

Ahmed Nabil Belbachir
Editor

Smart Cameras

 Springer

Editor

Ahmed Nabil Belbachir
AIT Austrian Institute of Technology
Donau-City-Str. 1
1220 Wien
Tech Gate Tower, 5. OG
Austria
nabil.belbachir@ait.ac.at

ISBN 978-1-4419-0952-7 e-ISBN 978-1-4419-0953-4
DOI 10.1007/978-1-4419-0953-4
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2009937051

© Springer Science+Business Media, LLC 2010

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Chapter 4

Detectors, Pixels, and Signal Processing

Christoph Posch

4.1 Introduction

Imaging systems comprise the processes involved in the formation of a usually 2-D image from an object or a scene in the real world and in the majority of cases provide a direct point-to-point correspondence between points of the objects in the 3-D world and at the image plane.

Practically all smart cameras and machine vision systems include a sensor that converts electromagnetic radiation into electrical signals and devices which convert these analog signals further into digital signals that can be processed by a computer. The basic functional blocks are similar across most types of machine vision systems:

1. Collection of radiation and projection onto the image plane (e.g., lens, optics)
2. Detector for radiation-to-electrons conversion (e.g., photodiode)
3. Electronics for detector readout and signal conditioning (e.g., APS pixel circuit)
4. Device that implements a procedure to readout the pixel signals (e.g., CCD shift register)
5. Timing and control circuits for the sensor elements and the array (e.g., electronic shutter)
6. Signal processing electronics (e.g., correlated double sampling)
7. Analog-to-digital conversion
8. Interface blocks

After reviewing some basic properties of radiation in the context of imaging, the following section deals with the detector that receives electromagnetic radiation emitted by the observed object or scene and converts its energy into an electrical signal suitable for further processing. In an imaging system, usually a 2-D array of sensors is used for resolving the spatial distribution of the radiation¹. It is often this

¹ One prominent exception is the 1-D array of a line-scan sensor that is used for imaging objects moving perpendicular to the line of sensor elements.

C. Posch (✉)
AIT Austrian Institute of Technology, Austria
e-mail: christoph.posch@arcs.ac.at

distribution, or the relative difference in radiation received by the individual sensor elements in the array, that carries the desired information about the scene.

The unit of a radiation detector plus front-end electronics responsible for the formation of the radiation-dependent electrical signal is called photoreceptor and constitutes the basic sensing element of a vision or image sensor. Several basic types of photoreceptors are in use in today's imagers and vision sensors. Different sensor and readout techniques often require additional circuitry within each of the sensor elements in the array, leading to more complex pixel circuits.

The conditioning or some form of pre-processing of the raw detector signal is frequently being done at the pixel level. The range of pixel-level signal processing is wide – it spans from none, only an integrating photoreceptor, a simple buffer like in standard CCD and CMOS imagers, to complete digital or programmable analog microprocessors in each pixel, found in some advanced vision-systems-on-chip. Also a conversion of the (pre-processed) analog signal to some form of digital representation can be done at the pixel level. The last section of this chapter provides an overview of pixel circuitry for early vision, signal pre-processing, and signal conversion.

4.2 Radiation and Imaging

4.2.1 The Electromagnetic Spectrum

Machine vision is not confined to visible radiation emitted by objects or illuminating scenes. The development of a multitude of imaging detector technologies allows using other types of radiation to probe scenes and objects of interest. The spectrum of electromagnetic radiation relevant for imaging applications spans from the Terahertz band with wavelengths between 1 mm and 100 μm to x-rays with wavelengths of 10 nm and shorter. Even shorter wavelengths – gamma rays – are used in nuclear medicine, high-energy physics, and astronomy applications. Imaging with very long wavelengths of the microwave and radio band is found in imaging radar applications for, e.g., earth observation and again in medicine (magnetic resonance imaging – MRI) and astronomy. In between lies the very narrow band of visible light between 380 and 780 nm and the adjacent infrared (IR) and ultraviolet (UV) bands above and below. The most widely used region of the electromagnetic spectrum for camera and imaging applications including machine vision is the visible light band followed by near and thermal infrared. Figure 4.1 depicts the spectrum of electromagnetic radiation.

Electromagnetic radiation can be treated as *electromagnetic waves* or as a flow of *particles*, constituting discrete packets of energy, called *photons*. A photon travels through space at the speed of light and carries an amount of energy which is proportional to the frequency ν (or anti-proportional to the wavelength λ) of the associated electromagnetic wave:

$$E_{ph} = h\nu = \frac{hc}{\lambda}. \quad (4.1)$$

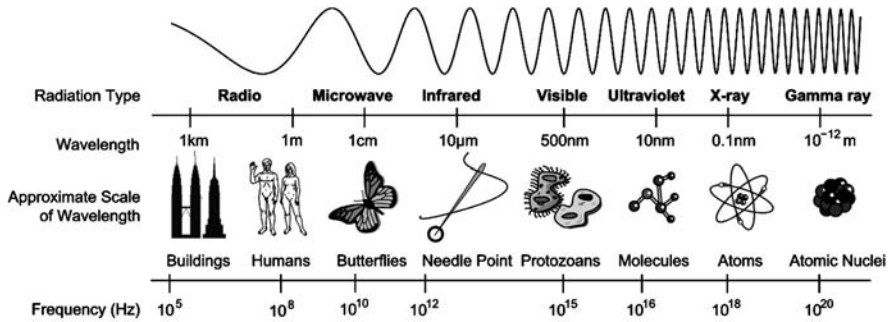


Fig. 4.1 Spectrum of electromagnetic radiation

Planck's constant $h = 6.626 \cdot 10^{-34} \text{ J} \cdot \text{s}$; the speed of light $c = 2.998 \cdot 10^8 \text{ m} \cdot \text{s}^{-1}$. The natural constant c relates the wavelength of an electromagnetic wave to its frequency. The energy of a single photon is commonly given in electron volts (eV), with 1 eV constituting the energy of an electron accelerated by an electric field with a potential difference of 1 V.

Although photons are neutral particles that do not carry electrical charge, using this unit is convenient as electromagnetic radiation is usually detected through the interaction of radiation with electrically charged particles in matter. In solid-state sensors, for example, the energy of impinging photons lifts electrons from the valence band into the conduction band of the semiconductor, allowing them to be turned into an electronic signal that can be read out and processed. The bandgap energy E_g , a material property, defines the minimum photon energy required for this process. Consequently, as a rule-of-thumb, the detector material is sensitive to radiation with photon energies $E_{ph} > E_g$. For example, silicon (Si) has a bandgap energy of 1.1 eV and hence can detect only wavelengths below 1.1 μm. A different semiconductor material, indium antimonide (InSb), with a bandgap energy of only 0.18 eV is sensitive to wavelength below 6.9 μm. As a consequence, InSb is used as detector material for infrared cameras in the 3–5 μm wavelength range, while silicon sensors are commonly used for visible light.

4.2.2 Quantities and Units

The energy of the single photon depends on the wavelength of the associated electromagnetic wave, yet the intensity of light, connected to the perceived brightness, depends on the number of photons per units of time and area. For visible light, the human eye sensation of electromagnetic waves of different wavelengths is called color. Visible chromatic (color) light is sensed by the human eye in the wavelength range between 380 and 780 nm. Light that contains a relatively balanced mix of wavelength in this range is perceived colorless or monochromatic. The only attribute of monochromatic light is the amount of light or the monochromatic intensity. The associated perception is a gray level between white and black.

While brightness is a highly subjective measure of light perception connected to the achromatic notion of intensity, physical attributes of radiation can be measured. Radiant flux is the total amount of energy that flows from a radiation source and is measured in watts (W), and irradiance (W/m²) the energy received per unit area by, e.g., a radiation detector (Table 4.1). By weighting the spectral response of the human eye, the radiometric quantities can be converted to the photometric quantities luminous flux (lm) and illuminance (lm/m² or lux).

