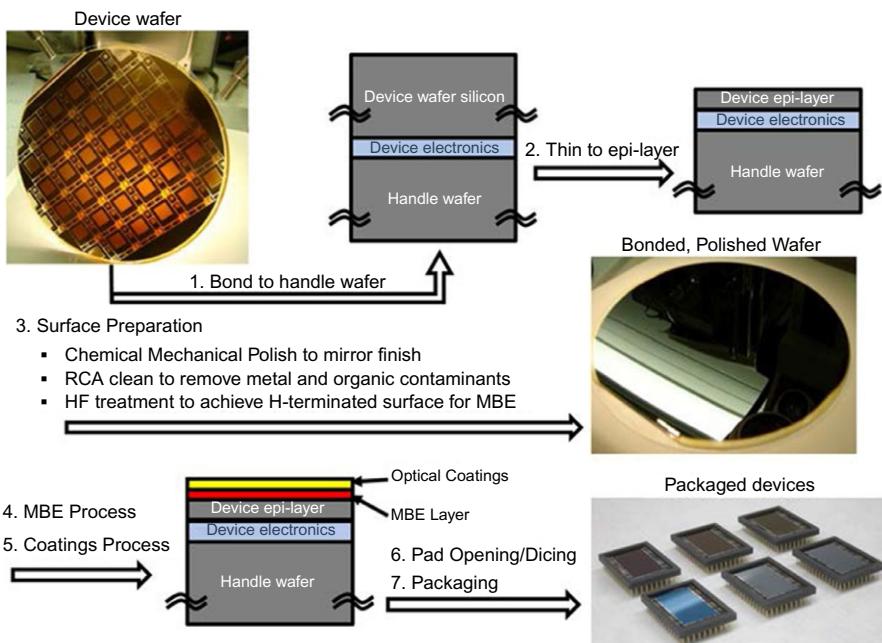
Delta and superlattice-doped devices have been demonstrated to have high efficiency, tailororable response, ultra-stable response, long-term stability, and applicability to virtually all silicon detector architectures including CMOS, CCD, APDs, and other architectures. High-throughput, high yield, industry and foundry-compatible processes have been developed to address detector production for large focal plane arrays (FPAs) in space applications, ground observatories, and for high-volume terrestrial applications with lower cost.

High-throughput infrastructure including a production-grade silicon MBE and surface preparation equipment for wafer batch processing with high yield have been developed at the JPL ([Nikzad, 2017](#)). The MBE machine employs a computer-controlled automated sample transfer system with a cluster tool that moves wafer platens between chambers. The preparation chamber and the storage chamber are differentially pumped to allow rapid processing. Each 10-in. platen can be configured for single wafers up to 8 in. in diameter or for multiple smaller wafers. The entire system is under computer control, enabling the development of automated high-throughput, multi-wafer processing. For AR coatings or detector-integrated filter deposition, JPL's facilities include ALD systems. ALD is arbitrarily scalable to larger wafers and compatible with batch processing. Multiple precursor compounds and multiple ports for reactive gases as well as individual and modular growth chambers allow for high-quality and high-throughput processing. [Fig. 15.13](#) shows the overview of post-fabrication processing steps of back illumination using delta and superlattice doping, in which fully fabricated wafers of silicon detector arrays of virtually any structure are transformed into high-efficiency detectors with tailored response in UV and visible band passes.

## 15.7 Planetary and astronomy applications

Ever since Galileo used his innovative telescope design to discover 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 benefited from 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 Lander (MSL).

In 1976, only 7 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 HST used CCDs for the first time in space. At  $800 \times 800$  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

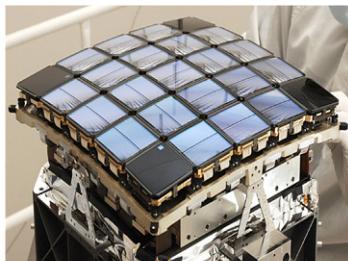


**Fig. 15.13** JPL's end-to-end post-fabrication process flow to produce high-performance UV/optical/NIR sensitized devices with 100% internal QE and tailorble external QE. The infrastructure for high volume, wafer-level processing allows for high-throughput and cost-effective processes.

wide field cameras (WFCs) 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 to be dark and revealing a seemingly infinite field of galaxies. These HST deep field images have changed the way we perceive our universe.

Cassini spacecraft used a CCD with a  $1024 \times 1024$  format that was a variation on the design of the HST-WF/PC2 CCD. Both of these CCDs were operated as front illuminated; however, because of the need for UV sensitivity, the HST device was coated by a phosphor that would absorb the UV and reemit in the green ( $\sim 500$  nm) wavelengths where CCD sensitivity was high. While WF/PC II camera produced beautiful images and enabled discovery, its CCD had only 10%–12% efficiency in the 121–310 nm spectral range. The subsequent HST cameras have employed back-illuminated CCDs in the advanced camera and the WFC3.

Beyond HST, UV and optical telescopes for future missions are planned with larger apertures and with larger FPAs that are populated with large-area silicon imaging arrays. For example, NASA has funded four large telescope mission studies in preparation for the 2020 decadal survey, in which the National Research Council will prioritize future NASA astrophysics missions. Two of the four mission concepts, Large



Kepler FPA. Photo courtesy of NASA

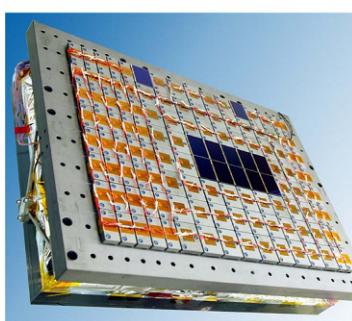


Photo credit: ESA



Photo credit: SDSS official website

**Fig. 15.14** 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  $1024 \times 2048$  pixels), Sloan Digital Sky Survey or SDSS (30 CCDs of  $2048 \times 2048$  pixel format), and finally Giai with 106 CCDs of  $4500 \times 1966$  pixels.

UV Optical Infrared survey mission (LUVOIR) and Habitable Exoplanet mission (HabEx) are being designed with telescopes with 4–15 m apertures. The design teams for these mission concepts are planning for large FPAs to be populated with high-performance UV or UV/optical detector arrays. Fig. 15.14 shows a few examples of large FPAs already in space or in operation in ground observatory. The demand for large FPAs 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). Nearly all planetary science studies employ UV instruments for studying the surface and atmosphere of planetary object.

In the last several years, delta-doped silicon arrays of various formats, designs, and architecture have been deployed in sounding rocket (Nikzad, 2017) and high-altitude balloons (Hamden et al., 2016; Kyne et al., 2016; Nikzad, 2017) and they are planned for being used and launched in CubeSats, SmallSats, as well as future large aperture telescope flagship missions (Nikzad, 2017; Martin et al., 2017).

## 15.8 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, that is, semiconductor inspection.

In the introduction, we enumerated a few of the most commonly used machine vision applications, including surveillance cameras, security cameras, and cameras in automobiles. State-of-the-art wafer and reticle inspection systems use deep ultra-violet (DUV) lasers at 263 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. Back-illuminated CMOS imaging arrays meet the requirements for resolution, frame rate, and noise, however, 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. Fig. 15.15 shows the camera. The stability of superlattice-doped detectors' response to 193 and 260 nm photons has been measured under continuous exposure over a period of several months, exhibiting unprecedented durability with no measurable changes in sensitivity or performance (Nikzad et al., 2012; Hoenk et al., 2013).

## 15.9 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 e.g., the introduction in Llosa, 2005). Leveraging from the advances made in silicon imaging for astronomy, highly efficient detectors have been implemented in medical imaging applications, making it possible to reduce exposure of patients to radiation, reduce diagnosis turnaround time, and produce more compact medical radiation devices.

More recently optical and UV imaging is being explored for intraoperative tumor delineation or postoperative diagnosis. The use of semiconductor imaging could allow

**Fig. 15.15** A superlattice-doped CMOS camera optimized in deep UV (260 nm) as a prototype for wafer and reticle inspection applications.



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 are 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 technologies, silicon detectors can offer a powerful multimodal imaging tool for medical applications ([Nikzad, 2017](#)).

## References

- Bai, Y., Montroy, J., Blackwell, J., Farris, M., Kozlowski, L., Vural, K., 2000. Development of hybrid CMOS visible focal plane arrays at Rockwell. Proc. SPIE 4028, 174. Infrared Detectors and Focal Plane Arrays VI.
- 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. Nucl. Sci. 49, 1221.

- Blacksberg, J., Hoenk, M.E., Elliott, T.S., Holland, S.E., Nikzad, S., 2005. Enhanced quantum efficiency of high purity silicon imaging detectors by ultra-low temperature surface modification using Sb-doping. *Appl. Phys. Lett.* 85, 254101.
- Blacksberg, J., Nikzad, S., Hoenk, M.E., Holland, S.E., Kolbe, W., 2008. Near-100% quantum efficiency of delta doped large-format UV-NIR silicon imagers. *IEEE Trans. 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-performance single-photon detection. *Optics Photon. News* 22 (5), 14–15.
- Charbon, E., 2014. Single-photon imaging in complementary metal oxide semiconductor processes. *Phil. Trans. R. Soc. A* 372, 20130100. <https://doi.org/10.1098/rsta.2013.0100>.
- Farr, W.H., 2009. Negative avalanche feedback detectors for photon-counting optical communications. *Proc. SPIE* 7199, 71990Q-1.
- Funsten, H.O., Ritzau, S.M., Harper, R.W., 2004. Fundamental limits to detection of low-energy ions using silicon solid-state detectors. *Appl. Phys. Lett.* 84, 3552.
- Greer, F., Hamden, E., Jacquot, B.C., Hoenk, M.E., Jones, T., Dickie, M., Monacos, S.P., Nikzad, S., 2013. Atomically precise surface engineering of silicon CCDs for enhanced UV quantum efficiency. *J. Vac. Sci. Technol. A* 31, 01A103.
- Grunthaner, P.J., Grunthaner, F.J., Fathauer, R.W., Lin, T.L., Schowengerdt, F.D., Pate, B., Mazur, J.H., Bean, J.C., Schowalter, L.J., 1988. *Silicon Molecular Beam Epitaxy*. The Electrochemical Society, New Jersey, p. 375.
- Hamden, E., Greer, F., Hoenk, M.E., Blacksberg, J., Jones, T.J., Dickie, M., Jacquot, B., Monacos, S., Martin, C., Morrissey, P., Schiminovich, D., Nikzad, S., 2011. Antireflection coatings designs for use in UV imagers. *Appl. Opt.* 50, 4180.
- Hamden, E.T., Jewell, A.D., Shapiro, C.A., Cheng, S.R., Goodsall, T.M., Hennessy, J.J., Hoenk, M.E., Jones, T.J., Gordon, S., Ong, H.R., Schiminovich, D., Martin, D.C., Nikzad, S., 2016. Charge-coupled devices detectors with high quantum efficiency at UV wavelengths. *J. Astron. Telesc. Instrum. Syst.* 2(3), 036003, <https://doi.org/10.1117/1.JATIS.2.3.036003>.
- Hennessy, J., Nikzad, S., 2018. Atomic layer deposition of lithium fluoride optical coatings for the ultraviolet. *Inorganics* 6 (2), 46.
- Hennessy, J., Jewell, A.D., Hoenk, M.E., Nikzad, S., 2015. Metal-dielectric filters for solar-blind silicon ultraviolet detectors. *Appl. Opt.* 54, 3507–3512.
- Hennessy, J., Jewell, A.D., Nikzad, S., 2017a. ALD metal fluorides for ultraviolet filter and reflective coating applications. *ECS Trans.* 80 (3), 107–118.
- Hennessy, J., Moore, C.S., Balasubramanian, K., Jewell, A.D., France, K., Nikzad, S., 2017b. Enhanced atomic layer etching of native aluminum oxide for ultraviolet optical applications. *J. Vac. Sci. Technol. A* 35, 041512.
- Hoenk, M.E., Carver, A., Jones, T., Dickie, M., Cheng, P., Greer, F., Nikzad, S., 2013. DUV stability of superlattice doped CMOS detector arrays. In: 2013 International Image Sensor Workshop, 12–16 June, Snowbird, Utah.
- Hoenk, M.E., Grunthaner, P.J., Grunthaner, F.J., Fattahi, M., Tseng, H.-F., 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. *Appl. Phys. Lett.* 61, 1084.
- Hoenk, M.E., Jones, T.J., Dickie, M.R., Greer, F., Cunningham, T.J., Blazejewski, D.R., Nikzad, S., 2009. Delta-doped back-illuminated CMOS imaging arrays: progress and prospects. *Proc. SPIE* 74190, 74115.

