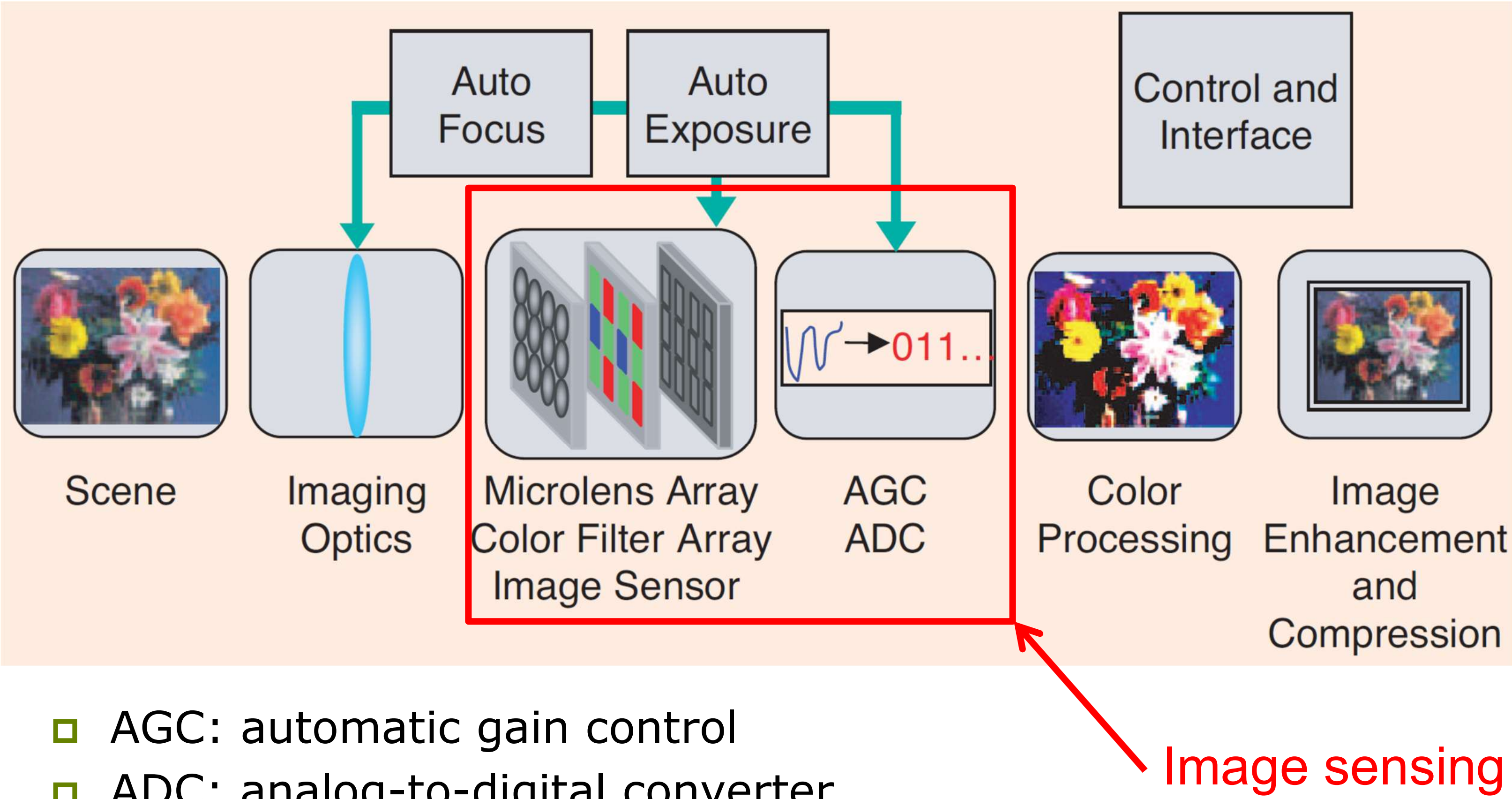
TSBB21, Lecture 2
Image Sensing

- Image sensing and the digital image sensor
- The photovoltaic effect, the photodiode
- Passive and active pixel sensors
- The readout problem, the CCD array and the CMOS sensor array
- Rolling or global shutter
- Blooming
- Fill factor and micro-lenses
- Noise sources: temporal and FPN
- Shot noise (also known as photon noise, Poisson noise, or Quantum noise)
- SNR
- Color cameras
  - 3 chip, 1 chip cameras, Bayer filters, …
- The video camera:
  - Interlaced vs. progressive scan
- Shading correction
- Reading material and thanks to:
  - Gamal & Eltoukhy: CMOS Image sensors, a few images
- Thanks to:
  - Gonzales & Woods: Digital Image Processing, Global Edition, 4th edition, a few images.
  - Klas Nordberg, who initiated this course. Many slides in this lecture are similar to his slides.