Table 4.1 Basic radiometric and photometric quantities

Radiometric quantity	Unit	Definition	Photometric quantity	Unit	Definition
Radiant flux	W	Power (emitted or received)	Luminous flux	lm (lumen)	Luminous power (emitted or received)
Irradiance	W/m ²	Power received at unit area	Illuminance	lm/m ² = lx (lux)	Luminous power received at unit area

For photopic vision , the conversion from a radiometric to a photometric quantity is given by

$$Q_v = 683 \text{ lmW}^{-1} \int_{380 \text{ nm}}^{780 \text{ nm}} Q_{e,v} V(\lambda) d\lambda,$$

(4.2)

with $V(\lambda)$ being the photopic spectral luminous efficiency function of the eye. It accounts for the wavelength-dependent response variation of the eye. The peak of

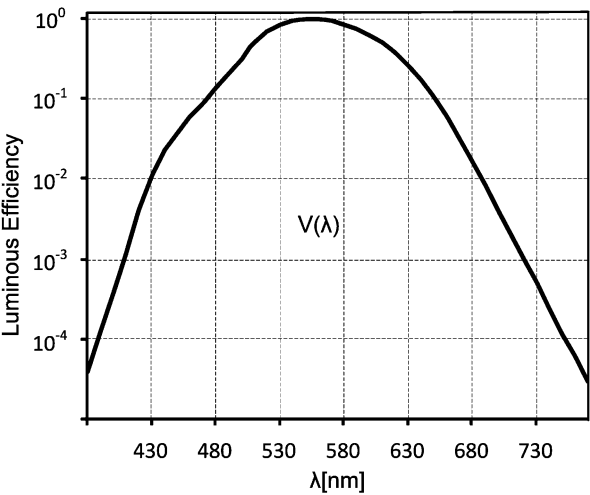


Fig. 4.2 Spectral luminous efficiency function of the “standard” light-adapted eye for photopic vision $V(\lambda)$

the relative spectral luminous efficiency curve for photopic vision lies at 555 nm (Fig. 4.2). The curve is standardized by the International Lighting Commission (CIE); tabulated values of $V(\lambda)$ can be found in [121].

Converting radiant energy into an electrical signal is the task of the radiation detector. The following section outlines the basic properties of imaging detectors relevant for a quantitative radiometric interpretation of images. The focus is on the technologies used in the visible light range.

4.3 Detectors for Imaging

Various technologies for detecting electromagnetic radiation have been the subject of extensive research over the past decades. Almost all detectors fall into one of three fundamental categories, whose distinction is based on the physical process involved in the conversion of radiative energy into an electrical signal.

- **Quantum detectors**

This detector type is based on the quantum photoelectric effect and directly responds to single photons. Each absorbed photon generates free charge carriers that can be read out as an electrical signal. Photon detectors are the most important sensor types for machine vision applications. They cover the entire range of electromagnetic radiation from x-rays, to ultraviolet and visible light, down to the infrared region. Important examples are solid-state sensors.

- **Thermal detectors**

Thermal detectors rely on a detector material with a temperature-dependent electrical property. Absorbed electromagnetic radiation changes the temperature of the detector element and hence the linked electric variable. The signal of a thermal detector is proportional to the instantaneous temperature of the detector element that is permanently trying to establish thermal equilibrium with its environment. The photo sensitivity of thermal detectors is largely independent of the wavelength. Microbolometer arrays for thermal infrared imaging are widely used examples of the thermal detector type.

- **Coherent detectors**

Coherent receivers directly measure the electromagnetic field of the incident radiation by mixing the electromagnetic field of the incoming photons with an internal reference field of similar frequency, produced by a local oscillator. Coherent detectors are used for relatively low-frequency radiation from the radio and microwave band to sub-millimeter waves.

In the next section, we will focus on the quantum detector, the most common detector type used in machine vision systems for the infrared, visible light, and ultraviolet bands of the electromagnetic spectrum and cover function and applicability with regard to the underlying physical mechanisms and radiometric properties.

4.4 Semiconductor Photodetectors

Quantum photodetectors based on semiconductor materials are the most widely used detector type for wavelength of the infrared, visible light, and ultraviolet bands. The photodetector turns an optical signal, received in the form of electromagnetic radiation or photons, into a preliminary electrical signal in the form of free charge carriers. This operation comprises the basic functions of (1) generation of free charge carrier pairs in response to incident light and (2) separation and collection of electrons and holes. The generation of an electrical output signal, that can be transmitted and further processed, requires the interaction with other electrical components.

4.4.1 Introduction to Solid-State Photosensing

Without going into detail about semiconductor physics, the fundamentals of solid-state photo-transduction are briefly reviewed in this section and some important terms and expressions are introduced and defined. A classic and very comprehensive reference on this subject is [517]. Further excellent references are available [597, 500].

Quantum mechanics shows that electrons bound to an isolated atomic nucleus can occupy only discrete energy levels with forbidden gaps in between. When many atoms form a crystal, the discrete levels are widened to form continuous energy bands, again with forbidden gaps between the bands. At a temperature of 0 K, the lowest energy band that is not fully occupied and all higher bands are called conduction bands. All lower bands, completely filled with electrons, are called valence bands. In electrical conductors, the lowest conduction band is partially occupied, while all conduction bands are empty for insulators and semiconductors at 0 K.

The separation between the highest valence band and the lowest conduction band is the so-called bandgap energy E_g which plays an important role in the classification of materials and their aptitude for being used as radiation detectors. E_g is crucially larger in electrical insulators than in semiconductor materials.

In electrical conductors, negatively charged electrons are the only charge carriers available. In a semiconductor crystal, when its temperature rises above 0 K, electrons in the valence band can acquire enough energy to leave the valence band and jump into the conduction band. The freed position in the valence band can now be considered to be positively charged. These states are called holes. Both the conduction band electrons and the valence band holes contribute to the electrical conductivity and hence are subject to charge transport in the semiconductor. The holes do not actually move. However, moving electron can fill a hole, leaving behind a hole at the place in the crystal lattice it moved away from. Thus the holes change their place and appear as if they were actual positively charged moving particles.

When an electric field $\mathbf{E}$ is applied to a semiconductor, the electrically charged electrons feel a force equal to $|q\mathbf{E}|$ where q is the elementary charge. The average velocity of carriers in a semiconductor due to the applied electric field $\mathbf{E}$ depends on

the carrier mobility μ , a material constant, and the electric field: $|\mathbf{v}| = \mu |\mathbf{E}|$. The mobility of holes μ_p is usually smaller than the mobility of electrons μ_n .

A semiconductor crystal without impurities or defects is called an intrinsic semiconductor. As described above, in intrinsic semiconductors charge carrier pairs are generated through thermal excitation. Hence the number of available mobile charge carriers depends on the temperature. The number of holes and electrons is equal. In silicon (Si), for example, the intrinsic carrier concentration at 300 K is $1.45 \cdot 10^{10} \text{ cm}^{-3}$.

The conductivity of semiconductors can be modified by introducing impurities into their crystal lattice. The process of controlled adding of impurities to a semiconductor is called doping. Doped semiconductors are commonly referred to as extrinsic semiconductors. By adding impurity to pure semiconductors, the electrical conductivity may be enhanced not only quantitatively, but also by the type of impurity atom, thus increasing the number of either electrons or holes. Hence, in extrinsic semiconductors the numbers of electrons and holes are no longer equal. As a result, two types of doped semiconductors can be fabricated, namely n-type with electrons as the abundant carrier and p-type with an abundance of holes. The predominant carrier in a semiconductor material is called majority carrier, the other minority carrier. Semiconductors of both types can be combined to achieve certain functions or characteristics. The simplest example is the p-type/n-type junction of a diode.

Besides through thermal energy, electrons can also be excited into the conduction band by energetic particles, e.g., photons, absorbed in the semiconductor material. This process is known as induced electron-hole pair generation. The minimum energy a photon must carry to trigger such a band transition is equal to the bandgap energy E_g . In the absence of an electric field to separate the charge carriers, electron-hole pairs recombine after an average period of time, called carrier lifetime, and emit the previously absorbed energy in the form of light or heat. In order to detect and make use of an optical signal, the photon-generated free charge carriers must be prevented from recombining and be collected. Succeeding in doing so allows the charge carriers generated by impinging radiation to be turned into an electrical signal, hence the semiconductor crystal to work as a radiation detector.

4.4.2 Figures of Merit of Photodetectors

Common parameters to quantify the operation of photodetectors are *responsivity*, *quantum efficiency*, *noise-equivalent power* (NEP), and *detectivity*.

Responsivity denotes the input-output relationship or gain of a detector. The responsivity R_{ph} of a photodetector is the ratio of light-induced photocurrent density J_{ph} and the incident optical power per unit area P_o :

$$R_{ph} = \frac{J_{ph}}{P_o} \left[\frac{\text{A}}{\text{W}} \right]. \quad (4.3)$$

Quantum efficiency η is defined as the number of photon-generated free charge carriers per impinging photon usually given in percent. The number of charge carriers is the photocurrent divided by the electron charge J_{ph}/q and the number of photons is the incident optical power divided by the energy of one photon $P_o/h\nu$:

$$\eta = \frac{\frac{J_{ph}}{q}}{\frac{P_o}{h\nu}} = R_{ph} \frac{h\nu}{q}. \quad (4.4)$$

For a given quantum efficiency, R_{ph} increases linearly with the wavelength until the cutoff wavelength at the bandgap energy of the material is reached, e.g., 1.12 μm for Si, and R_{ph} drops to zero.