- Hoenk, M.E., et al., 2014. Superlattice-doped silicon detectors: progress and prospects. Proc. SPIE, 9154, 915413.
- Holland, S.E., Groom, D.E., Palaio, N.P., Stover, R.J., Wei, M., 2003. Fully depleted, back-illuminated charge-coupled devices fabricated on high-resistivity silicon. IEEE Trans. Electron Devices 50, 225.
- Hynecek, J., 2001. Impactron—a new solid state image intensifier. IEEE Trans. Electron Devices 48 (10), 2238–2241.
- Janesick, J.R., 2001. Scientific Charge Coupled Devices. SPIE Press, Bellingham, Washington.
- Janesick, J., Elliott, T., Pool, F., 1989. Radiation damage in scientific charge-coupled devices. IEEE Trans. Nucl. Sci. 36 (1), 572–578.
- Janesick, J., Putman, G., 2003. Developments and applications of high-performance CCD and CMOS imaging arrays. Annu. Rev. Nucl. Part. Sci. 53, 263–300.
- Jerram, P., Pool, P., Bell, R., Burt, D., Bowring, S., Spencer, S., 2001. LLLCCD—low light level imaging without the need for an intensifier. Proc. SPIE 4306, 178.
- Jewell, A., Hennessy, J., Jones, T.J., Cheng, S.R., Carver, A., Ardilla, D., Shkolnik, E., Hoenk, M., Nikzad, S., 2018. Ultraviolet detectors for astrophysics missions: a case study with the star-planet activity research cubesat (SPARC). Proc. SPIE 10709, 107090C.
- Kuschnerus, P., Rabus, H., Richter, M., Scholze, F., Werner, L., Ulm, G., 1998. Characterization of photodiodes as transfer detector standards in the 120 nm to 600 nm spectral range. Metrologia 35, 355–362.
- Kyne, G., Hamden, E.T., Lingner, N., Morrissey, P., Nikzad, S., Martin, D.C., 2016. The faint intergalactic-medium red-shifted emission balloon: future UV observations with EMCCDs. Proc. SPIE 9915, <https://doi.org/10.1117/12.2232879>.
- Lesser, M., Iyer, V., 1998. Enhancing back illuminated performance of astronomical CCDs. Proc. SPIE 3355, 446.
- Llosa, 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).
- Ma, J., Fossum, E.R., Sept. 2015. Quanta image sensor jöt with sub 0.3e- r.m.s. read noise and photon counting capability. IEEE Electron Device Lett. 36 (9).
- Mandai, S., Charbon, E., 2012. A wide spectral range single-photon avalanche diode fabricated in an advanced 180nm CMOS technology. Opt. Express 20, 5849.
- Marini, J., Mahabob, I., Rocco, E., Bell, L.D., Shahedipour-Sandvik, F., 2018. Polarization engineered N-polar Cs-free GaN photocathodes. J. Appl. Phys. 124, 113101.
- Martin, S., Rud, M., Scowen, P., Stern, D., Nissen, J., Krist, J., 2017. HabEx space telescope optical system. Proc. SPIE 10398, 14.
- Morrissey, P., Kaye, S., Martin, C., Sheung, J., Nikzad, S., Jones, T.J., Blacksberg, J., Hoenk, M.E., Bell, L.D., 2006. A novel low-voltage electron-bombarded CCD readout. Proc. SPIE 6266, 626610, <https://doi.org/10.1117/12.673208>.
- Nanver, L.K., Sammak, A., Sakic, A., Mohammadi, V., Derakhshandeh, J., Mok, K.R.C., Qi, L., Golshani, N., Scholtes, T.M.L., de Boer, W.B., 2012. Applications of PureB and PureGaB ultrashallow junction technologies. In: 2012 IEEE 11th International Conference on Solid-State and Integrated Circuit Technology (ICSICT), pp. 1–4.
- Niclass, C., Favi, C., Kluter, T., Gersbach, M., 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.
- Nikzad, S., 2017. Space technology and medicine. SPIE Prof. <https://doi.org/10.1117/2.4201704.17>.

- Nikzad, S., Croley, D., Elliott, S., Cunningham, T.J., Proniewicz, W.K., Murphy, G.B., Jones, T.J., 1999. Single-event keV proton detection using a delta-doped charge-coupled device. *Appl. Phys. Lett.* 75 (17), 2686–2688.
- Nikzad, S., Cunningham, T.J., Hoenk, M.E., Ruiz, R.P., Soules, D.M., 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. *Appl. Phys. Lett.* 89, 182114.
- Nikzad, S., Hoenk, M.E., Carver, A.G., Jones, T.J., Greer, F., Hamden, E., Goodsall, T., 2013. High throughput, high yield fabrication of high quantum efficiency back-illuminated photon counting, far UV, UV, and visible detector arrays. In: Proc. of the 2013 International Image Sensor Workshop. vol. 1, p. 37.
- Nikzad, S., Hoenk, M., Greer, F., Jacquot, B., Monacos, S., Jones, T., Blacksberg, J., Hamden, E., Schiminovich, D., Martin, C., 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. *Appl. Opt.* 51, 356–369.
- Nikzad, S., Hoenk, M.E., Grunthaner, P.J., Terhune, R.W., Grunthaner, F.J., Winzenread, R., Fattah, M., Tseng, H.-F., 1994a. Delta-doped CCDs: high QE with long-term stability at UV and visible wavelengths. *Proc. SPIE* 2198, 907.
- Nikzad, S., Hoenk, M.E., Grunthaner, P.J., Terhune, R.W., Winzenread, R., Fattah, M., Tseng, H.-F., Grunthaner, F.J., 1994b. Delta-doped CCDs as stable, high sensitivity, high resolution UV imaging arrays. *Proc. SPIE* 2217, 355.
- Nikzad, S., Hoenk, M.E., Jewell, A.D., Hennessy, J.J., Carver, A.G., Jones, T.J., Goodsall, T.M., Hamden, E.T., Suvarna, P., Bulmer, J., Shahedipour-Sandvik, F., Charbon, E., Padmanabhan, P., Hancock, B., Bell, L.D., 2016. Single photon counting UV solar-blind detectors using silicon and III-nitride materials. *Sensors* 16, 927.
- Nikzad, S., Jones, T., Elliot, S.T., Cunningham, T., Deelman, P., Walker, A., Oluseyi, H., 2000. Ultrastable and uniform EUV and UV detectors. *Proc. SPIE* 4139, 250. Instrumentation for UV/EUV Astronomy and Solar Missions.
- Nikzad, S., Yu, Q., Smith, A.L., Elliott, S.T., Jones, T.J., Tombrello, T.A., 1998. Direct detection and imaging of low-energy electrons with delta-doped charge-coupled devices. *Appl. Phys. Lett.* 241.
- Scowen, P., Nikzad, S., Hoenk, M.E., et al., 2010. Large focal plane arrays for future missions. In: Astro2010: The Astronomy and Astrophysics Decadal Survey, Technology Development Papers.
- Theuwissen, A.J.P., 1995. Solid-State Imaging With Charge-Coupled Devices. Kluwer Academic Publishers, Dordrecht, Boston, London.
- Trauger, J.T., 1990. Sensors for the Hubble Space Telescope wide field and planetary cameras (1 and 2). In: CCDs in Astronomy, Proceedings of Astronomical Society of the Pacific, pp. 217–230.
- Tripathi, N., Bell, L.D., Nikzad, S., Shahedipour-Sandvik, F., 2010a. Effect of n+GaN cap thickness on Cs-free GaN photocathode characteristics. *Appl. Phys. Lett.* 97, 052107.
- Tripathi, N., Bell, L.D., Nikzad, S., Tungare, M., Suvarna, P.H., Shahedipour Sandvik, F., 2010b. Novel Cs-free GaN photocathodes. *J. Electron. Mater.* 40, 382.
- Wilkinson, F.J., Farmer, A.J.D., Geist, J., 1982. The near ultraviolet quantum yield of silicon. *J. Appl. Phys.* 54, 1172.
- Williams, G., Janesick, J., Dosluoglu, T., Reinheimer, A., 1997. Advances in Pixelvision, Inc. Back-Illuminated CCD Imaging Technologies for Low Light Level Imaging Applications. DTIC Online.

## Further reading

- Baggett, S.M., MacKenty, J.W., Kimble, R.A., et al., 2010. WFC3 detectors: on-orbit performance. Proc. SPIE 7731, 1.
- Collins, N.R., Boehm, N., Delo, G., et al., 2009. Wide field camera 3 CCD quantum efficiency hysteresis: characterization and mitigation. Proc. SPIE 7439, 1.
- Deelman, P., Nikzad, S., Blouke, M., 2000. Delta doped CCDs with integrated UV coatings. Proc. SPIE 3965, 462.
- Grunthaner, P.J., Grunthaner, F.J., Fathauer, R.W., Lin, T.-L., Hecht, M.H., Bell, L.D., Kaiser, W.J., Schowendgerdt, F.D., Mazur, J.H., 1989. Hydrogen-terminated silicon substrates for low-temperature molecular beam epitaxy. Thin Solid Films 183, 197.
- 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 Jr., E.J., Shaya, E.J., Seidelmann, K.P., Westphal, J.A., 1996. Hubble Space Telescope WFPC2 imaging of M16: photoevaporation and emerging young stellar objects. Astron. J. 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., Jacquot, B.C., Monacos, S.P., Nikzad, S., 2010. The photometric stability of delta-doped detectors. Proc. SPIE 7742.
- Jacquot, B.C., Hoenk, M.E., Jones, T., Cunningham, T.J., Nikzad, S., 2012. Direct detection of 1000–5000 eV electrons with delta doped silicon CMOS and electron-multiplying CCD imagers. IEEE Trans. Electron Devices 59, 1988.
- Jenkins, J.M., Barouki, W.J., Dunham, E.W., McDonald, J.S., 1997. High precision photometry with back-illuminated CCDs. In: Sodarblum, D. (Ed.), Planets Beyond our Solar System and Next Generation of Space Missions. In: ASP Conference Series, vol. 119, pp. 277–280.
- Jiang, Z.M., Xiu, L.S., Jiang, X.M., Zheng, W.L., Lu, X.K., Zhu, H.J., Zhang, X.J., Wang, X., 1997. Extremely narrow Sb delta-doped epitaxial layer characterized by X-ray reflectivity. Chin. Phys. Lett. 14 (9), 686.
- Kasdin, J., 2009a. Exoplanets and disks: their formation and diversity. AIP Conf. Proc. 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](http://www.astro.princeton.edu/<dns/Theia/nas_theia_v14.pdf).
- Knezevic, T., Suligoj, T., Sakic, A., Nanver, L.K., 2012. Modeling of electrical characteristics of ultra-shallow pure amorphous boron p+n junctions. In: Proceedings of the 35th International Convention MIPRO 2012, Opatija, Croatia, 21–25 May, pp. 36–41.
- Lepri, S., Nikzad, S., Jones, T., Blacksberg, J., Zurbuchen, T.H., 2006. Response of a delta-doped charge-coupled device to low energy protons and nitrogen ions. Rev. Sci. Instrum. 77, 053301.
- Mackay, C.D., Tubbs, R.N., Bell, R., Burt, D., Jerram, P., Moody, I., 2008. Sub-electron read noise at MHz pixel rates. Proc. SPIE 4306, 289.
- Nikzad, S., Jewell, A.D., Hoenk, M.E., Jones, T.J.J., Hennessy, J.J., Goodsall, T.M., Carver, A.G., Shapiro, C., Cheng, S.R., Hamden, E.T., Kyne, G., Martin, D.C., Schiminovich, D., Scowen, P., France, K., McCandliss, S., Lupu, R.E., 2017. High-efficiency UV/optical/NIR detectors for large aperture telescopes and UV explorer