Figure from:
Gamal & Eltoukhy

The imaging system pipeline

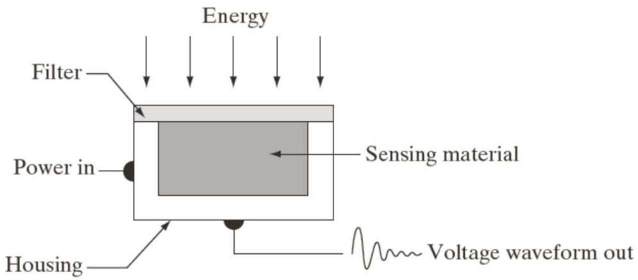


- AGC: automatic gain control
- ADC: analog-to-digital converter
- Color Processing: includes white balancing and color correction

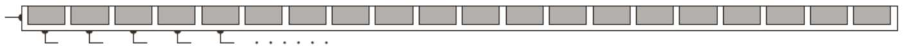
Figure from:
Gonzales & Woods

The digital image sensor

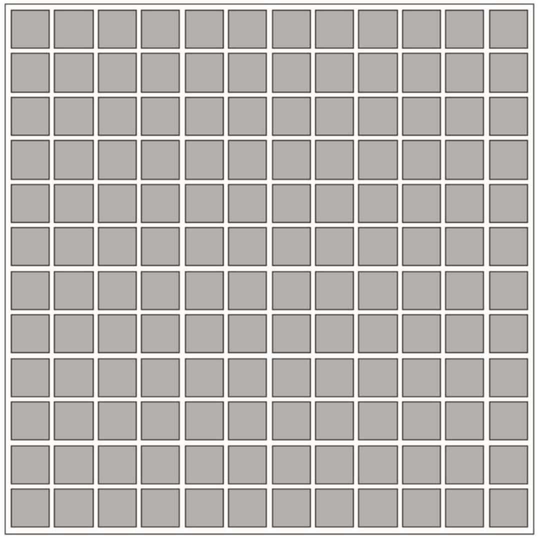
Single sensor element:



Line sensor:

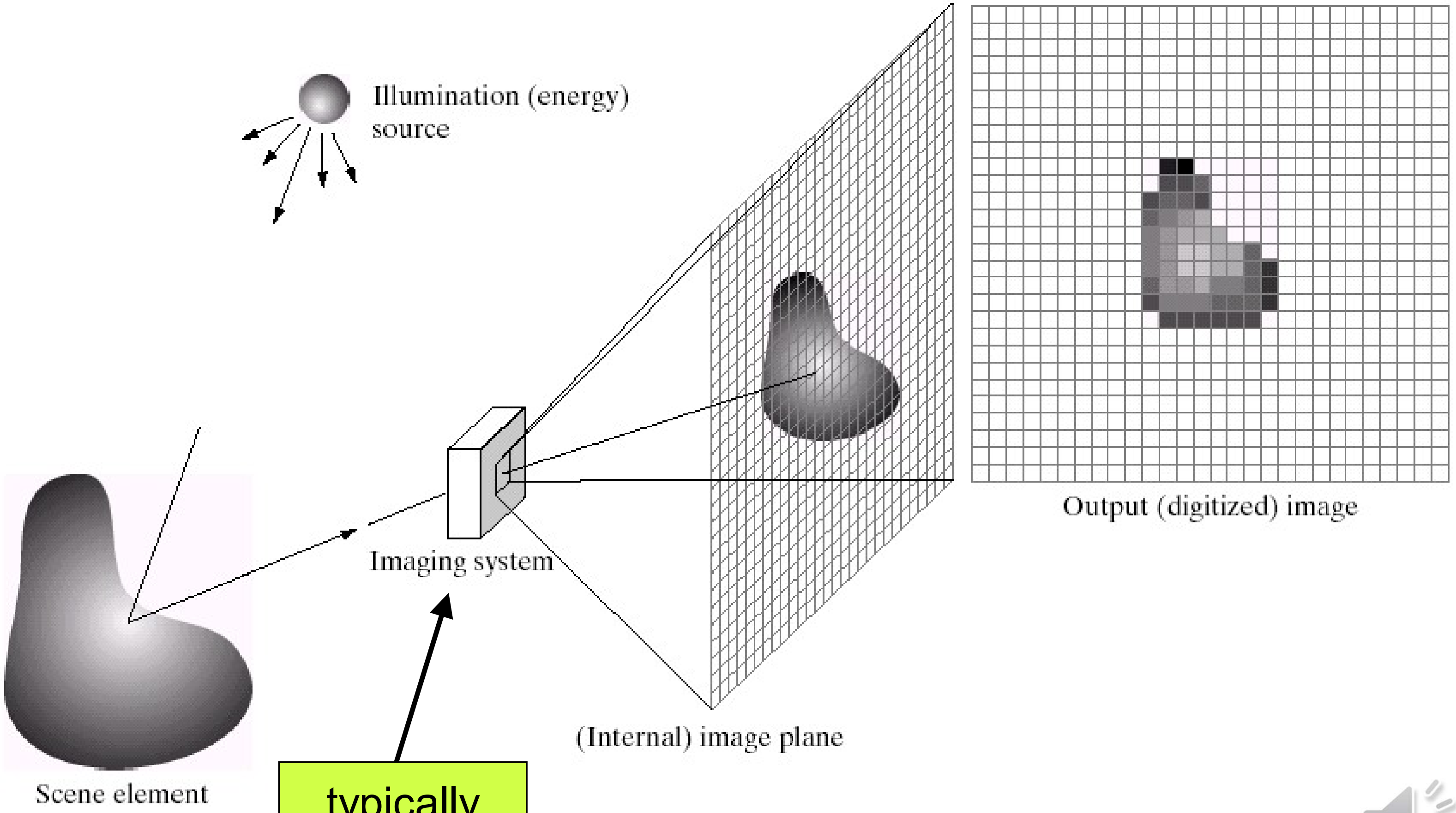


Array sensor:



The digital image acquisition process

Figure from:
Gonzales & Woods



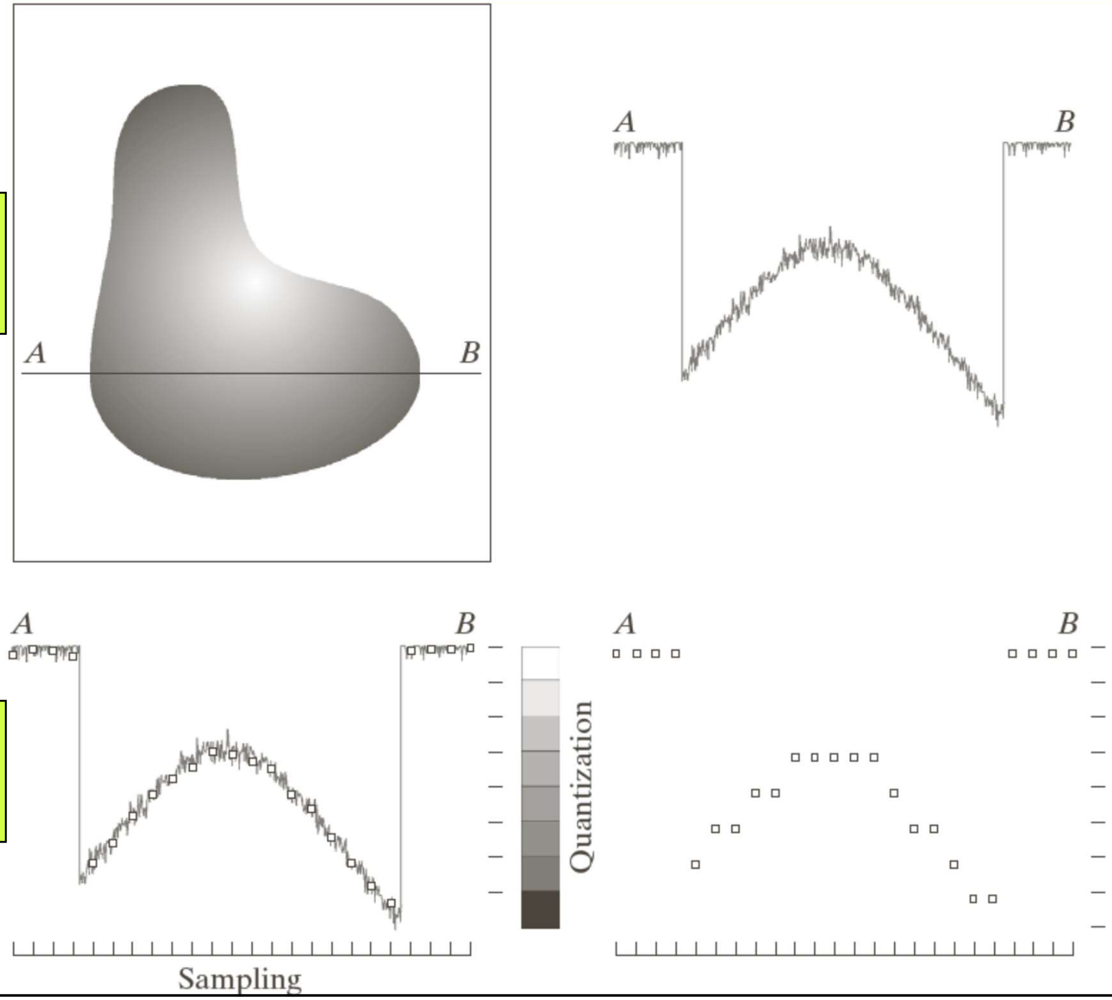
typically
a camera

Image sampling and quantization

Figure from:
Gonzales & Woods

Continuous Image

Line from image



After sampling

After quantization

The photovoltaic effect / the photoelectric effect

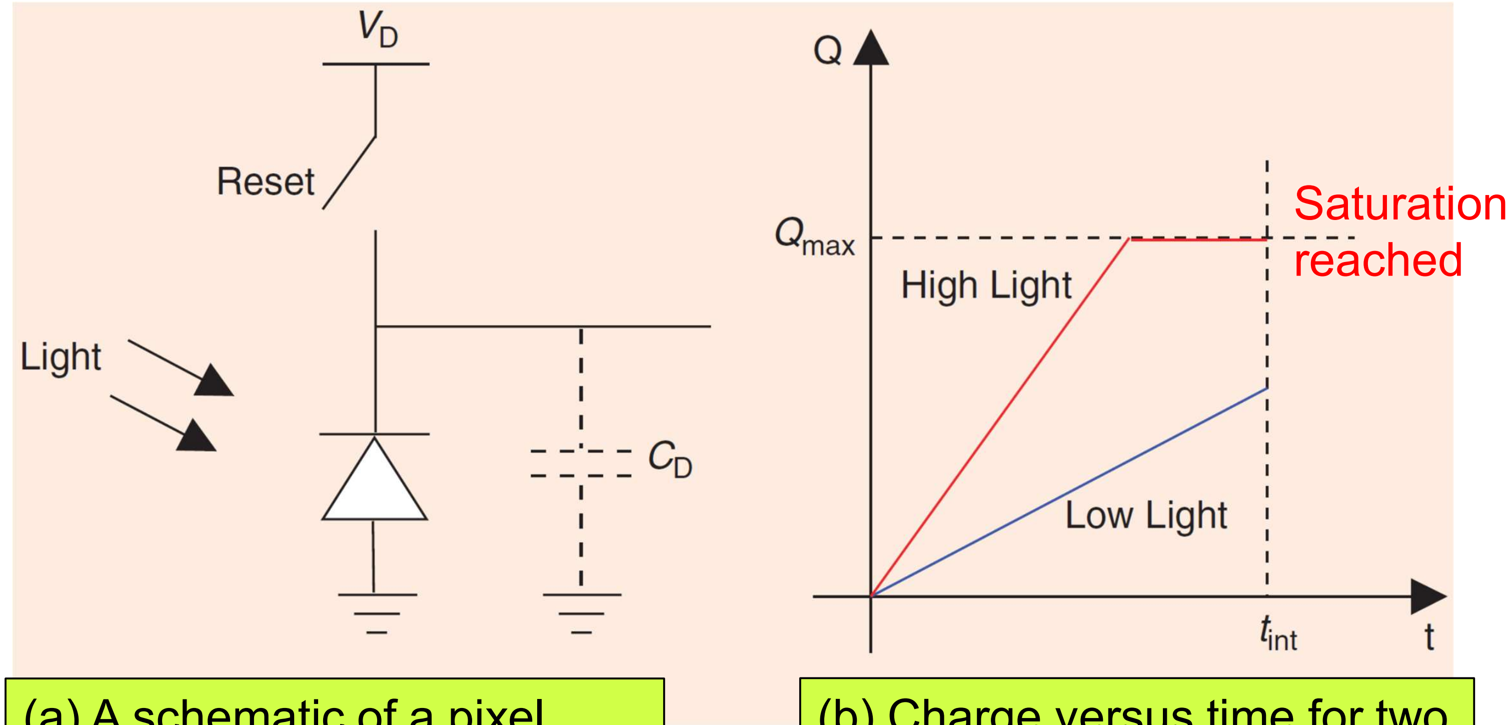
- The main or desired interaction in a photo-detector is **absorption**
  - a photon is converted to an electron/hole pair
- Electrons bond to atoms with a certain energy Q_g
  - Photon absorption occurs with a certain probability when the photon energy
    $E = h\nu \geq Q_g$
- When a photon is absorbed: the electron is released from the atom and becomes "free"
- Leaves a "hole", a missing electron, in the atom. The hole is also "free" to move around
- This is called the "photovoltaic effect" or "photoelectric effect".