The lower limit of light detection is determined by the noise characteristics of the device. The photodetector noise i_n is the sum of thermal noise and shot noise from dark current and photocurrent,

$$i_n = \sqrt{i_{th}^2 + i_{sd}^2 + i_{sph}^2} \left[\frac{\text{A}}{\sqrt{\text{Hz}}} \right], \quad (4.5)$$

with the thermal noise of an equivalent shunt resistance $i_{th} = \sqrt{\frac{4kTB}{R_{sh}}}$ and current shot noise $i_s = \sqrt{2qI_sB}$ of dark current and photocurrent and B the noise bandwidth. The noise-equivalent power (NEP) is defined as the intensity of incident light required to generate a current equal to the noise current – or yielding a signal-to-noise ratio (SNR) of unity in a 1 Hz bandwidth:

$$NEP = \frac{I_n}{R_{ph}} [\text{W}]. \quad (4.6)$$

The detectivity D^* is defined as the ratio of detector area A and bandwidth B to the NEP:

$$D^* = \frac{\sqrt{AB}}{NEP} \left[\text{cm} \cdot \frac{\sqrt{\text{Hz}}}{\text{W}} \right]. \quad (4.7)$$

The response time is another important performance parameter for photodetectors particularly when used in high-speed optical communications systems. Refer to [517] for a comprehensive study on semiconductor photodetectors.

4.4.3 Types of Semiconductor Photodetectors

4.4.3.1 Photoconductor

Photoconductors respond to incident light by changing resistance/conductance of the detector material. Intrinsic photoconductors are the simplest way to build a solid

state photodetector, consisting of just a piece of semiconductor with ohmic contacts at both ends (Fig. 4.3a). At normal temperatures, relatively few electrons occupy the conduction band of the semiconductor, resulting in a low electric conductivity of the material. Impinging photons generate charge carriers that increase the conductivity. The current flowing through the device in response to an applied external voltage is dependent on the incident optical power. The minimum photon energy that can be detected is determined by the bandgap energy of the detector material. Intrinsic photoconductor detectors exhibit good uniformity and quantum efficiencies typically of the order of 60%. The main shortcomings of photoconductors are high noise and slow response speed. Photoconductors are not widely used in image sensors.

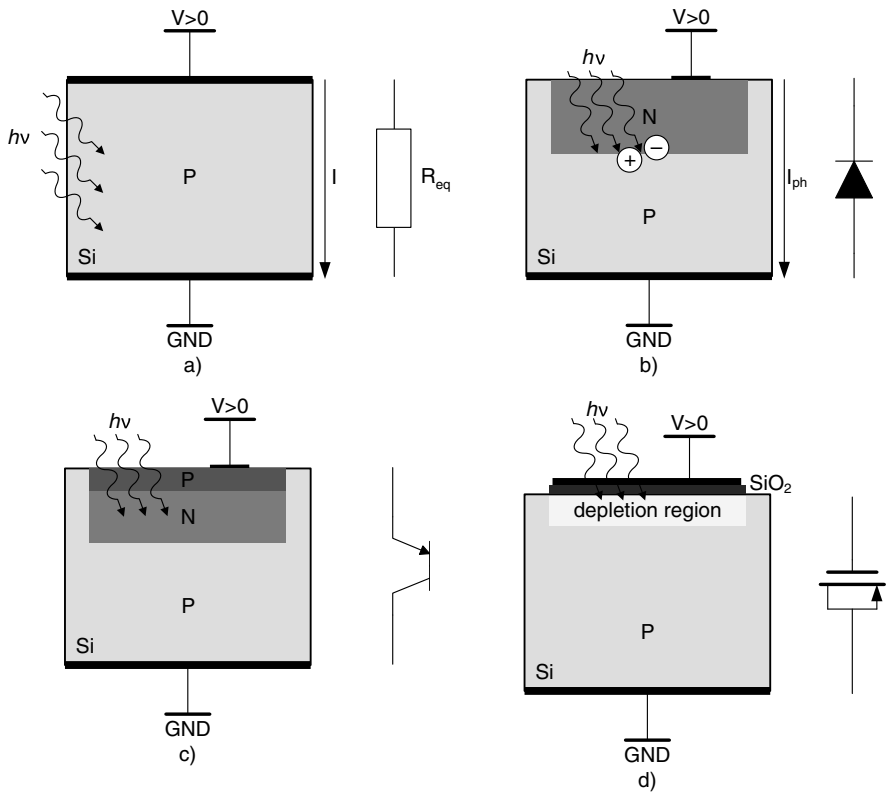


Fig. 4.3 Semiconductor device cross sections of (a) photoconductor, (b) photodiode, (c) (vertical p-n-p) phototransistor, and (d) photogate

4.4.3.2 Photodiode

The photodiode (Fig. 4.3b) is the most widely used detector type in CMOS image sensors for visible light and adjacent NIR and UV. A photodiode is based on a junction of p-type and n-type semiconductor or a p-type – insulator – n-type structure (PIN diode). Close to the contact region the material is depleted of free

charge carriers, forming a depletion layer with an electric field that separates photogenerated electron–hole pairs. For fast transit times needed for high-speed operation, the depletion region should be made narrow, while in order to maximize quantum efficiency, the depletion layer should be as thick as possible to allow a large fraction of the incident photons to be absorbed. This leads to a design trade-off between speed and quantum efficiency.

The photodiode can be operated in three modes of biasing: photovoltaic (solar-cell mode), photoconductive (photodiode mode), and avalanche mode (Fig. 4.4). As a detector for imaging, the device is normally operated in photodiode mode and is reverse biased. The current flowing through the photodiode is the sum of diffusion (dark) current I_{dark} and photocurrent I_{ph} . The photocurrent has a linear relation to the incident light intensity. In the near infrared (NIR) region between 0.8 and 0.9 μm , silicon photodiodes with antireflective coating can reach quantum efficiencies of close to 100% [517].

In photovoltaic or solar-cell mode, no external voltage is applied to the diode. When photons of energies greater than the bandgap energy are absorbed in or close to a p–n junction of a photodiode, the resulting electron/hole pairs are pulled by the electric field of the contact potential across the p–n junction. Electrons are swept from the p-region into the n-region, and holes in the opposite direction. As the charge carriers are spatially separated across the detector, a resulting voltage appears. If the n- and the p-type regions are electrically connected, a current can flow between both regions, giving rise to the so-called photovoltaic effect.

In avalanche mode, photodiodes show a current gain effect as a result of impact ionization under avalanche breakdown conditions, i.e., each absorbed photon leads to a multitude of photogenerated charge carriers. Avalanche PDs are used, e.g., as particle detectors in high-energy physics experiments and related imaging applications like positron emission tomography (PET).

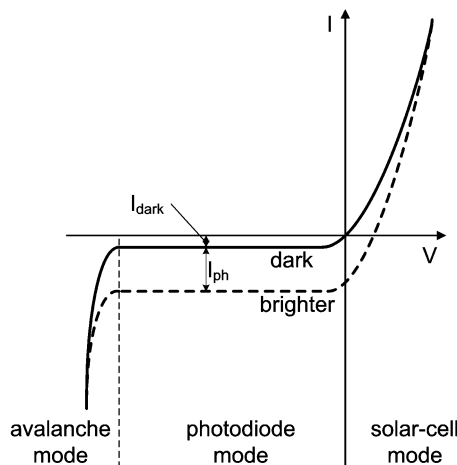


Fig. 4.4 Photodiode I–V characteristics

4.4.3.3 Phototransistor

A phototransistor is usually a vertical or lateral n-p-n or p-n-p bipolar transistor structure whose base-collector junction is exposed to light (Fig. 4.3c). It can be viewed as a photodiode with built-in amplifier and offers current gain >1 . Phototransistors suffer from poor high-frequency performance and are rarely used in solid-state imagers.