- missions: development of and field observations with delta-doped arrays. *J. Astron. Telesc. Inst. Syst.* 3, 036002, <https://doi.org/10.1117/1.JATIS.3.3.036002>.
- Nikzad, S., Yu, Q., Smith, A.L., Jones, T.J., Tombrello, T.A., Elliott, S.T., 1997. Direct detection and imaging of low-energy electrons with delta-doped charge- coupled devices. *Appl. Phys. Lett.* 73, 3417.
- Postman, M., 2009. Advanced technology large aperture space telescope (ATLAST): a technology roadmap for the next decade. In: NASA Strategic Mission Concept Study.
- Smith, D.R., Ingle, R., Holland, A., 1987. Ultraviolet and extreme ultraviolet response of charge-coupled devices. *Opt. Eng.* 26, 875.
- Smith, D.R., Ingle, R., Holland, A., 2006. Proton irradiation of EMCCDs. *IEEE Trans. Electron Devices* 53 (2), 205.
- Stahl, H.P., Stephens, K., Henrichs, T., Smart, C., Prince, F.A., 2010. Single- variable parametric cost models for space telescopes. *Opt. Eng.* 49, 073006.
- Westhoff, R.C., Burke, B.E., Clark, H.R., Loomis, A.H., Young, D.J., Gregory, J.A., Reich, R.K., 2009. Low-dark-current, back-illuminated charge-coupled- devices. *Proc. SPIE* 7249.
- Westhoff, R.C., Rose, M.K., Gregory, J.A., Berthiaume, G.D., Seely, J.F., Woods, T.N., Ucker, U., 2007. Radiation-hard, charge-coupled devices for the extreme ultraviolet variability experiment. *Proc. SPIE* 6686, 4.

This page intentionally left blank

# Index

Note: Page numbers followed by *f* indicate figures and *t* indicate tables.

## A

Abaciscus/abacus. *See* Mosaic paintings

Absorption coefficient, 28–29, 86,  
296–297  
of fluorophore, 379  
as function of energy in silicon, 293–294,  
294*f*  
linear absorption coefficient in silicon, 294,  
295*f*

Active pixel architectures  
4T active pixel, 53, 56  
photoactive pixel, 55  
3T pixel, 44, 51–52  
Active pixel sensor (APS), 124–125, 128,  
149–150, 196–198  
five-transistor global-shutter APS circuit,  
138, 139*f*  
4T buried APS, 152  
four-transistor CMOS APS, 129*f*  
frequency-domain lifetime imaging/  
sensing, 399  
limited DR, 142*f*  
operation, 142*f*  
reset signal, 150  
three-transistor CMOS APS, 126*f*

Active triangulation, 339

Acuity, 9–10

Advanced driver assistance system (ADAS)  
applications, 241, 244–245

AF-coated back-illuminated CCD,  
83–84

Airy disc, 441, 442*f*

Airy pattern, 120

ALD. *See* Atomic layer deposition (ALD)

Aliasing, 439–447

ALICE, 305, 310

Alumina ( $\text{Al}_2\text{O}_3$ ), 271, 272*t*

Aluminum nitride (AlN), 271, 272*t*

Amplitude-modulated continuous wave  
modulation (am-cw), 326–329

Analog processing element (APE), 169

Analog-to-digital converters (ADCs),  
127–128, 130, 161–165, 174–175,  
190, 200–201

ADC-noise, 328–329

array chip, 133

column parallel design, 132*f*

cyclic, 137*f*

DS ADC, 134, 137*f*

first-order sigma-delta, 137*f*

image sensor, 132*f*

pixel-level, 132–133

SAR ADC, 134, 135*f*

SS ADC, 134, 136*f*

stacked design architecture, 132*f*

topologies, 134, 137*t*

Android, 363

Antiparticle, 15, 27

Antireflective coating (ARC), 260, 261*f*, 336

BSI sensor, 110, 111*f*

layer, scaled-down FSI sensor, 101

APD. *See* Avalanche photodiode (APD)

A-Pix, 194

Apple, 363

AR, 361

ARkit API, 188

3D selfie

face ID/face unlock, 362

facial animation/animojis/morphing,

362

picture improvement (lighting/bokeh),

362

FaceTime, 185

iPhone-X Face ID, 211–212

Tim Cook, 363

Application-specific instruction set

architecture (ASIP), 359

Application-specific integrated circuits

(ASICs), 207, 302

APS. *See* Active pixel sensor (APS)

Aptina MobileHDR, 214

ARC. *See* Antireflective coating (ARC)

- ARkit API, 188
- Artificial retinas. *See* Silicon artificial retinas
- ASILs. *See* Automotive system integrity levels (ASILs)
- Asynchronous cellular logic array (ACLA), 169
- Asynchronous/synchronous processor array (ASPA), 169, 171*t*
- Atmospheric Imaging Assembly (AIA), 281, 283*f*
- Atomic layer deposition (ALD), 260, 490, 492
- Augmented reality (AR) applications, 188–189, 211
- Automotive image sensors
- camera locations, 241, 242*f*
  - requirements for automotive qualification test criteria, 253
  - camera module interface, 251–252
  - color, 252
  - combined colors with NIR light, 252
  - high dynamic range, 247–249
  - image sensor interface, 251
  - low light performance (sensitivity), 246–247
  - optics, 253
  - resolution, 246
  - system integrity, 250
  - temperature range, 249
- sensing applications, 241
- ADAS, 244–245
  - BSD and LCA, 245
  - driver drowsiness detection, 246
  - parking assist, 245
- vision applications, 241
- night vision systems, 243–244, 243*f*
  - rearview systems, 242
  - road crossing monitoring, 244
  - surround view systems, 242–243
- Automotive market
- ADAS, 364
  - autonomous vehicles, 364
  - exterior, 365
  - interior, 365
- driving and 3D sensor use, 367–368
- exterior, 368
  - interior, 368
- electric cars, 364
- exterior automotive 3D sensing, 367
- exterior automotive 3D technologies, 367
- interior automotive 3D sensing, 365–366
- interior vision rollout, 366
- Automotive system integrity levels (ASILs), 250
- Autonomous driving, 366–368
- Autoradiographic film, 302
- Avalanche photodiode (APD), 481–484
- Avogadro's number, 290
- B**
- Backgate effect, 306
- Back-illuminated CIS (BI-CIS), 65–66
- Backside illuminated (BSI) sensors, 92, 121–122, 123*f*, 163, 174–175, 190, 194–196, 196*f*, 322–324
- annealing, 487
  - AR coatings, 84–85
  - backside charging, 84, 85*f*
  - backside thinning, 84
  - back-surface band bending and charge trapping issue, 486–487, 486*f*
  - back-surface passivation processes, 486–487
  - charge diffusion, 85
  - components, 96, 96*f*
  - conduction band edge, 486–487, 487*f*
  - delta-doped arrays, 491
  - delta doping, 487–489, 492
  - interface solutions
    - dummy wafer, 113–114
    - pad opening scheme, 111, 112*f*
    - ROIC wafer, 113–114
    - 3D bonding, 113–114, 113*f*
    - through silicon via (TSV) solution, 111
    - wafer-to-wafer electrical test structure, 113–114, 114*f*  - optical absorption, 84
  - photon absorption, 102–103, 103*f*
  - photon loss reduction with advanced AR coatings, 490
- process integration, 101, 102*f*
- bonding, 105
  - frontside processing, 102–105
  - light coupling, 110–111
  - machine vision, 101
  - material selection, 101

- surface passivation, 109–110  
wafer thinning, 105–109  
quantum efficiency, 83–84, 95  
silicon-silicon oxide interface, 486  
solar blind UV silicon detectors,  
    490–491  
superlattice doping, 487–489, 488–489f,  
    492
- Bayer pattern, 120–121, 121f, 194, 444–445,  
    444–445f
- Beam-forming optics, 337
- Beam telescope, 304
- Bellus3D, 188, 212
- Bethe formula, 290, 291f
- Bill-of-materials (BOM), 361
- Binning  
    charge shifting and clocking, 80  
    CMOS-based FPDs, 420–421  
    high-resolution image sensors, 123–124  
    patterns, 454, 455f
- Biometrics, 211
- Black body, 10–11, 13–14, 450t
- Blind spot detection (BSD), 241, 245
- Bokeh photograph, 206–207, 362
- Boltzmann constant, 10–11, 42, 125–126,  
    328–329
- Bremsstrahlung*, 293
- Broadcast cameras, 437  
    CCD and CMOS image sensors,  
        452–455  
    opto-electrical matching and other  
        parameters (*see* Opto-electrical  
        matching, broadcast cameras)  
    performance  
        0 dB setpoint, 439  
        gain errors, 439  
        linear mode, 438  
        offset errors, 439  
        parameters, 438  
        standards for, 451–452
- BSI sensors. *See* Backside illuminated (BSI)  
    sensors
- Bump-bonding, 63, 305
- Buried-channel charge coupled devices  
(BCCD), 399
- Buried photodiode (BDP)  
    FLIM applications, 399–400  
    four-transistor CMOS APS, 128, 129f
- Burst-exposure imaging, 232
- C
- Calorimeter, 293, 298–300
- Camera  
    active illumination and beam-forming  
        optics, 337  
    calibration, ToF cameras, 337–338  
    image sensor and demodulation pixels  
        BSI, 334  
        contrast and sensitivity, figure of merit,  
            335–336  
        effective sensitivity, 334  
        micro-lenses, 334  
        pixel size, 336  
    modulation concepts  
        direct ToF, 325–326  
        indirect ToF, 326  
    performance and limitations  
        am-cw ToF technique, 326–327  
        correlation function, 327–328  
        dark condition, 333  
        factors, 330–331  
        noise sources, 328–329  
        Nyquist-Shannon sampling theorem, 328  
        power budget, 329–330  
        regular condition, 332  
        sunlight condition, 332  
        receiving optics, lens and filter, 336  
        3D ToF camera, 2D images of, 333–334
- Camera-enabled mobile phones, CMOS  
    image sensor  
    advantages, 187  
    architecture and product considerations  
        analog processing for mobile  
            applications, 222  
        camera phone image sensor roadmap,  
            220–221  
        FF selfie camera, 218–220  
        parameters, 218  
        pipeline architecture, 221–222  
        RF camera, 218–219  
        sensor-embedded image processing,  
            222–224  
        sensor high-speed interface, 224
- and CCD pixel size advancement trend,  
    190–191, 190f
- core image/video capture technology  
    requirements and advances  
        ADCs, 200–201  
        image computing function, 202