4.4.3.4 Photogate

A photogate is a MOS capacitor which converts impinging photons into stored charge rather than into photocurrent or voltage signals (Fig. 4.3d). The capacitor is charged in proportion to the light intensity to which it is exposed by accumulating photogenerated carriers in the depletion region under the gate when the device is biased accordingly. Unlike in most other photodetectors, there is no external photocurrent flowing during light exposure. Photogates are the basic element in a CCD image sensor, which is described in the following section.

4.5 Basic Pixel Structures

4.5.1 Charge-Coupled Device (CCD)

A CCD pixel is based on the photogate detector, a MOS capacitor with a polysilicon gate as the top terminal. In an imager, the CCD pixels are placed close to each other in rows that can act as analog shift register. During the exposure interval, an electrical potential applied to the gate electrode holds the accumulating, photo-generated charges in the potential well underneath the gate. Typical exposure times range from 100 μ s to 100 ms.

The accumulated charge can be displaced to a neighboring well by applying a synchronized sequence of voltage signals to the two gates. Figure 4.5 illustrates the transfer of charge in a CCD device. To initiate the charge transfer, the voltage applied to the gate G2 is quickly raised. Subsequently the voltage at neighboring gate G1 decreases slowly. Thereby, the charge stored under gate G1 during the exposure phase is transferred to the potential well under gate G2. By repeating this process, the accumulated photo charge of each pixel eventually reaches the end of the row where it gets processed by a charge-sensitive amplifier and is turned into a voltage or current signal. Transfer efficiencies between two wells of 99.99% can be achieved by using fringing fields and clock frequencies of several tens of megahertz [517].

4.5.2 CMOS Photoreceptors

The unit of optical detector plus front-end electronics in CMOS pixels is called photoreceptor. In this section, photoreceptor circuits based on photodiodes that

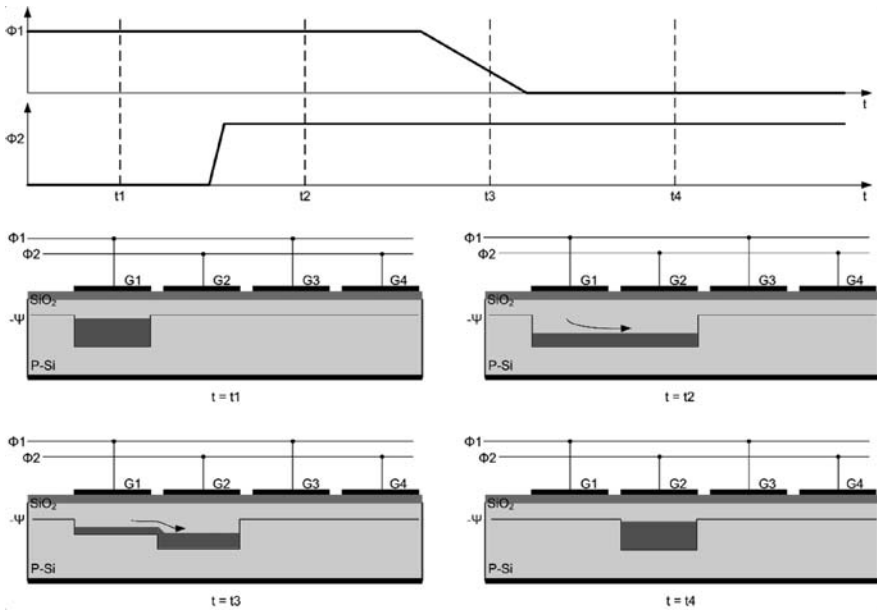


Fig. 4.5 Illustration of one CCD charge transfer cycle. Shown are the clock phases ϕ_1 , ϕ_2 applied to the gates G1–G4 and the surface potential ψ with charge packets moving between the changing potential wells as response to the clock signals

transduce the optical input into electrical charge, current, or voltage are investigated and compared. The most fundamental classification of photoreceptor circuits is based on their mode of operation, integrating/sampled or continuous time.

Sampling pixels integrate the induced photocurrent from the detector onto a capacitor during a given integration period. Often the parasitic capacitance of the photodiode itself is used as the integration capacitor. Before each integration cycle, the photodiode is reset and the capacitance is pre-charged to a reference value. The voltage across the capacitance at a given point in time after the reset constitutes the output signal and carries information of the incident light during this time. Continuous-time photoreceptors transduce the continuous photocurrent directly into an output voltage or current for direct processing. Many “smart” vision sensors use continuous-time photoreceptors to perform early processing of the visual information in the temporal and/or spatial domain.

Three elementary photoreceptor types, charge mode, current mode, and voltage mode, can be distinguished. The charge mode circuit is the basic structure of integrating/sampled pixels whereas current mode and voltage mode circuits constitute the basic continuous-time photoreceptors types. In [578], a detailed comparison of the elementary photoreceptor types based on achievable information rate is shown.

The integrating (charge mode) photoreceptor (Fig. 4.6a) consists of a photodiode and a reset transistor. Complete with a row select switch transistor and a source follower buffer, this structure forms the basic CMOS APS pixel which can be found in

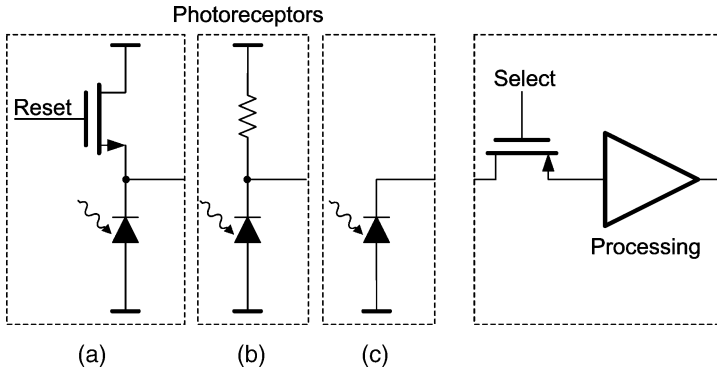


Fig. 4.6 Basic CMOS photoreceptors (a) charge mode (integrating), (b) voltage mode (continuous), and (c) current mode (continuous)

most mainstream CMOS image sensors. The linear voltage mode circuit (Fig. 4.6b) converts the photocurrent into an output voltage based on an ohmic resistance. By using a load element with a logarithmic I–V characteristic, logarithmic encoding of the illumination information can be achieved. This is useful, e.g., to increase the dynamic range of the sensor. Continuous-time voltage mode photoreceptors are used in many “smart” vision and image sensors and will be covered later in this section. In the current mode photoreceptor (Fig. 4.6c) the photocurrent is directly routed to a processing circuit. This scheme has certain advantages in relation to pixel-level signal processing and will be discussed in Section 4.6.1.

The remainder of this section deals with CMOS image sensors and concludes with a comparison of CCD and CMOS as basic technologies for image sensing. Overviews of CMOS image sensors and applications can be found in [188, 62, 585, 414].

4.5.3 CMOS Passive Pixel Sensor (PPS)

The photodiode-type integrating passive pixel approach, first suggested by Weckler in 1967 [558], is shown in Fig. 4.7. The pixel circuit consists only of a photodiode and an access transistor, the operation is similar to dynamic random access memory (DRAM). When the access transistor is activated, the photodiode is connected via a vertical column bus to a charge-integrating amplifier (CIA). With the access transistor off, the photodiode discharges at a rate proportional to the incident light. When the photodiode is accessed, the voltage on the photodiode is reset to the column bus voltage, and the charge, proportional to the optical signal, is converted to a voltage by the CIA.

The single-transistor photodiode passive pixel allows the highest fill factor, the ratio of photodiode area to total pixel area, of up to 90%. On the other hand, PPS sensors suffer from low sensitivity and readout noise due to the large bus capacitances which are directly connected to the pixel circuit during readout and the

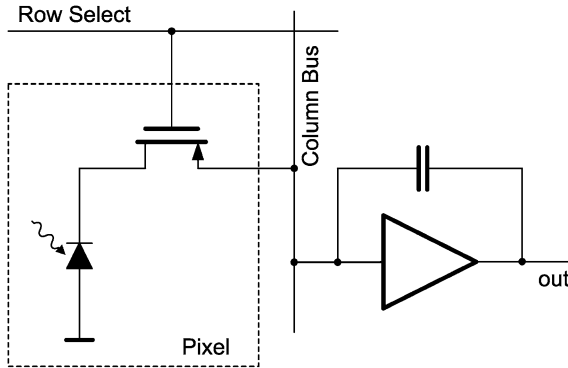


Fig. 4.7 Basic pixel circuit of a CMOS PPS with column charge amplifier readout

consequential limited scalability to larger array sizes or high frame rates. PPS arrays are prone to column fixed pattern noise (FPN) due to column amplifier mismatch.

4.5.4 CMOS Active Pixel Sensor (APS)

First experiences with the CMOS PPS led to the finding that a buffer/amplifier could potentially improve the performance of the pixel [403]. A sensor with an active amplifier within each pixel is referred to as an active pixel sensor or APS. The CMOS APS trades pixel fill factor for improved performance as compared to passive pixels due to the pixel-level amplifier. Current APS pixels typically reach fill factors of 50–70%. Loss in optical signal is more than compensated by a reduction in read noise, leading to an increase in signal-to-noise ratio (SNR) and dynamic range.

The applied active circuit is usually a simple source follower that acts as both an amplifier and a buffer that isolates the photodetector charge from the large capacitances of the readout buses. Power dissipation is kept low because each amplifier is only activated during readout. APS pixels mostly use photodiodes or photogates as detectors and operate in integration mode. While a PPS directly transfers the accumulated signal charges to the outside, an APS transduces the charges to a voltage at the pixel level. Conventional APS's suffer from a high level of fixed pattern noise (FPN).

4.5.4.1 Photodiode (PD) APS

The photodiode-type APS was first described by Noble in 1968 [403] and further investigated by Andoh [17] since the late 1980s. In its fundamental form, the pixel contains three transistors: a photodiode reset transistor (MR) for resetting the photodiode voltage, a source follower transistor (SF) for buffering the photodiode voltage to a column bus, and a row select transistor (MS) (Fig. 4.8). The conversion of signal charge to voltage takes place at the photodiode node capacitance which is the sum

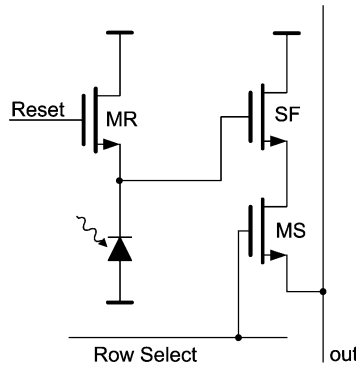


Fig. 4.8 CMOS photodiode APS schematic

of device capacitances of the diode and connected transistors. The load of the source follower is connected to the column bus and is common to all pixels in the column.

During the reset phase, the reset transistor is on and the photodiode is connected to a reference voltage. The photodiode node capacitance is charged accordingly. During integration, the reset transistor is switched off and the capacitance is discharged by the incident light proportional photocurrent. At the end of the desired exposure time, the resulting voltage is read out. The output voltage is higher for darker and lower for more brightly illuminated pixels.

The SNR performance (compare Section 4.4.2) decreases for smaller pixel sizes since photocurrent shot noise and reset voltage noise, also known as kTC noise, scale as $1/\sqrt{C_{PD}}$, where C_{PD} is the photodiode capacitance. The dynamic range (DR) of an integrating APS pixel is defined as the ratio of the maximum output voltage V_{max} and the minimum detectable signal V_{min} (compare Section 4.4.2) given in dB where V_{max} is limited by the full-well capacity, the number of charges that can be accumulated in the PD. The conversion gain is defined as the voltage change due to one accumulated elementary charge (V/e^-) and equals $1/C_{PD}$.

Consequently, in minimizing chip area, design trade-offs can be made between pixel fill factor (PD area), dynamic range (full-well capacity), and conversion gain.

4.5.4.2 Pinned Photodiode (PPD) APS

In order to decouple photodetection and photoconversion areas and resolve the trade-off of the PD APS, the “Pinned Photodiode” APS (PPD APS) has been developed. The accumulated photocharge is transferred to a floating diffusion (FD) via a transfer gate (TX) (Fig. 4.9). The conversion of charge to voltage is done at the FD capacitance. The circuit topology allows a straightforward correlated double sampling (CDS) implementation for noise suppression. The PPD APS is superior to the standard PD APS in terms of noise performance; however, due to the additional transistor, the achievable fill factor is smaller and the charge capacity is limited. PPD APS can compete with CCD imagers in terms of image quality. An important

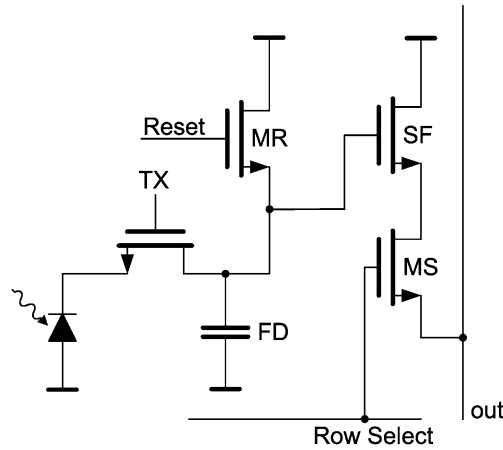


Fig. 4.9 Pinned photodiode APS schematic

drawback of the pinned photodiode is that it cannot be fabricated in a standard CMOS process [339, 415].

4.5.4.3 Photogate (PG) APS

The photogate APS was introduced in 1993 [373, 372, 374]. It combines the CCD principle of operation of integration and charge transport in the pixel with random access readout. Initially the signal charge is accumulated under a photogate. For readout, a floating diffusion is reset and the resulting voltage measured. The signal charge is then transferred to the output diffusion by pulsing the photogate accordingly. Again the voltage is measured and the difference between the reset level and the signal level is read out, realizing true correlated double sampling (CDS) with known benefits of suppressing reset kTC noise and offset FPN. The PG APS uses five transistors per pixel (Fig. 4.10). The improvement in noise performance is partially outweighed by a reduction in quantum efficiency due to the optical properties of the polysilicon gate. The photogate pixel can be operated at higher charge capacity as compared to the pinned photodiode. PG APS is suitable for high-performance and low-light applications but requires a specialized fabrication process.

4.5.4.4 Logarithmic PD APS

Increasing the dynamic range of CMOS image sensors (typically 60–70 dB) toward the dynamic range of natural scenes of >120 dB has been the subject of extensive research over the past decade. A logarithmic transformation of the photo-signal (the pixel output signal is proportional to the logarithm of the photo-signal) allows for an increase in intra-scene dynamic range as the photo-signal is companded. A MOS transistor that operates in the sub-threshold region, e.g., shows a logarithmic I–V characteristic. This property can be exploited to achieve logarithmic transformation

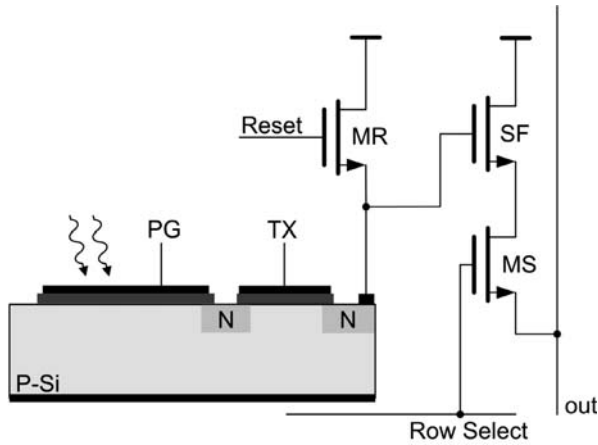


Fig. 4.10 Photogate APS schematic

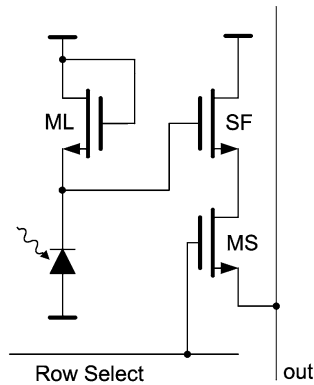


Fig. 4.11 Logarithmic photodiode APS schematic

of the photo-signal. Figure 4.11 shows an example circuit which exhibits the logarithmic characteristic for typical light levels and photocurrents in the pA to nA range where the load transistor ML operates in sub-threshold. The circuit is of the continuous-time (non-integrating) type.

Drawbacks of the continuous-time approach include slow response time for low light levels and a small voltage swings; however, the main obstacle faced when using continuous-time logarithmic photoreceptors has been the large FPN. Various pixel-level and chip-level calibration techniques and approaches employing combinations of linear and logarithmic pixel response have shown promising results. Intra-scene dynamic ranges of more than six decades have been achieved with logarithmic PD APS at FPN levels of a few percent [283, 466, 333, 112].