- Camera-enabled mobile phones, CMOS  
 image sensor (*Continued*)  
 optics and packaging, 202–203  
 pixels and pixel array, 191–193,  
 196–200  
 semiconductor process, 193–196  
 SOC sensor, 201–202, 201*f*  
 emerging CMOS sensor-embedded  
 technologies (*see* Sensor-embedded  
 technologies, mobile imaging)  
 future trends  
 new computational imaging multi-array  
 sensors, 226–228  
 new exotic materials, new color patterns,  
 and new nonplanar, 3D pixels,  
 225–226  
 smartphone optical fingerprint sensing,  
 230  
 smartphone opto-fluidic sensors for  
 biomedical applications, 230  
 smartphone spectrometer sensing, 230  
 software camera phone and moment  
 capturing, 228–229  
 image quality performance race, 189–190  
 innovative imaging applications, 187–189,  
 188*f*  
 market growth, 186, 186*f*  
 market trends, 185–186  
 miniature cameras, 186  
 primary scene-facing camera, 185  
 Camera-on-a-chip CMOS solutions, 59  
 Canesta acquisition, 370  
 CAPD. *See* Current-assisted photodiode  
 (CAPD)  
 Car area network (CAN), 242  
 Cassini spacecraft, 493  
 CDS. *See* Correlated double sampling (CDS)  
 Cellular neural networks (CNNs) paradigm,  
 167, 168*f*  
 Cephalography, 415  
 CFA. *See* Color filter array (CFA)  
 Charge amplifier pixel, 420–421, 422*f*  
 Charge-carrier multiplier (CCM), 61  
 Charge collection time, 31  
 Charge-coupled device (CCD) image sensors,  
 58, 68–69, 119, 161, 164, 190–191,  
 322, 443  
 area arrays, 77  
 frame store CCD, 77–78, 77*f*  
 full frame CCD, 77, 77*f*  
 interline transfer CCDs, 77*f*, 78  
 orthogonal transfer array, 78  
 OTCCD, 77*f*, 78  
 broadcast cameras, 452–455  
 “buried”, 39–40  
 camera phones, 189  
 charge sensing and amplifiers, 81–82  
 charge shifting and clocking, 79–80, 79*f*  
 vs. CMOS imagers, 91–92  
 degrees of freedom, 59  
 dental X-ray imaging, 414–415  
 illumination modes, 82, 83*f*  
 backside illumination, 83–85  
 frontside illumination, 82–83  
 imaging parameters and characterization  
 charge transfer efficiency, 89–90  
 dark signal/current, 90–91  
 fixed pattern noise, 91  
 full well, 88–89  
 linearity, 90  
 photon transfer curve, 88–89  
 quantum efficiency, 86–88  
 read noise, 88  
 interlined CCD, 40, 52–53  
 layout, 76–77, 76*f*  
 linear arrays, 77  
 photon detection, 75–76  
 pixels, 76–77  
 principles of, 31–34  
 for space applications (*see* Space imaging)  
 SPADs, 61  
 time delay and integration operation, 80–81  
 two-, three-, and four-phase CCDs, 76–77  
 Charged particles detection, CMOS,  
 290–293  
 applications, 301–302  
 electron microscopy, 300–301  
 history, 302–307  
 particle physics, 298–300  
 Charge-transfer efficiency (CTE), 31–32,  
 89–90  
 Charge transfer inefficiency (CTI), 258, 266,  
 268*f*  
 Charge transfer process, 349  
 Chip-on-board (COB) method, 202  
 Chip-scale-packaging (CSP), 202  
 CineAlta, 437  
 Cinematography, 101, 109, 437, 446, 452

- CIS. *See* Complementary metal-oxide-semiconductor (CMOS) image sensors
- Clear-View, 188
- Color-difference signals (CrCbs), 438
- Color filter array (CFA), 193, 252
- CMOS camera semiconductor process, 194
  - light coupling, 111, 112*f*
- “Color” X-ray sensor, 423, 435
- Complementary metal-oxide-semiconductor (CMOS) image sensors, 58, 61, 161, 479
- advantage of, 161–162
  - automotive applications (*see* Automotive image sensors)
  - backside illuminated sensor (*see* Backside illuminated (BSI) sensors)
  - vs. CCDs, 91–92
  - fabrication, 58
  - fluorescence lifetime (*see* Fluorescence lifetime imaging (FLIM))
  - functions, 163
  - HDTV (*see* High-definition TV imaging)
  - high-performance scientific imaging (*see* High-performance scientific imaging, CIS)
  - high-performance silicon imagers (*see* High-performance silicon imagers)
  - high-resolution, 119–120
    - pixel size improvements, 121–123, 122–123*f*
    - pixel size limits, 120–121, 121*f*
    - windowing and binning, 123–124
  - high-speed, 130–131
    - digital conversion and output, 133–138, 135–137*f*, 137*t*
    - global shutters, 138, 139*f*
    - pixel improvement, 138
    - signal chain, 131–133, 132–133*f*
    - TDI sensors, 139–140
  - light-sensitive pixels, 95
  - low-noise, 124, 125*f*
    - noise correction circuits, 128–130, 129–130*f*
    - noise sources, 125–128, 126*f*
  - low-power, 140–141
  - mobile devices (*see* Camera-enabled mobile phones, CMOS image sensor)
- n* metal layers, cross section of, 33–34, 33*f*
- photodetection efficiency, 34
- polysilicon layer, 152
- smart vision chips (*see* Smart camera on a chip)
- for space applications (*see* Space imaging)
- SPAD arrays, 60
- WDR, 141–142, 142*f*
  - combined linear and logarithmic response, pixels with, 145–146, 146*f*
  - frequency-based WDR image sensors, 150–151
  - integration time control pixels, 147–150
  - logarithmic sensors, 143–144, 143–145*f*
  - multiple sampling image sensors, 151
  - threshold comparison pixels, 147–150
  - well capacity adjustment, 150, 150–151*f*
- X-rays (*see* X-ray sensor)
- Compressive sampling (CS) methods, 234
- Compton scattering, 26–27, 26*f*, 294
- Computational imaging, 230
  - burst-exposure imaging, 232
  - embedded 3D depth sensing and, 231
  - flat cameras and coded aperture, 234–235
  - HDR photography, 187
  - multi-array sensor, 226–228
  - Pixel Visual Core, 231–232
- Computed radiography (CR), 415
- Confocal scanning FLIM systems, 396
- Conservation of momentum, 290
- Contrast transfer function (CTF), 439–440
- Conversion gain (CG), 30, 61, 421–423
- Core image/video capture technology, mobile applications
  - ADCs, 200–201
  - categories, 191, 192*f*
  - image computing function, 202
  - optics and packaging, 202–203
  - pixels and pixel array
    - APS sensor, 196–198
    - architectures and technology advancement, 191, 192*f*
    - CDS for fixed pattern noise reduction, 196–198
    - digital CDS method, 196–198
    - image capture performance, 191
    - “More-than-Moore” technologies, 198–200
  - optical formats, 191–193, 196, 197*t*

- Core image/video capture technology, mobile applications (*Continued*)  
 optical pathway, 191, 200  
 photodiode's well capacity, 200  
 pinned photodiode architecture and methods, 196–198  
 scaled performance, 193  
 shading and pixel crosstalk, 200  
 size, die size, and resolution, 196, 197*f*  
 SNR performance, 196–198, 198–199*f*  
 transistor sharing, 200
- semiconductor process  
 color filter arrays, 194  
 dark current, 193  
 defects, 193  
 Dzero, 193  
 hot pixel, 193  
 light guides, 194–195, 195*f*  
 manufacturing yield and cost, 194  
 wafer-level BSI process technology, 194–196, 196*f*
- SOC sensor, 201–202, 201*f*
- Correlated double sampling (CDS), 51–52, 128–130, 129*f*, 196–198, 271, 454–455
- Cross-correlation function, 345, 347
- CTE. *See* Charge-transfer efficiency (CTE)
- CTI. *See* Charge transfer inefficiency (CTI)
- Current-assisted photodiode (CAPD), 350–351
- Cyclic ADC, 137*f*, 200
- Cyclic redundancy codes (CRC), 250
- D**
- Dark current (DC), 30, 90–91, 103–104, 125, 140–141, 152–153, 152*f*, 193, 265*f*, 266, 267*f*
- Dark current nonuniformity (DCNU), 125
- Dark current shot noise, 48, 328–329
- Dark noise, 91, 328–329, 332–333, 335–336, 395–396
- Dark shot noise, 30
- Dark signal, 90–91, 262
- Dark signal nonuniformity (DSNU), 68–69, 91
- DDS. *See* Digital double sampling (DDS)
- Deep learning smart image sensors, 235
- Deep trench isolation (DTI), 104–105, 122, 123*f*, 205
- Delphi experiment, 303
- Delta-doped arrays, 488–489*f*, 491
- Delta doping, 487–489, 492
- Demodulation  
 contrast, 342–344  
 correlation theory, 345–347  
 device approaches  
 CAPD, 350–351  
 CCD pixel, 348–349  
 continuously modulated optical radiation field, 347  
 electro-optical shutter mechanism, 347–348  
 PMD, 349–350  
 QEM, 353–355  
 requirements, demodulation pixels, 347–348  
 sampling demodulation/just demodulation pixels, 347  
 static drift-field demodulation, 352–353
- efficiency, 342–344
- pixel arrays, 322
- sampling theory, 342–345
- ToF image sensor architecture  
 amplifiers and ADC, 357  
 clock generation and phase shifting, 357–358  
 micro-sequencer, 359  
 pixel electronics, 356–357  
 pixel processor, 358–359
- Density effect, 290
- Dental X-ray imaging, 413–415
- Deoxyribonucleic acid (DNA) sequencing, 378
- Depleted p-channel field effect transistor (DEPFET), 310
- Depletion region. *See* Space-charge region (SCR)
- Design for manufacturing (DFM), 423
- Detectable quantum efficiency (DQE)  
 Thales HD sensor, 429, 433*f*  
 Trixell's PX4700 flat panel sensor, 427–428, 427*f*
- Diffractive optical element (DOE), 337
- DiffuserCam, 234–235
- Diffusion cloud, diameter of, 85

- Digital double sampling (DDS), 462–464, 465*f*
- Digital imagers, 473
- Digital pixel sensor (DPS), 132–133, 170–171, 174
- Digital signal regeneration stage, 44, 52
- Digital TDI, 257
- Digital-to-analog converter (DAC), 134
- Direct bonding<sup>TM</sup>, 105
- Direct Bond Interconnect (DBI), 210
- Direct conversion X-ray sensors, 429–434
- Discrete Fourier transform (DFT), 342–344
- Displacement damage (DD), 258
- dark current and hot pixels, 266, 267*f*
- uncorrected CTI, in proton-irradiated Hubble CCD, 266, 268*f*
- Double data sampling (DDS), 390
- Double delta sampling (DDS), 51–52
- Double sampling, 417–418
- DPS. *See* Digital pixel sensor (DPS)
- Driver drowsiness detection, 246
- DSNU. *See* Dark signal nonuniformity (DSNU)
- 3D stacking, 208
- challenge, 210–211
- 3D IC stacking, 62, 65
- digitizing, I/O, and image computation, new methods for, 209
- new camera capabilities/features, 209–210
- new generations of pixel arrays, 208–209
- DTI. *See* Deep trench isolation (DTI)
- Dual cameras, smartphones, 207
- Dual gain 3T pixel, 421–423, 422*f*
- Dual-slope ADC (DS ADC), 134, 137*f*
- DxO mobile mark, 208
- Dynamic pixel management, 445
- Dynamic range (DR), 119, 122, 140–143, 142*f*, 147–148, 150–151, 150*f*, 153–154, 262, 325–326, 428
- Dynodes, 20
- E**
- Earth observation (EO) satellites, 256
- Edge-emitting semiconductor lasers (EEL), 337
- Electromagnetic interference/  
electromagnetic compatibility (EMI/EMC), 251
- Electron-bombarded arrays, 481, 482*f*
- Electron counting, 306–307
- Electron energy, 17, 18*f*
- Electron-hole pair (EHP), 17–18, 26, 30, 474–475
- Electron microscopy, 300–301
- Electron multiplying charge-coupled devices (EMCCD), 377, 395–396, 483, 483*f*
- Electro-optical shutter mechanism, 347–348
- Electrostatic potential barrier, 42
- Energy-gap energy, 17–18
- Energy loss distribution (straggling function), 290–292, 291*f*
- Equivalent noise charge (ENC), 47
- Euclid CCD, 280, 281*f*
- European Machine Vision Association (EMVA1288), 247
- Explosives sensing, 385
- Exposure bracketing, 213
- External clock TDC (EC-TDC), 393–395, 394*f*
- Extraoral dental X-ray imaging, 414–415
- Extreme ultraviolet (EUV), 307–308
- F**
- Fabry-Perot interferometer, 49
- Facebook, 363
- FaceTime, 185
- Fano factor, 295–296
- Fermi-Dirac distribution, 42
- Fermions, 25–26
- Fiber optic plate (FOP), 297
- Field-free region, 85
- Field of view (FoV), 242–243, 330
- Figure of merit (FOM), 335–336, 450–451
- Fish-eye lenses, 243, 246
- Fixed-pattern noise (FPN), 32, 91, 124–125, 125*f*, 127–129, 143–146, 152–154, 163–164, 438–439
- Fixed pattern phase noise, 358
- FlatCam, 235
- Flat panel display (FPD)
- CIS-based FPD technology
- advantages, 416
- binning, 420–421
- charge amplifier pixel, 421, 422*f*
- Dalsa's dual gain 3T pixel, 421–423, 422*f*