Continuous-time logarithmic PD photodetectors can be operated in a high-gain feedback configuration to reduce charging/discharging times of parasitic and

junction capacitances. Bandwidth and speed of the photodetector are greatly enhanced [331]. Figure 4.12 shows a logarithmic photoreceptor circuit based on a common-source gain stage (M1, M2) feedback loop. This type of photodetector is used in smart sensors with pixel-level analog signal processing and will reappear in Section 4.6.5.

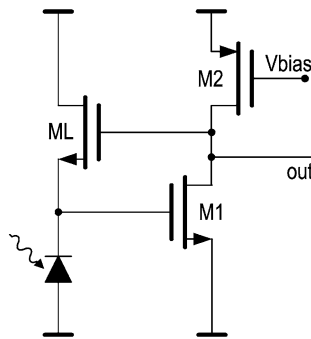


Fig. 4.12 Logarithmic photoreceptor with common-source feedback

4.5.5 CMOS Versus CCD

CMOS and CCD image sensors are both made from silicon and convert incident light into electrical charge based on the similar physical processes. Both technologies support photogates and photodiodes as detector elements and have similar fundamental properties of sensitivity in the visible and NIR ranges of the electromagnetic spectrum.

CCDs provide excellent image quality at excellent noise performance. Although it is possible, it is in fact not economical to integrate other camera functions like clock drivers, logic circuits for timing control, signal processing blocks, on the CCD chip. Most CCD camera systems contain several chips, at least three or more. The amplitude and shape of the various clock signals controlling the charge transfer in CCDs are critical. Generating correctly sized and shaped clocks is usually done in a specialized clock driver chip. Multiple, uncommon supply voltages and high power consumption are standard in most CCD systems.

CMOS imagers sense light in the same way as CCD but this is where the similarities end. The accumulated charge packets are not transferred, but converted as early as possible by charge-sensitive circuits. CMOS sensors can be manufactured on standard fabrication processes. The main problem with CMOS sensors is the matching of the individual amplifiers in the pixels leading to significant fixed pattern noise (FPN); however, the issue has been resolved to a great extent by using on-chip correction techniques. Higher dark currents contribute to the inferior noise performance. A major benefit of CMOS cameras over CCD lies in the possibility of implementing virtually all of the electronic camera functions onto the sensor chip. It is common to integrate the A/D conversion but also signal processing, timing

logic, exposure control, etc., on-chip resulting in complete single-chip cameras. The major benefits are one-rail supply requirements and significant power savings over CCD. Competitiveness and abundance of CMOS technologies combined with the high level of camera integration entails substantial cost advantages over CCD. The main differences and common features of CCDs and CMOS image sensors are summarized in Table 4.2.

Table 4.2 Comparison between CCD and CMOS sensors	
CCD	CMOS
Excellent image quality	Medium to high image quality
High SNR (low noise <40 electrons)	Lower SNR (>20 noise electrons)
Low FPN (dark non-uniformity <1%)	Higher FPN (requires on-chip correction)
Low dark current (<10 pA/cm ²)	Higher dark current (nA/cm ²)
Rare technology	Mainstream technology
No on-chip ADC, signal processing, control	On-chip device integration – smart sensors
Serial scanning readout	Random access readout
Complicated driving and interfacing	Simple data interface
Multiple high-voltage supplies	Single supply operation
Higher system cost	Lower system cost
Common characteristics:	
Spectral response (400–1000 nm)	
Minimal pixel size (2–5 μm)	
Charge storage per unit area	
Chip size and number of pixels	

4.6 Pixel-Level Signal Processing

“Smart” vision sensors, as opposed to standard image sensors like CMOS APS or CCD as covered in the preceding sections, perform processing tasks on the optical information early in the signal chain, typically at the pixel level. Other synonyms for this type of sensor technology/signal processing are “early vision,” for its relation to biological paradigms, or “focal-plane processing” (FPP).

This section covers some selected circuitry for early vision and signal pre-processing found in the pixels of various “smart” vision sensors. The range of pixel-level signal processing is wide – it spans from simple amplifiers to complete programmable digital or analog microprocessors in each pixel. Two fundamental classes of pixel-level signal processing can be distinguished – intrapixel, where the processing is performed on the individual pixel signals, and interpixel, where the processing is performed locally or globally on signals from several pixels. An excellent overview of this topic has been published by Moini [385] and covers all significant developments up to 1997 in great detail. Reference [159] provides an introductory overview of the topic. Later developments are covered in [408].

Pixel-level processing implies that photo-induced signals from non-integrating photoreceptors can be directly processed in continuous time, and photocurrents and

derived voltages from integrating photoreceptors are available for processing during integration. Many pixel-level processing efforts, mainly of the interpixel type, aim to improve the quality of an image sensor by enhancing noise performance, SNR, dynamic range, or speed. The transition from PPS to APS by adding a source follower to the pixel circuit is a simple but prominent example.

But many more useful functions can be implemented or supported at the pixel level such as electronic shuttering [579], correlated double sampling (CDS) [372], and frame differencing [145].

Interpixel processing opens the arena for computational vision sensors that are, e.g., able to extract information or features from the scene and perform tasks like motion detection, edge detection, optical flow, or segmentation [385, 408, 159]. Massively parallel pixel-level processing allows reducing the required speed of the single processing element to the point where analog circuits operating in the sub-threshold domain can be used to perform complex computations while consuming very little power [369]. This paradigm gives rise to the development of continuous-time dynamic vision sensors that, like their biological role models, use asynchronous pulses (spikes) to transmit information about scene dynamics – changes in the scene – with very high temporal resolution and low communication bandwidth requirements, e.g., [317].

Illumination can be encoded in time or frequency by controlling the timing of pulse signals proportional to the pixel's illumination. Pulse frequency or time between pulses can be converted to digital data at the pixel level or can be communicated off-chip using synchronous or asynchronous digital buses.

Current mode pixels allow straightforward implementations of arithmetic functions like additions or multiplications, hence simple to medium-complex image processing tasks are to be performed directly in the sensor.

In this section, several approaches that implement basic vision tasks or some form of processing of the visual information at the pixel level are presented. This section does not claim completeness for the great number of different designs that have been implemented. Some of the circuits presented have only recently been implemented in smart camera or imager products for the first time or are about to be.

4.6.1 Current Mode Pixel and Analog Focal-Plane Processing

Current mode imaging provides an attractive alternative to the common voltage mode APS if focal-plane image processing is desired. In current mode circuits, basic arithmetic functions like addition, subtraction, or multiplication can be implemented in a straightforward way, but also integration and non-linear transformations are available in the current mode domain. Analog memories based on current copier circuits, current scaling circuits, and current mirrors with programmable mirror ratios have been demonstrated [450, 206]. In the context of image sensor arrays, these paradigms can be applied to continuous-time or integrative photocurrents to implement various intrapixel and interpixel signal processing tasks.

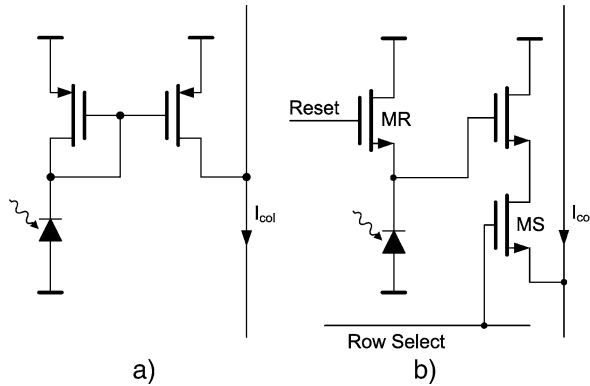


Fig. 4.13 Continuous-time (a) and sampled (b) current mode pixel readout circuits

The basic circuit of a continuous-time current mode pixel with current mirror readout is shown in Fig. 4.13a. Non-unity current mirror ratios can be used to amplify the photocurrent. Like other continuous-time photo-circuits, this architecture suffers from low sensitivity at low light, non-linearities, and high fixed pattern noise. Figure 4.13b shows the schematic of an integrating PD current mode pixel. Current mode imagers are still far less common than voltage mode CMOS APS sensors due to their generally lower image quality [368]. However, methods like multiple sampling or velocity-saturated readout for reducing FPN have shown promising results [395, 427, 220, 585, 586]. A low-power technique for mega-pixel current-mediated imagers [368] is presented in [70].

In the field of focal-plane image processors and hardware array implementations of cellular neural networks, current mode pixel circuits and signal processing are omnipresent. An in-depth coverage of current mode techniques suitable for focal-plane processing can be found in [450] and their application for image sensing and processing in [165]. In [202], basic current mode early vision in a 256×256 pixel array is presented, and in [148], an analog array processor with binary imaging and programmable inter pixel processing for simple vision tasks is described. A biology-inspired analog parallel array processor based on current mode techniques is shown in [94]. Complex spatiotemporal processing has been realized in various designs in the past years [101, 169, 219, 206, 360, 359, 391].

The SCAMP family of software-programmable single-instruction-multiple-data (SIMD) processor arrays for focal-plane low-level image processing applications, based on current mode analog circuit techniques is presented in [151, 152]. The ACE16k [323, 451], the third generation in a family of complex vision-systems-on-chip [164, 163, 148], contains an array of 128×128 mixed-signal SIMD processing elements with configurable optical sensors which is able to perform various early vision tasks at a speed corresponding to 300 giga operations per second (compare Table 4.3).

Table 4.3 Comparison of vision chips with programmable analog and/or digital processing elements in each pixel [335]

Vision chip	PVLSAR 2,2 [413]	NSIP [158]	SPE [300]	ACE16k [323]	SCAMP [152]	ASPA [335]
Array size	128 × 128	32 × 32	64 × 64	128 × 128	128 × 128	128 × 128
Cell size, μm ²	60 × 60	118 × 118	67.4 × 67.4	75.7 × 77.3	49.4 × 49.4	100 × 117
Technology, μm	0.8	0.8	0.35	0.35	0.35	0.35
Processing type	Digital	Digital	Digital	Analog	Analog	Digital
Memory per cell	5 bit	8 bit	24 bit	8 analog registers, 4 bit binary	9 analog registers	64 bit
Performance GOPs	49	1.9	6.4	330	20	157
Die size, mm ²	76	25	29.2	145.2	50	213.5
Power per chip	1 W	0.1 W	N/A	2.9 W	240 mW	5.4 W
Perf/area MOPS/mm ²	92.6	1.1	343	3800	512	820.8
Perf/energy GOPs/W	5.4	25.2	N/A	180	85.3	29

4.6.2 Time-Based (PM) Pixel

In time-based or pulse modulation (PM) imaging the incident light intensity is not encoded in amounts of charge, voltage, or current but in the timing of pulses or pulse edges.

Two basic approaches to time-based imaging can be discerned. The first technique is called pulse width modulation (PWM) or synonymously time-to-threshold and time-to-first-spike (TTFS) imaging. Instead of integrating photogenerated current for a fixed period of (exposure) time, the integration is stopped when the pixel signal reaches a certain value. In the basic PWM implementation, the pixel voltage is reset to a fixed DC voltage and then the photodiode capacitance is discharged by the photogenerated current. A comparator compares the pixel output voltage to a voltage reference. The comparator output switches when the threshold is reached. The time between reset and the switching of the comparator carries the information on the average pixel illumination during that period. In [158, 401, 394, 75, 523, 254] early image sensors with PWM pixel circuitry are presented.

The second fundamental approach to pulse modulation imaging is called pulse frequency modulation (PFM) and encodes the information about pixel illumination in the instantaneous frequency of digital pulses or inter-pulse intervals. The pulses are generated by a photocurrent-controlled oscillator in each pixel. Again the pre-charged photodiode capacitance is discharged by the photocurrent. When the fixed threshold is reached, a pulse is generated, the comparator resets the photodiode and the integration cycle starts again. The time between the pulses, respectively the pulse frequency, is proportional to the instantaneous pixel illumination. The first PFM

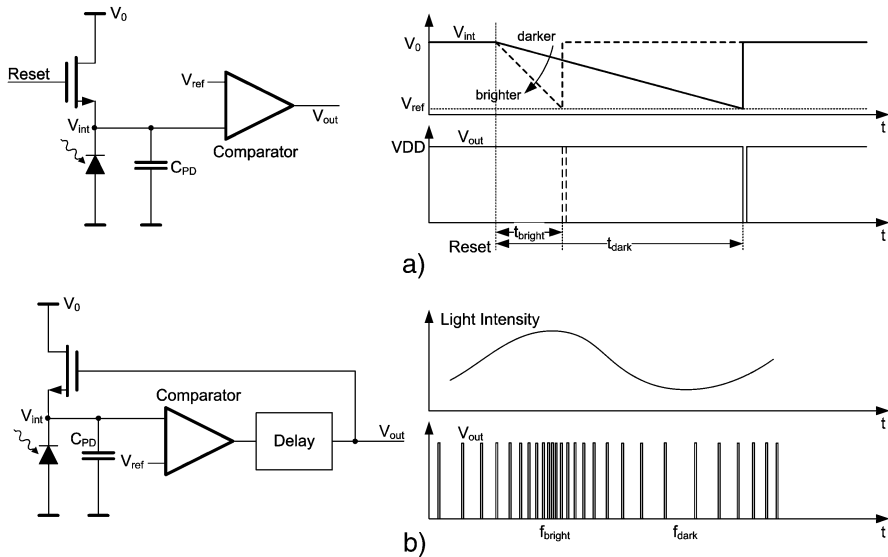


Fig. 4.14 Basic schematics and signal waveforms of (a) pulse width modulation (PWM) and (b) pulse frequency modulation (PFM) pixels

circuit was reported by Frohmader et al. [192] in 1982. The first PFM-based image sensor was proposed in 1993 [524] and demonstrated in 1994 [584].

Basic PWM and PFM circuits are shown in Fig. 4.14 (a) and (b).

The time-based scheme can represent a much wider dynamic range of pixel intensities than conventional CMOS APS image sensors. Since the integration time is inversely proportional to the photogenerated current, the dynamic range of the imager is limited to the dynamic range to which integration time can be measured and hence is no longer limited by the power supply rails. Or, from another point of view, the time-based approach allows each pixel to choose its own optimal integration time, by this means achieving high dynamic range and improved SNR.

PWM or PFM techniques are frequently used as part of a pixel-level analog-to-digital conversion scheme. This subject is addressed in the following section.

PWM/PFM pixels can be combined with asynchronous, digital array readout (address-event representation, AER) [65, 66] to instantaneously and asynchronously transmit the pixel pulses off-chip when they occur. This approach has yielded a variety of devices like “time-to-first-spike” (TTFS) type imagers [435, 340, 221, 492] or the Biomorphic Digital Image Sensor [133]. PFM digital pixels are also suitable for retinal prosthesis, where asynchronous and independent operation are desired [271, 272, 399].

Some of these vision devices have left the realm of synchronous, frame-based operation that is common to practically all image sensors covered up to now. Section 4.6.5 expands on the asynchronous, event-based vision approach.

4.6.3 Pixel-Level ADC

Many CMOS imagers provide their data in digital form, implying analog-to-digital conversion on-chip, usually one per array or one per pixel column. Miniaturization trends in semiconductor technology allow ever greater numbers of devices per unit area, and inspired research toward putting an ADC into each pixel of the imager array, thus realizing a digital pixel sensor (DPS). Digital pixel sensors are not meant to compete with conventional APS pixels in terms of fill factor or sensor resolution (hence image quality), but are designed to serve in specialized areas like high-speed imaging, video-oriented applications, and machine vision.

One-bit pixel-level AD conversion, or thresholding, has been reported in [60]. The first true pixel-level AD conversion technique appears in [188] and [581] with advancements in [582] and [583]. In a further development, Kleinfelder et al. [297] report a 10 kfps pixel-level single-slope conversion imager with global gray counter and in-pixel 8-bit memory.

PWM or PFM encoding of the illumination information is the basis of various pixel-level AD conversion schemes [58, 530]. In this case pixel pulses are not directly transmitted off-chip but their timing is converted to digital numbers and stored on-chip until read out.

For PWM, the pulse duration can be measured in cycles of a clock which is common to all pixels in the array using pixel-level counters [252]. In [291, 292] PWM pixel-level AD conversion with a global Gray code counter and pixel-level digital memory is presented. Bermak et al. [59] show a digital pixel sensor array with PWM-based AD conversion and linearization of the $1/x$ PWM encoding of the illumination information. Xu et al. [577] present a complementary pixel architecture (CAPS) for low-voltage (1V) operation in $0.25\ \mu\text{m}$ CMOS with PWM encoding and Kagawa et al. [273] show a 3-transistor PWM pixel.