- Flat panel display (FPD) (*Continued*)  
 four transistor (4T) pixel architecture, 419–420  
 on-chip ADC, 423  
 1T pixel architecture on a:Si TFT FPD, 417–418  
 “photon counting” pixel, 423  
 three transistors (3T) pixel architecture, 418–419  
 visible light CIS pixel architectures, 417  
 yield and cost considerations, 423–425
- fluoroscopy, 416  
 mammography, 416
- Flexible industrial token (FIT) generator, 459, 462
- Flexible vertical token (FVT) registers, 459, 462
- Floating diffusions (FDs), 52, 55–56, 121, 419–420
- Fluorescence lifetime imaging (FLIM), 377–378  
 applications, 385  
 CMOS detectors and pixels  
   advantages, 386  
   frequency-domain lifetime pixels, 386–389  
   non-imaging applications, 386  
   TCSPC pixels, 393–395  
   time-gating pixels, 389–393  
 equipment needed for, 378–379  
 fluorescence lifetime, 379–381  
 Jablonski diagram, 379, 379*f*  
 lifetime measurement techniques, 381–384  
 microarray applications, 378  
 point scanning, 381  
 sources of information, 378  
 streak cameras, 384  
 stroboscopic excitation, 385  
 system-on-chip  
   CMOS miniaturized PMTs, 396–398  
   confocal scanning FLIM systems, 396  
   frequency-domain lifetime imaging/  
     sensing, 399  
   latest SPAD-TCSPC arrays, 399–400  
   photon-capture devices, comparison of, 398*t*  
   time-domain, embedded algorithms, 400–405  
   wide-field FLIM systems, 395–396  
     up-conversion methods, 384–385  
     wide-field imaging, 381
- Fluorescence resonance energy transfer (FRET), 380–381, 385, 395, 404
- Fluoroscopy, 413, 416
- Focal plane arrays (FPAs), 492, 494*f*
- Fovea, 3–5  
 Fovea centralis, 6–7
- FPD. *See* Flat panel display (FPD)
- FPN. *See* Fixed-pattern noise (FPN)
- Frame interline transfer (FIT), 452
- Frame store CCD, 77–78, 77*f*
- Frame transfer (FT), 452
- Frame transfer charge-coupled device (FT-CCD), 440
- Frequency-domain FLIM, 377–378, 381  
 CMOS imaging sensors with, 399  
 pixels, 386–389  
 principle of, 383–384, 383*f*
- Front-facing (FF) selfie cameras, 185–186, 203–205, 218–220
- Front-illuminated CCDs, 82–83
- Front side illuminated (FSI) sensor, 121–122, 123*f*, 174, 195–196  
 components, 95–96, 96*f*  
 photon absorption, 102–103, 103*f*  
 scaled-down FSI sensor  
   antireflecting coating layer, 101  
   Apina sensor, 100, 100*f*  
   back-end scheme, 97–98, 97*f*  
   copper inter-dielectric layers, 98  
   light pipe concept, 99–100, 99*f*  
   metallization, 98–99  
   micro-lens, 97–98  
   passivation layer, 97–98  
   PD active region, 97–98
- Full frame CCD, 77, 77*f*
- Full well, 88–89
- Full well capacity (FWC), 51, 88, 91–92, 348
- Fully pinned photodiode, 419–420
- G**
- Gaia space telescope, 280–281, 282*f*
- Galactic cosmic rays (GCRs), 258, 263
- Gated-image intensifier (GII), 386
- Gated photodiode, 389–390, 390*f*
- Gating methods, 401–402
- Gaussian process, 125–126

- Genesis, 437, 446  
Gesture control, 365–366  
Global shutter (GS) sensors, 115, 138, 139f,  
    215–216  
Google, 363  
Google Glass, 362–363  
Ground sampling distance (GSD), 257
- H**
- Hasselblad X1D-50c, 208  
HDR. *See* High dynamic range (HDR)  
HDTV. *See* High-definition TV imaging  
Head-mounted displays (HMD), 369  
Head-up display (HUD), 362  
 $^3\text{H}$ . Film, 302  
High-definition-dynamic pixel management  
    (HD-DPM) imager, 452–453, 454f  
High-definition TV imaging  
    aliasing and OLP filtering, 443–447  
    bit size, pixel count, and other issues,  
        464–467  
    broadcast cameras (*see* Broadcast cameras)  
    MTF, aliasing and resolution, 439–443  
    SNR (*see* Signal-to-noise ratio (SNR),  
        HDTV)  
    three-dimensional television, 467  
    UHDTV, 467–468  
High dynamic range (HDR), 358–359  
    computational imaging, 187  
    examples of, 248  
    image capture, limitations of  
        sensor-based HDR in-pixel method,  
            213–214  
        sensor-based HDR in-sensor method,  
            214–215  
    intra-scene dynamic range, 247–248  
    *vs.* linear image, 247–248, 247f  
    multiple lower dynamic range images,  
        248  
    night scene, dynamic range of, 247–248,  
        248f  
    non-linear response curve with kneepoints,  
        248–249  
High-k materials, 35  
Highly doped epitaxial-silicon layers,  
    109–110  
High-performance scientific imaging, CIS  
    charged particles detection, 290–293  
applications, 301–302  
electron microscopy, 300–301  
history, 302–307  
particle physics, 298–300  
detection  
    particles and photons, indirect  
        detection of, 296–297  
    of photons, 293–296  
    in silicon, 290–297  
substrates, 297  
X-ray detection  
    advanced applications, 307–308  
    silicon and non-CMOS detectors,  
        309–310  
    stitching, 308–309  
High-performance silicon imagers  
    back illumination (*see* Backside  
        illuminated (BSI) sensors)  
    biological and medical applications,  
        495–496  
    commercial applications, 495  
    in Earth-orbiting telescopes, 473  
    photon counting detectors, 480–485  
    planetary and astronomy applications,  
        492–494  
    readout structures, 479, 480f  
    scientific imaging detectors, 478–479  
    silicon digital imagers, 473  
    solid-state imaging (*see* Solid-state  
        imaging detectors)  
High-resolution CMOS image sensors,  
    119–120  
pixel size improvements  
    APS, 121  
    BSI, 121–122  
    electrical crosstalk, 122, 123f  
    FSI, 121–122  
    light guides, 121, 123f  
    microlenses, 121, 123f  
    optical crosstalk, 122, 123f  
    row-selection transistors, 121, 122f  
pixel size limits  
    airy disk, 120  
    airy pattern, 120  
    Bayer pattern, 120–121, 121f  
        first-order approximation, 120  
    windowing and binning, 123–124  
High-spatial resolution autoradiograms,  
    302

- High-speed CMOS image sensors, 130–131  
 digital conversion and output, 133–134  
 ADC topologies, 134, 137*t*  
 cyclic ADC, 137*f*  
 DS ADC, 134, 137*f*  
 SAR ADC, 134, 135*f*  
 SD ADC, 137*f*  
 SS ADC, 134, 136*f*  
 typical circuit implementation, 135*f*
- global shutters  
 fast moving scenes, rolling shutter on, 138, 139*f*  
 five-transistor circuit, 138, 139*f*
- pixel improvement, 138  
 signal chain, 131–133, 132–133*f*  
 TDI sensors, 139–140
- High-speed image processing chips, 170–174
- High Speed Serial Pixel Interface (HiSPi), 200–201
- HiRho technology, 260
- Hold switch, 356–357
- Holoportation, 211
- Horizontal optical low pass filter (H-OLP), 439
- Horizontal synchronization signal output (HSYNC), 251
- Horizontal timing generator (HTIME), 462
- Hot/white pixels, 193
- Hubble space telescope  
 Crab Nebula, 256*f*  
 proton-irradiated CCD, CTI in, 266, 268*f*  
 Teledyne e2v CCD43 assembly, 280, 280*f*
- Human vision system  
 accommodation process, 8  
 acuity, 9  
 air-cornea interface transmittance, 5  
 binocular vision/stereopsis, 8–9  
 collagen fibrils, 5  
 density of rods and cones, 6, 7*f*  
 duplicity theory, 5–6, 6*t*  
 field of view, 5  
 fovea centralis, 6–7  
 human eye diagram, 3–5, 5*f*  
 light rays penetrating eye, geometry of, 7–8, 8*f*  
 luminance range, 10  
 passive cone photoreceptors, 3–5
- pupil, 5  
 quanta of light (photons), 9  
 scotopic (nocturnal) and photopic (diurnal) visions, 5–6  
 video stream, 9–10
- Hybrid bonding technology, 360
- Hybrid CCD-on-a-CMOS, 213
- Hybrid pixel detectors, 301–302
- Hyperspectral imaging, 257, 258*f*  
 Sentinel 3 image, 277–278, 279*f*  
 Sentinel 5P image, 278, 279*f*  
 Sentinel 2 VNIR Flight Focal Plane, 277, 278*f*
- I**
- IC-TDC. *See* Internal clock TDC (IC-TDC)
- IEM. *See* Integration for extraction method (IEM)
- Image intensifier tube (IIT), 413, 416
- Image signal processor (ISP), 202
- Image system processor (ISP), 222–223
- Image tube-based technologies, 480
- Impinging photons, 26–27
- IMX400 chip, 210
- In cabin monitoring (ICM), 366
- Infrared (FIR) cameras, 243–244
- Inner tracking system (ITS), 305
- In-pixel method, sensor-based HDR, 213–214
- In-sensor method, sensor-based HDR, 214–215
- Integrated circuits (ICs), 22, 31, 196–198, 271
- Integrated signal processor (ISP) chip, 252
- Integration for extraction method (IEM), 402–404, 403*f*
- Intelligent photography, 211
- Intensified CCD (ICCD), 377
- Interline CCD image sensors, 52
- Interline transfer (IT) CCDs, 77*f*, 78, 216, 452
- Internal clock TDC (IC-TDC), 393–395, 394*f*
- Internal extrinsic transition, 27
- Internal quantum efficiency (IQE), 476, 488–489*f*
- International Linear Collider, 305–306, 310
- International Standardization Organization (ISO), 253–254

International Technology Roadmap for Semiconductor (ITRS), 193  
International Telecommunications Union (ITU), 451  
Internet of things (IoT), 360  
Intraoral dental X-ray imaging, 414–415  
Intra-scene dynamic range, 247–248

## J

JANUS camera, CIS115, 274–275, 275f  
Jell-O, 215  
Johnson/Nyquist noise. *See* Thermal noise  
Jots, 61  
Just noticeable differences (JND), 437–438

## K

$k_B T/C$  ( $k$ -T-over- $C$ ) noise, 47  
Klik, 188

## L

Lambert reflector, 330  
Landau distribution, 290–292  
Lane change assist (LCA), 241, 245  
Large-area sensors, 289  
Large electron positron (LEP) collider, 303  
Large Hadron Collider (LHC), 298–300, 299f, 303  
Laser annealing, 109  
Laser detection and ranging (LADAR), 319  
Light coupling, 110–111  
Light detection and ranging (LIDAR), 244–245, 319, 336, 367–368  
Light-emitting diodes (LEDs), 206, 322–324, 337, 365, 383–384  
Light flux, 87  
Light guides (LGs), 194–195, 195f  
Linear absorption coefficient, 294, 295f  
Linear energy transfer (LET), 268  
Linear mode, 438  
Linear sensors, 257, 301–302  
Logarithmic pixels, 143f, 144–145, 145f, 153f  
Lot acceptance test (LAT) program, 262–263  
Low Earth orbit (LEO), 258, 263  
Low-level image processing, 170–173  
Low-light-level CCDs (LLLCCDs), 61  
Low-noise CMOS image sensors  
    FPN correction, 124, 125f  
    noise correction circuits

CDS, 128–130, 129f  
four-transistor CMOS active pixel sensor, 128, 129f  
input referred noise, 130  
signal chain, 129–130, 130f  
noise sources  
    ADCs, 127–128  
    DCNU, 125  
    first-order model, 126–128  
    three-transistor CMOS active pixel sensor, 125, 126f  
Low noise X-rays, 310  
Low-power CMOS image sensors, 140–141  
Low-voltage differential signaling (LVDS), 243, 251, 462–463, 463f  
Lumogen, 296–297

## M

Magic Leap, 363  
Mammography, 413, 416  
Marquardt-Levenberg algorithms, 396  
Mass spectroscopy (MS), 301–302  
Matrix-assisted laser desorption/ionization (MALDI), 301–302  
Maximum electrical field ( $E_{max}$ ), p-n junction, 40  
Medipix chip, 309–310  
MEDIPIX IC, 63  
MegaFrame TDC pixels, 393–395  
3-Meg imager, 113–114  
Metal grid, 110  
Metal-oxide-semiconductor-capacitor (MOS-C) structure-based photodetectors  
    cross section and band diagram, 35, 36f, 37–39, 38f  
    disadvantages, 39–40  
    “inversion” layer, 39  
    on p-type silicon, 35  
    quasi-Fermi level, 39  
    space-charge region, 35–39  
    threshold voltage, 39  
Metal-oxide semiconductor field-effect transistor (MOSFET), 264, 265f, 269  
Metal-oxide-semiconductor (MOS) transistors, 302–303  
MetImage pixel format, 277, 277f  
Michelson-Morley experiment, 14

- Micro-channel plate photo-multiplier tube (MCP-PMT), 378–379, 381–382
- Microchannel plates (MCP), 20–22, 21*f*, 322
- Microlenses, 82, 111, 121, 194, 322–324
- Micro-pixel avalanche photodiodes (MAPD), 60–61
- Microsoft, 363
- HoloLens, 369–370
  - Skype technology, 219
  - X-box Kinect, 211, 216, 360
- MIMOSA28, 304–305
- Minimum ionizing particles MOS array (MIMOSA), 304
- Mobile devices, CMOS image sensor.
- See* Camera-enabled mobile phones, CMOS image sensor
- Mobile imaging
- applications of, 185
  - camera phones (*see* Camera-enabled mobile phones, CMOS image sensor) challenge of, 187
- Mobile market
- apps and functionality, 362
  - AR
    - consumer mobile AR entries, 363
    - and mobile phone evolution, 362–363
    - phone accessory, 363
    - 3D sensors and consumer AR headsets, 363–364
  - front-facing “selfie” 3D imaging, 362
  - 3D sensors and mobile volume, 364
  - world-facing 3D, 361
- Mobile silicon imaging
- camera advances
    - dual cameras, 207
    - DxO mobile mark, performance assessment, 208
  - computational imaging
    - burst-exposure imaging, 232
    - flat cameras and coded aperture, 234–235
  - Pixel Visual Core, 231–232
  - deep learning smart image sensors, 235
  - 3D stacking methods and attributes, 208–211
  - mega-application-phases, 231
  - multiwavelength sensors, 234
  - new mobile UXs, 233–234
- pixel advances
- small CMOS pixels, 203–206
  - special phase detection focus pixels, 206
  - special pixels distance-for-Bokeh
    - photograph blurring, 206–207
- single-photon CMOS sensors, 234
- smartphones
- AR/VR viewers, image sensors for, 233
  - edge-based machine perception, 232–233
- 3D camera advances
- improved NIR sensitivity, 212–213
  - machine perception and 3D depth imaging, 211–212
- Modulated electron multiplied (MEM)-FLIM sensor, 388–389, 388*f*
- Modulation transfer function (MTF), 85, 102–103, 120, 416, 438
- airy disc at different *f*-numbers, 441, 442*f*
  - aliasing and OLP filtering, 443–447
  - curves of electrical and optical elements in camera, 442, 442*f*
  - diffraction limited lens, 441
  - direct conversion X-ray sensors, 433–434
  - Nyquist frequency, 426–428, 441, 441*f*
  - pixel aperture, 440, 441*f*, 443–444, 443*f*
  - square wave, 439–440
  - Thales HD sensor, 429, 432*f*
  - Trixell’s PX4700 flat panel sensor, 426–428, 427*f*
  - X-ray phantom, 426, 426*f*
  - zone chart, 440, 440*f*, 444–445
- Monolithic active pixel sensors (MAPS), 305
- Moore’s law, 174, 190–191, 190*f*, 220, 438
- Moore’s Law of Imaging, 191
- Mosaic paintings, 3–5, 4*f*
- MOS-CCD EPIC camera, 281, 282*f*
- MOSFET. *See* Metal-oxide semiconductor field-effect transistor (MOSFET)
- MTF. *See* Modulation transfer function (MTF)
- Multi-array (MA) sensor, 226–228
- Multifrequency algorithms, 357
- Multimedia optical synchronous transport (MOST), 252
- Multi-shot method, 213
- Multiwavelength sensors, 234

**N**

- Natural sampling, 344  
Near-infrared (NIR) sensing, 205–206, 212–213  
Near-infrared (NIR) spectrum, 319  
Neutral density (ND) filters, 439  
New car assessment program (NCAP), 244–245  
Night vision systems, 243–244, 243f  
Noise electron density (NED), 457–458  
Noise equivalent dose (NED), 428  
Noiseless co-addition. *See* Binning  
Non-Bayer color filter, 225–226  
Non-gating embedded algorithms, 402–405  
Nonlinear least square methods (NLSM), 400–401  
n-type metal-oxide-semiconductor (nMOS)  
transistor, 143, 143f, 148, 149f, 463  
Nyquist frequency, 426–428, 441, 441f, 444–445  
Nyquist's criteria, 443  
Nyquist-Shannon sampling theorem, 328, 344–345  
Nyxel pixel, 212

**O**

- Ocean and Land Color instrument (OLCI), 277–278  
Ocelot, 60  
One-dimensional (1D) velocity sensor, 166–167  
“One measurement sequence”, 359  
One transistor (1T) pixel architecture on a:Si  
TFT FPD, 417–418  
Optical fingerprint sensors, 230  
Optical image stabilization (OIS), 232  
Optical low pass (OLP) filter, 440, 443–447  
Opto-electrical matching, broadcast cameras  
image diagonals in broadcast, 447–448, 448t  
Noise and sensitivity improvements, 450–451, 451f  
photon density for black body radiator  
temperature, 449–450, 450t  
photon flux of blackbody radiator, 448–449, 448f

- photons and electrons in HDTV pixel, 449–450, 450t  
three-imager cameras, 447  
Opto-fluidic sensors, 230  
“Opto” process, 304–305  
Orthogonal transfer array (OTA), 78  
Orthogonal transfer CCD (OTCCD), 77f, 78  
OSIRIS camera, Rosetta spacecraft, 274, 274f  
Oxygen sensing, 385

**P**

- Packaging  
CMOS sensor-based camera, 202–203  
for space applications, 271–272  
Pair creation, 294  
Pair production, 26–27, 26f  
Parallel binning, 80  
Partially pinned photodiode (PPPD), 420–421  
“Particle-antiparticle” pair, 27  
Particle physics, 298–300  
Particle-wave duality principle, 300  
Partition noise, 46–48  
Passive triangulation, 338–339  
Pauli exclusion principle, 25–26  
67P/Churyumov-Gerasimenko, 274, 274f  
Phase-detection autofocus (PDAF) approach, 206–207  
Phase-fluorescent sensor, 387–388, 387f  
Phase-locked loop (PLL) circuit, 357–358  
Photocurrent integration time, 31  
Photodetector technology, 18–22  
Photodiode (PD)  
cyclic reset operation, 44  
definition, 42–44  
output voltage signal, 42–44  
photo-carriers, 95  
spurious noise components, 46  
3T pixel, 46, 47f, 48  
Photoelectric effect, 15–16, 17f, 26, 294  
Photogates, 39–40, 49  
Photogate 4T active pixel structure, 56  
Photolithography, 104–105, 308–309  
Photometry, 9–12  
Photomultiplier tubes (PMTs), 19–20, 20f, 377–378, 396–398

- Photon counting detectors  
 all-solid-state imagers with avalanche gain, 480  
 electron-bombarded arrays, 481, 482*f*  
 image tube-based technologies, 480  
 quanta image sensor, 480  
 solid-state detectors with gain  
   avalanche photodiodes, 481–484  
   electron multiplying CCDs, 483, 483*f*  
   S/N ratio, 481–483  
   SPAD, 484–485, 484–485*f*
- Photonic mixer device (PMD), 322, 349–350
- Photon transfer method, 296
- Photopic relative luminous efficiency function, 12
- Photopic relative luminous sensitivity, 11, 11*f*
- Photo response nonuniformity (PRNU), 91, 110, 439
- Photosensing, 12–18
- Pinned photodiode (PPD), 53–56, 54*f*, 193, 264, 353, 419–420
- Pixel IMaging for Mass Spectroscopy (PImMS) sensor, 305–306
- Pixels
- FLIM
    - frequency-domain lifetime pixels, 386–389
    - TCSPC pixels, 393–395
    - time-gating pixels, 389–393
  - high-performance pixel structures
    - digital signal regeneration stage, 52
    - floating diffusion, 52, 55–56
    - global shutter efficiency, 56
    - image lag, 52–53
    - interline CCD cell, 52, 53*f*
    - no image lag photodiode structure, 52
    - pinned photodiode, 53–56, 54*f*
    - SN capacitance, 56
  - mobile applications
    - APS sensor, 196–198
    - architectures and technology
      - advancement, 191, 192*f*
    - CDS for fixed pattern noise reduction, 196–198
    - digital CDS method, 196–198
    - image capture performance, 191
    - “More-than-Moore” technologies, 198–200
  - new exotic materials, new color patterns, and new nonplanar, 3D pixels, 225–226
  - optical formats, 191–193, 196, 197*t*
  - optical pathway, 191, 200
  - photodiode’s well capacity, 200
  - pinned photodiode architecture and methods, 196–198
  - scaled performance, 193
  - shading and pixel crosstalk, 200
  - size, die size, and resolution, 196, 197*f*
  - small CMOS pixels, 203–206
  - SNR performance, 196–198, 198–199*f*
  - special phase detection focus pixels, 206
  - special pixels distance-for-Bokeh photograph blurring, 206–207
  - transistor sharing, 200
- noise considerations
- complex impedance, 46–47
  - correlated double-sampling, 51–52
  - double delta sampling, 51–52
  - impinging radiation wavelength-dependent transmission curve, 49, 50*f*
  - $k_B T/C$  (*k*-T-over-*C*) noise, 47
  - low-frequency/flicker noise, 45–46
  - partition noise, 46–48
  - PD 3T pixel, 46, 47*f*, 48
  - polysilicon layers, 49
  - random telegraph signal (RTS) noise, 45
  - reset noise, 46
  - SF noise, 48
  - shot noise, 45, 48–49
  - signal noise contributions, 50–51
  - signal-to-noise ratio, 44, 46, 49–51
  - silicon-nitride-based passivation layer, 49, 50*f*
  - spectral density, 45
  - standard deviation, 45
  - thermal noise, 45–46
  - variance, 45
  - white noise reduction, 52
- smart camera on a chip, 164–165
- source-follower in-pixel amplifier, 44
- three transistor active pixel, 44
- Pixel Visual Core, 231–232
- Planck’s law, 10
- Planck’s quantum theory, 16
- Plank’s constant, 300
- Plumbicon, 438