In the case of PFM, the pixel-generated pulses are counted for a given period of time. Early image sensor implementations include a 32×32 PFM pixel array with the pulse counting done off-chip [584] and a 128-pixel PFM line sensor [147]. McIlrath et al. [367] use a sampled free-running photocurrent-controlled oscillator in each pixel to implement a first-order sigma-delta converter. Andoh et al. [16] report a 64×64 PFM pixel array with 8-bit counter in each pixel and line-wise successive readout. Wang et al. [556] present a 130 dB dynamic range 28×28 PFM sensor with shift-register counters and Boussaid et al. [71] present a scalable PFM pixel architecture with in-pixel 8-bit counter/memory and random access readout. In [493] a recent development of a 64×64 prototype PFM-based digital sensor with in-pixel 8-bit Gray-code counter/memory and intermediate (integration time) readout is reported.

4.6.4 Pixel-Level Digital Processing

Again driven by shrinking feature sizes of modern semiconductor processes and inspired by the progress made in pixel-level AD conversion, recent efforts were

targeted toward pixel-level digital signal processing. The most advanced devices integrate a fully programmable digital signal processor in each pixel of a sensor-processor array. Much like analog and cellular neural network (CNN) array processing architectures, this approach allows massively parallel operation and provides complex vision and image processing at high frame rates.

Ishikawa et al. [252] report an early 16×16 SIMD vision processor array with 512 digital transistors per pixel and 20 MHz clock frequency. In [270, 301] a 64×64 SIMD vision processor array for real-time image processing is shown.

Dudek et al. report the ASPA focal-plane digital processor array with asynchronous processing capabilities [334, 335].

Table 4.3 contains a comparison of analog and digital vision chips with programmable focal-plane processing [335].

4.6.5 Asynchronous Event-Based Vision

Classical image processing is based on the evaluation of data delivered by an image sensor system in the form of frames. Conventional, clocked image sensors acquire the visual information from the scene either sequentially for each sensor element (pixel) or each pixel line/column or in various patterns, pixel parallel, but always time-quantized at some frame rate. Each frame of data, that is recorded, transmitted, and needs to be post-processed in some fashion, carries the information from all pixels, regardless of whether or not this information has changed since the last frame had been acquired, usually not long ago. This method obviously leads, depending on the dynamic contents of the scene, to a high degree of redundancy in the image data. In addition, each frame imposes the same exposure time on every pixel, making it difficult to deal with scenes of high dynamic range.

The paradigm of asynchronous, event-based vision, as opposed to frame-based image sensing, is modeled after biological vision systems like the human retina. Biological visual systems have no notion of a frame – retinal outputs are asynchronous, massively parallel, and data driven. The output data volume of a self-timed, event-based vision sensor depends on the dynamic contents of the target scene – pixels that are not visually stimulated do not produce output. The result is a near-complete suppression of image data redundancy as compared to traditional, frame-based image sensors, drastically reducing demands on processing power, memory, transmission bandwidth, energy supply, etc.

The output of an asynchronous, address-event representation (AER) vision sensor is in the form of digitally coded addresses, specifying the x, y -address of the pixel in the array, communicated via an asynchronous, arbitrated digital data bus. The events are generated locally by the pixels, meaning that pixels operate autonomously and individually decide when to quantize the vision information, usually after some form of local gain control, and spatiotemporal processing and redundancy reduction.

Various different designs of AER vision devices have been developed during the past few years including the following:

- Imaging with TTFS encoding [492, 221]
- Imaging with PFM encoding in standard CMOS [133] and silicon-on-sapphire for backlight operation [132]
- Spatial contrast [126]
- Temporal contrast [317]
- PWM imaging with temporal contrast-event controlled conditional intensity readout [432]

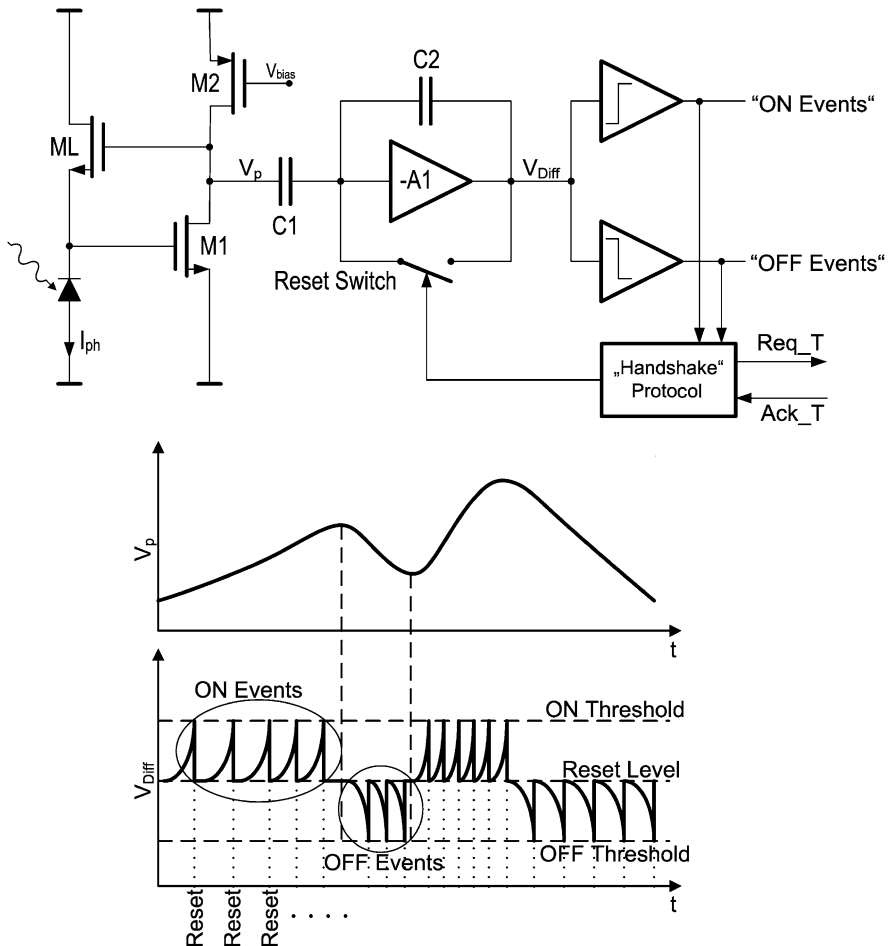


Fig. 4.15 DVS pixel schematic and typical signal waveforms of photoreceptor voltage V_p and spike events V_{Diff} over time. The rate of change is encoded in the inter-event intervals while the polarity of the gradient, going brighter or darker, is determined by the signal reaching the upper or lower threshold, respectively (“ON” or “OFF” events)

- Orientation-selective filtering [111]
- Exotic devices like foveated imagers with multiple pixel types [36] and dark-active imaging [222]

One recent example of an asynchronous AER event-based vision device is the temporal contrast *Dynamic Vision Sensor* (DVS) developed by T. Delbruck et al. [316, 317]. This sensor is modeled after the transient pathway of the human/vertebrate retina and performs local relative change detection in an array of autonomous pixels. Kramer [304] reported the predecessors to this device. The data delivered by this type of sensor contain high temporal resolution information about changes in the observed scene which can, in most cases, be attributed to object motion. Because output bandwidth is automatically dedicated to dynamic parts of the scene, these devices are especially suited for applications involving motion detection and analysis. Absolute (DC) levels are discarded in the pixel, hence there is no intensity information contained in the sensor data.

Figure 4.15 shows the DVS pixel schematic and typical signal waveforms. The pixel circuit contains a fast logarithmic photoreceptor circuit with common-source feedback (PD, ML, M1, M2), a differencing circuit that amplifies changes with high precision (A1, C1, C2) and two common-source comparators for positive and negative gradient events. The photoreceptor circuit automatically controls individual pixel gain by its logarithmic response while, at the same time, responding quickly to changes in illumination. The continuous-time photoreceptor FPN is removed by capacitive coupling and by balancing the output of the differencing circuit to a reset level after the generation of each event to re-establish the operating point of the circuit. The gain of the change amplifier is determined by the well-matched capacitor ratio $C1/C2$. Because the differencing circuit removes DC and due to the logarithmic conversion in the photoreceptor, the pixel is sensitive to temporal contrast and quantizes relative rather than absolute changes in pixel illumination.

Several devices have been built based on the temporal contrast DVS pixel circuit – low- and medium-resolution array sensors [316, 317], line sensors with on-chip event time-stamping for high-speed industrial vision [431], microbolometer-based transient IR sensors [363], and a QVGA vision sensor with asynchronous, redundancy-free PWM grayscale imaging [432].