- PMD. *See* Photonic mixer device (PMD)
- PMTs. *See* Photomultiplier tubes (PMTs)
- p-n* junction-based photodetectors, 40–44
- Poisson probability distribution, 30, 48–49
- Positron emission tomography (PET), 20
- Power dissipation, 259, 269–270
- PPD. *See* Pinned photodiode (PPD)
- Primary scene-facing camera, mobile phone, 185
- Processing elements (PEs), 161–162, 164–165
- Programmable versatile large scale artificial retina (PVLSAR), 169, 171<sup>t</sup>
- Proximity sensors, 217
- Pseudo-noise (PN), 326
- p-type metal-oxide-semiconductor (PMOS) transistors, 148, 149<sup>f</sup>, 305
- Pulse frequency modulation (PFM), 134–136
- Pulse width modulation (PWM), 134–136, 141
- Pure View 808, 219
- Pyramid surface light-diffraction structures, 212
- Q**
- Quanta image sensor (QIS), 61, 234, 480
- Quantum dot photodetector, 212
- Quantum efficiency (QE), 86–88, 194, 475–476
- active silicon thickness, 260, 261<sup>f</sup>
  - AR coatings, 260, 261<sup>f</sup>
  - surface passivation, 260
- Quantum efficiency modulation (QEM), 353–355
- Quantum yield (QY), 87–88, 476
- Quincunx, 444
- R**
- Radiant flux, 12
- Radiation hardness, 258
- displacement damage, 266–268, 267–268<sup>f</sup>
  - ionizing damage, 264–266, 265<sup>f</sup>
  - radiation environment, in space, 263–264
  - single event effects, 268–269, 270<sup>f</sup>
- Radio detection and ranging (RADAR), 244–245, 367
- Radiography, 413, 415–416
- Radiometry, 12
- Random telegraph signals (RTSs), 45, 264, 266
- Rayleigh-Jeans law, 13–14
- “RCCC” CFA, 252
- Read-after-exposure global shutter approach, 357
- Read noise, 61, 88, 91, 328, 450–451, 455
- Readout integrated circuit (ROIC), 66, 113–114
- Rear-facing (RF) scene camera, 185–186, 203–205, 218–219
- Rearview systems, 242
- Reduced twisted pair gigabit Ethernet (RTPGE), 252
- Reset noise, 46, 51, 125–126
- Reset switch, 145, 356–357
- Reset transistor, 47–48, 56, 81–82, 81<sup>f</sup>, 148, 418
- Retma charts, 447
- Reverse start-stop TCPSC principle, 381, 382<sup>f</sup>
- Rewind, 188
- RGBZ sensor, 212
- Ring oscillator-based TDCs, 394<sup>f</sup>, 395
- ROIC. *See* Readout integrated circuit (ROIC)
- Rolling shutter, 452, 462
    - fast moving scenes, effects on, 139<sup>f</sup>
    - limitations of, 215–216
- RTSs. *See* Random telegraph signals (RTSs)
- Rutherford-Bohr model of the atom, 14–15, 17–18, 17<sup>f</sup>
- S**
- SCAMP-3 vision chip, 169, 170<sup>f</sup>, 171<sup>t</sup>
- Scanistor, 32
- Scanning electron microscopes (SEM), 300–301
- Scientific CMOS (sCMOS) sensors, 377
- Scientific imaging detectors, 478–479
- Scintillation, 78, 414
- Secondary emission, 20
- Secondary ion mass spectrometry (SIMS), 301–302
- SEE. *See* Single event effects (SEE)
- SEL. *See* Single event latchup (SEL)
- Self-scanned silicon image detector arrays, 32

- Semiconductor imagers, 473, 481–483  
 Semiconductor process, mobile applications  
     color filter arrays, 194  
     dark current, 193  
     defects, 193  
     3D stacking methods and attributes, 208–211  
     Dzero, 193  
     hot pixel, 193  
     light guides, 194–195, 195*f*  
     manufacturing yield and cost, 194  
     wafer-level BSI process technology, 194–196, 196*f*  
 Sensing systems, 241  
     ADAS, 244–245  
     BSD and LCA, 245  
     driver drowsiness detection, 246  
     parking assist, 245  
 Sensor-embedded image processing, 222–224  
 Sensor-embedded technologies, mobile  
     imaging, 204*f*  
     to achieve superior image and video capture experiences, 203, 204*f*  
     3D and depth capture, limitations of, 216–217  
     HDR image capture, limitations of  
     in-pixel method, 213–214  
     in-sensor method, 214–215  
     mobile silicon imaging (*see* Mobile silicon imaging)  
     rolling shutter and GS image capture methods, limitations of, 215–216  
 Sentinels  
     Sentinel 3 image, algal blooms, UK, 277–278, 279*f*  
     Sentinel 5P image, global carbon monoxide pollution, 278, 279*f*  
     Sentinel 2 VNIR Flight Focal Plane, 277, 278*f*  
 Serial binning, 80  
 Serial registers, 80  
 Shallow trench isolation (STI), 205, 264  
 Shockley-Read-Hall (SRH) generation-recombination centers, 30  
 Shot noise, 45, 48–49, 80, 124, 328–329, 425–426  
 Sigma-delta ADC (SD ADC), 137*f*  
 Signal-to-noise ratio (SNR), 247, 328  
     HDTV, 456–464  
     CCD imager, 459  
     CMOS imager, 458–459, 461*f*  
     detection node capacitance, 457–458  
     floating diffusion dark current, histogram of, 463, 464*f*  
     image after DDS, in global shutter mode, 463–464, 465*f*  
     luminance definition for, 456  
     LVDS, 462–463, 463*f*  
     minimum channel length *vs.* oxide thickness, 457, 457*f*  
     NED, 457–458  
     state-of-the-art amplifier noise performance, 458, 460*f*  
     static noise measurement of on-chip amplifier, 458, 459*f*  
     5T-pixel with limiter transistor, 462, 462*f*  
     Xensium-FT, performance parameters of, 463–464, 465*f*  
 pixels, 44, 46, 49–51, 196–198, 198–199*f*  
 of SPADs, 396–397  
 X-ray sensor, 420–421, 425–426  
 Silicon artificial retinas  
     arrays of identical pixels, 165–166  
     characteristics, 165  
     Mead's artificial retina, 166, 166*f*  
     two-dimensional (2D) silicon retina, 166–167  
     VLSI computational sensor, 166–167  
 Silicon-based planar-technologies, 34–35  
 Silicon carbide, 271–272, 272*f*  
 Silicon dioxide (SiO<sub>2</sub>), 264  
 Silicon image sensors  
     CCD and CMOS photosensing technologies, 31–34  
     hybrid and 3D detector technologies  
     contemporary digital still camera sensors, 65  
     3D IC stacking, 62  
     3D integration technology, 65–66, 68, 68*f*  
     FE-TC4, readout IC prototype, 66, 67*f*  
     flip-chip technology, 62–63  
     “More Moore” developments, 61–62  
     through silicon via (TSV) technique, 63–64, 64*f*  
     miniaturization, 56–59

- MOS-C structure-based photodetectors, 34–40  
pixel structures  
    high-performance pixel structures, 52–56  
    noise considerations in, 44–52  
*p-n* junction-based photodetectors, 40–44  
silicon phototransduction, 25–31  
single-photon counting, 59–61  
Silicon-nitride-based passivation layer, 49, 50  
Silicon on insulator (SOI), 101, 212, 297  
Silicon photomultipliers (SiPMs), 60–61  
Silicon phototransduction  
    absorption coefficient, 28–29  
    charge collection time, 31  
    Compton scattering, 26–27, 26*f*  
    dark current, 30  
    diffusion length, 30  
    electron-hole pair, 26, 30  
    extrinsic photoelectric effect, 27–28  
    impinging radiant flux, 29  
    impinging radiation, 26–27, 26*f*  
    internal extrinsic transition, 27  
    internal photoelectric effect, 26  
    light absorption depth, 29  
    pair production, 26–27, 26*f*  
    “particle-antiparticle” pair, 27  
    Pauli exclusion principle, 25–26  
    photodetector output current, 30–31  
    photon flux, 28–29  
    spectral responsivity, 31  
    wavelength and photon energy, 27, 28*f*  
Simultaneous location and mapping (SLAM), 362  
Single event effects (SEE), 258, 263–264, 268–269, 270*f*  
Single event functional interrupt (SEFI), 269  
Single event latchup (SEL), 258, 269, 270*f*  
Single event transient (SET), 269  
Single event upset (SEU), 258, 269  
Single instruction multiple data (SIMD)  
    processor arrays, 169, 170*f*  
Single-photon avalanche diodes (SPADs), 59–60, 153–154, 175, 217, 234, 360, 378–379, 484–485, 484–485*f*  
CMOS miniaturized PMTs, 396–398  
CMOS SPAD plus TDC pixels, column of  
    on-FPGA CMM for, 404–405, 405*f*  
    on-FPGA IEM for, 402–404, 403*f*  
    digital gating pixel of, 391, 392*f*  
    gated analog NMOS-only SPAD pixels, 392–393, 392*f*  
    SPAD-TCSPC arrays, 399–400  
    X-ray sensors, 435  
Single-photon CMOS sensors, 234  
Single-photon counting, 59–61  
Single-photon imaging, 163, 175  
Single-slope ADC (SS ADC), 134, 136*f*  
Smart camera on a chip, 161–164, 162*f*  
    advantages, 164  
    computational chip, 167  
        ACLA, 169  
        ASPA, 169  
        CNNs paradigm, 167, 168*f*  
        on-pixel programmable analog processor, 169, 171*t*  
        on-pixel programmable digital processor, 169, 171*t*  
        PVLSAR, 169  
        SIMD processor arrays, 169, 170*f*  
    drawbacks of, 167  
    high-speed image processing chips, 170–174  
pixel-level processing, 164–165  
silicon artificial retinas  
    arrays of identical pixels, 165–166  
    characteristics, 165  
    Mead’s artificial retina, 166, 166*f*  
    two-dimensional (2D) silicon retina, 166–167  
    VLSI computational sensor, 166–167  
spatial vision chips, 165–166  
spatio-temporal image processing vision chips, 165–166  
ViSoc single chip smart camera, 164  
Smartphones  
    AR/VR viewers, image sensors for, 233  
    Bokeh depth-effect portrait mode, 206–207  
    compressive sampling (CS) methods, 234  
    dual cameras, 207  
    DxO Sensor mark, 208  
    edge-based machine perception, 232–233  
    innovative imaging applications, 187–189, 188*f*  
    multiwavelength sensors, 234  
    NIR sensing, 206  
    optical fingerprint sensing, 230

- Smartphones (*Continued*)  
 optical image stabilization, 232  
 opto-fluidic sensors for biomedical applications, 230  
 proximity sensors, 217  
 RF camera/world-facing camera, 205, 218  
 shipments, 185  
 small pixel, 203–205  
 spectrometer sensing for analysis of materials, 230  
 user-facing video camera, 185
- Smart vision chips. *See* Smart camera on a chip
- SNR. *See* Signal-to-noise ratio (SNR)
- SOC. *See* System-on-a-chip (SOC)
- SocialCamera app, 189
- Society of Motion Pictures Engineers (SMPTE), 451
- SOI. *See* Silicon on insulator (SOI)
- Solar blind UV silicon detectors, 490–491
- Solar Dynamics Observatory (SDO), 281, 283f
- Solid-state imaging detectors  
 fully depleted CCDs, 476–478  
 performance metrics, 475  
 photon absorption and carrier generation, 474–475, 475f  
 photon absorption length, 476–478, 477f  
 quantum efficiency, 475–476  
 quantum yield, 476  
 silicon transmittance, 476, 477f  
 spectral response, 476–478
- Solid-state imaging principle, 18–19, 19f
- Source follower (SF), 48, 129–130, 418, 418f, 457–458, 460t
- Source-follower in-pixel amplifier, 44
- South Atlantic anomaly (SAA), 263
- Space-charge region (SCR), 35–39
- Space imaging, 255  
 altitude and orbit control systems and star trackers, 273–274  
 Earth observation and solar system exploration  
 hyperspectral imagers, 277–278  
 images of Pluto, RALPH instrument, 275, 276f  
 MetImage pixel format, 277f  
 step-and-stare imaging, 274–275, 274–275f
- electro-optical performance  
 dark signal, 262  
 dynamic range, 262  
 measurement repeatability, 262–263  
 noise, 262  
 quantum efficiency, 260–261, 261f
- heritage, 259
- high detection efficiency, 259
- high-performance applications, in astronomy, 278–280  
 Euclid CCD, 280, 281f  
 focal plane of Gaia, 280–281, 282f  
 Sun imaging, 281–282, 283f  
 Teledyne e2v CCD43 assembly, Hubble WFC3, 280, 280f  
 X-ray imaging, 281, 282f
- long-term availability, 259
- packaging, 259, 271–272, 272f, 272t
- power dissipation, 259, 269–270
- radiation hardness, 258  
 displacement damage, 266–268, 267–268f  
 ionizing damage, 264–266, 265f  
 radiation environment, in space, 263–264  
 single event effects, 268–269, 270f
- reliability, 259, 272–273
- scanning imagers  
 hyperspectral imaging, 257, 258f  
 linear sensors, 257  
 time delay integration imaging, 257
- snapshot/staring mode operation, 256, 256f
- system considerations, 259, 271
- SPADs. *See* Single-photon avalanche diodes (SPADs)
- Spatial vision chips, 165–166
- Spatio-temporal image processing vision chips, 165–166
- Star tracker sensors, 273–274
- Stereopsis, 8–9, 207, 212
- Stereoscan, 300–301
- Stereo vision, 338–339
- Still X-ray imaging, 413
- Stitching, 289, 308–309, 424–425
- Storage gates (STG), 388–389, 388f
- Stray light, 336–337
- Streak cameras, 384
- Stroboscopic excitation, FLIM, 385
- Structure light (SL) stereopsis, 212

- Sub-diffraction-limit (SDL) photodetectors, 61
- Successive approximation register ADC (SAR ADC), 134, 135*f*
- Superlattice doping, 487–489, 488–489*f*, 492
- Surface doping methods, 486
- Surface passivation
- dielectric with negative fixed charge, 110
  - highly doped epitaxial-silicon layers, low-temperature growth of, 109–110
  - laser annealing, 109
- Surround view systems, 242–243
- Switch-capacitor-amplifiers (SCAs), 462
- System-on-a-chip (SOC), 161–162, 252
- FLIM
    - CMOS miniaturized PMTs, 396–398
    - confocal scanning FLIM systems, 396
    - frequency-domain lifetime imaging/sensing, 399
    - latest SPAD-TCSPC arrays, 399–400
    - photon-capture devices, comparison of, 398*t*
    - time-domain, embedded algorithms, 400–405
    - wide-field FLIM systems, 395–396
  - mobile sensors, 189, 201–202, 201*f*, 222–223
- T**
- Tasimeter, 19
- TCSPC. *See* Time-correlated single-photon counting (TCSPC)
- TDCs. *See* Time-to-digital converters (TDCs)
- TFT. *See* Thin-film transistors (TFT)
- Thales HD sensor
- dark current distribution, 429, 432*f*
  - DQE curve of, 429, 433*f*
  - MTF curve of, 429, 432*f*
  - parameters of, 428–429, 432*t*
  - SNR in dB *vs.* radiation dose, 429, 433*f*
- Thermal cameras. *See* Infrared (FIR) cameras
- Thermal noise, 45–46
- Thin-film transistors (TFT), 415–418
- Thin tissue autoradiography, 302
- 3D camera
- improved NIR sensitivity, 212–213
  - machine perception and 3D depth imaging, 211–212
- Three-dimensional (3D) imaging, 163
- Three-dimensional television, 467
- Three transistors (3T) pixel, 44, 51–52, 418–419
- dual gain 3T pixel, 421–423, 422*f*
  - noise considerations, in PD 3T pixel, 46, 47*f*, 48
- Through-silicon via (TSV) technology, 111, 198–200, 210
- Tim Cook, 363
- Time-correlated single-photon counting (TCSPC), 59–60, 377–378, 381
- MCP-TCSPC system, 395–396
  - pixels, 393–395
  - reverse start-stop principle, 381, 382*f*
  - SPAD-TCSPC arrays, 399–400
- Time delay integration (TDI), 139–140, 257, 275–276
- Time-gated FLIM, 377–378
- pixels
    - gated analog NMOS-only SPAD pixels, 392–393, 392*f*
    - gated photodiode, 389–390, 390*f*
    - gated pinned-photodiode
      - implementations, 390, 391*f*
      - gating schemes, 389, 389*f*
    - SPAD, digital gating pixel of, 391, 392*f*  - principle, 382–383, 383*f*
- Time-of-flight (ToF) approach, 59–60, 163, 175
- applications
- automotive market, 364–368
  - drones, 371–372
  - emerging and secondary markets, 368
  - IoT/home appliances, 373
  - iPhone X, 360
  - Microsoft Xbox Kinect, 360
  - mobile market, 361–364
  - retail, 372
  - robotics, 370–371
  - virtual reality, 368–369
  - VR/AR segmentation, 369
  - VR players and market potential, 3D
    - players, 370
- cameras, concept and design considerations (*see* Camera)
- CMOS ToF image sensors

Time-of-flight (ToF) approach (*Continued*)

  cw-iToF approach, 342–347  
  demodulation device approaches  
    (see Demodulation)

  light detection and ranging, 319  
  optical ToF measurement, 321  
  radar systems, 319  
  sonar systems, 319  
  speed of light, 319–321  
  3D depth imaging, 211  
  3D TOF

    camera, 2D images of, 333–334  
    range, 321–324  
    R&D, 372  
    robotics, 371  
    “selfie” sensor, 360  
    sensors and mobile volume, 364  
    VR, 370

  triangulation-based approaches, 338–341

Time-resolved fluorescence analysis,

  378–379, 384–385

Time-to-digital converters (TDCs), 60,

  387–388

CMOS SPAD plus TDC pixels, column of  
  on-FPGA CMM for, 404–405, 405/  
  on-FPGA IEM for, 402–404, 403/  
  MegaFrame TDC pixel architectures,

  393–395

Total ionizing dose (TID), 264, 265/  
TowerJazz semiconductors, 305–306

Transfer gates, 348–349, 353

Transimpedance amplifiers (TIA), 387–388

Transmission electron microscopes (TEM),  
  300

Trixell’s PX4700 flat panel sensor, 426–428,  
  427/  
  TroPOMI, 278

TV lines (TVL), 446–447

Two-dimensional (2D) silicon retina,  
  166–167

Two-gate rapid lifetime determination (RLD)  
  method, 389, 401

## U

ULTIMATE sensor, 304–305

Ultrahigh-definition television (UHDTV),  
  467–468

Ultraviolet catastrophe, 13–14

Unmanned aerial vehicles (UAV), 371–372

Up-conversion methods, 384–385

User-facing video camera, smartphones,  
  185

US Jet Propulsion Laboratory (JPL), 32–33

UV-sensitive sensors, 101

## V

Vertical cavity surface-emitting lasers  
  (VCSELs), 322–324, 337

Vertical optical low pass filters  
  (V-OLP), 439

Vertical synchronization signal output  
  (VSYNC), 251

Very-large scale integration (VLSI)  
  electronics, 161

Video graphics array (VGA), 191–193,  
  242

Vignetting, 243

VIPER camera, 437, 452–455

Virtual reality (VR), 368–369

  and AR segmentation  
    console AR, 370–371  
    mobile VR, 370  
    tethered VR, 370  
  players and market potential, 3D for, 370  
  3D for, 370

Vision applications, 241

Vision chips. *See* Smart camera on a chip

Vision systems

  night vision systems, 243–244, 243/  
  rearview systems, 242  
  road crossing monitoring, 244  
  surround view systems, 242–243

ViSoc single chip smart camera, 164

Visual stimulus, 10–11

## W

Wafer-level camera module (WLM)  
  technology, 219–220

Wafer-level optics (WLO), 219–220

Wafer-level packaged (WLP) devices, 111

Wafer thinning, BSI sensors, 105–109

Wafer to wafer bonding technique, 105

Wafer-to-wafer electrical test structure,  
  113–114, 114/  
  Weber-Fechner law, 437–438

White noise, 52

- Wide dynamic range (WDR) CMOS image sensors, 141–142  
combined linear and logarithmic response, pixels with, 145–146, 146*f*  
frequency-based WDR image sensors, 150–151  
integration time control pixels, 147–150  
logarithmic sensors  
    in-pixel calibration circuit, 144, 144*f*  
    nMOS, in weak inversion, 143, 143*f*  
    one pixel circuit, 144, 145*f*  
    two-parametric calibration scheme, 144  
multiple sampling image sensors, 151  
threshold comparison pixels, 147  
well capacity adjustment  
    pixel output relationship, 150, 151*f*  
    stepped reset signal, 150, 150*f*
- Wide field cameras (WFCs), 492–493  
Wide-field FLIM systems, 381, 395–396  
Wien’s displacement law, 13–14
- X**
- X-box Kinect, 211, 216, 360  
Xensium-FT, 463–464, 465*t*  
XMM Newton, 281, 282*f*  
X-ray detection  
    advanced applications, 307–308  
    silicon and non-CMOS detectors, 309–310  
    stitching, 308–309  
X-ray sensor  
    applications and requirements, 428–429,  
        429*t*  
    CIS-based FPD technology  
        advantages, 416  
        binning, 420–421  
        charge amplifier pixel, 421, 422*f*
- Dalsa’s dual gain 3T pixel, 421–423,  
    422*f*  
four transistor (4T) pixel architecture,  
    419–420  
on-chip ADC, 423  
1T pixel architecture on a:Si TFT FPD,  
    417–418  
“photon counting” pixel, 423  
three transistors (3T) pixel architecture,  
    418–419  
visible light CIS pixel architectures, 417  
yield and cost considerations, 423–425
- direct X-ray sensors, 429–434  
fluoroscopy, 413, 416  
intraoral and extraoral dental X-ray  
    imaging, 414–415  
mammography, 413, 416  
medical radiography, 413, 415–416  
parameters  
    NED and DR, 428  
    resolution, MTF, and DQE, 426–428  
    SNR, 425–426
- SPADs, 435  
still imaging, 413  
Thales HD sensor  
    dark current distribution, 429, 432*f*  
    DQE curve of, 429, 433*f*  
    MTF curve of, 429, 432*f*  
    parameters of, 428–429, 432*f*  
    SNR in dB vs. radiation dose, 429,  
        433*f*  
vendors and their parameters, 428–429,  
    430–431*t*
- Z**
- Zone chart, 440, 440*f*, 444–445

This page intentionally left blank

## WOODHEAD PUBLISHING SERIES IN ELECTRONIC AND OPTICAL MATERIALS

*High-Performance Silicon Imaging: Fundamentals and Applications of CMOS and CCD Sensors, Second Edition*, covers the fundamentals of silicon image sensors, addressing existing performance issues and current and emerging solutions. Silicon imaging is a fast-growing area of the semiconductor industry. Its use in cell phone cameras is already well established, with emerging applications including web, security, automotive, space, science, and digital cinema cameras.

This new edition has been revised to reflect the latest state-of-the-art developments in the field, including three-dimensional (3D) imaging, scientific applications, advances in achieving lower signal noise and single-photon capabilities, and new applications for consumer markets. Chapter topics include charge-coupled device (CCD) image sensors, circuits for high-performance complementary metal-oxide-semiconductor (CMOS) image sensors, and CMOS-based optical time-of-flight 3D imaging and ranging.

### KEY FEATURES

- Covers the fundamentals of silicon-based image sensors and technical advances, focusing on performance issues.
- Looks at image sensors in applications, such as mobile phones, scientific imaging, and TV broadcasting, and in automotive, consumer, scientific, and biomedical applications.
- Addresses the theory behind 3D imaging and 3D sensor development, including challenges and opportunities.

This book is an excellent resource for both academics and engineers working in the optics, photonics, semiconductor, and electronics industries.

**Daniel Durini** is currently a full research professor in areas of microelectronics and radiation detection at the National Institute of Astrophysics, Optics, and Electronics (INAOE) in Puebla, Mexico. He served at the Fraunhofer Institute for Microelectronic Circuits and Systems (IMS) in Duisburg, Germany, from 2004 to end of 2013, where he led during the last four years a group dedicated to developing special CMOS process modules for high-performance photodetection devices, pixel structures, and imagers. Prior to his current position, he was with the Central Institute of Engineering, Electronics, and Analytics, ZEA-2—Electronic Systems of the Research Centre Forschungszentrum Jülich in Germany, where he spear-headed, between 2015 and beginning of 2018, the development of Detector Systems dedicated to scientific applications. He has authored and co-authored more than 45 technical papers, three book chapters, and holds six granted patents and two patent applications in the area of CMOS image sensors and radiation detection.



**WP**

WOODHEAD  
PUBLISHING

An imprint of Elsevier  
[elsevier.com/books-and-journals](http://elsevier.com/books-and-journals)

ISBN 978-0-08-102434-8



9 780081 024348