The Photodiode

- The *photovoltaic effect* (*photoelectric effect*) is utilized in a *photodiode* to produce a voltage
  - For image sensors in the range: near IR –Visible light - UV
  - For solar cells
- As the range of photocurrents produced under typical illumination conditions is too low (in the range of femto- to picoamperes) to be read directly, it is typically *integrated* and read out as charge or voltage at the end of the exposure time.

A pixel consisting of a photodiode and an integrating capacitor

Figure from:
Gamal & Eltoukhy



(a) A schematic of a pixel operating in *direct integration*.

(b) Charge versus time for two different photocurrent values.

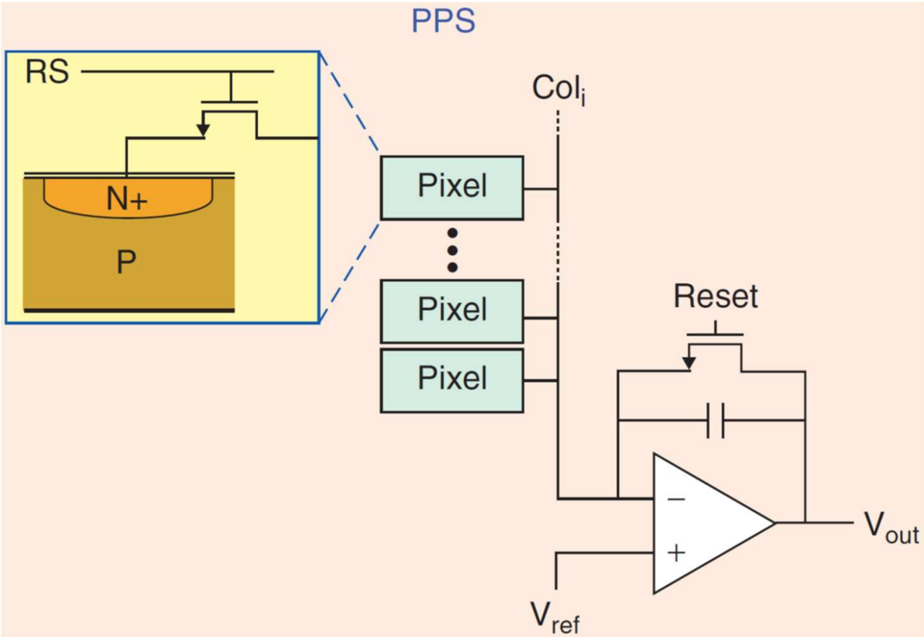
CMOS, short explanation

- CMOS = complementary MOS
- MOS= metal-oxide semiconductor
- Both PMOS (P-doped MOS) and NMOS (N-doped MOS) are included
- Used for Analog circuits such as image sensors
- Also used for Digital logic circuits

Figure from:
Gamal & Eltoukhy

The passive pixel sensor (PPS)

- The *passive pixel sensor (PPS)* is the earliest CMOS image-sensor architecture.
- The PPS pixel includes a photodiode and a row-select transistor. The readout is performed one row at a time.
- At the end of integration, charge is read out via the column charge-to-voltage amplifiers.



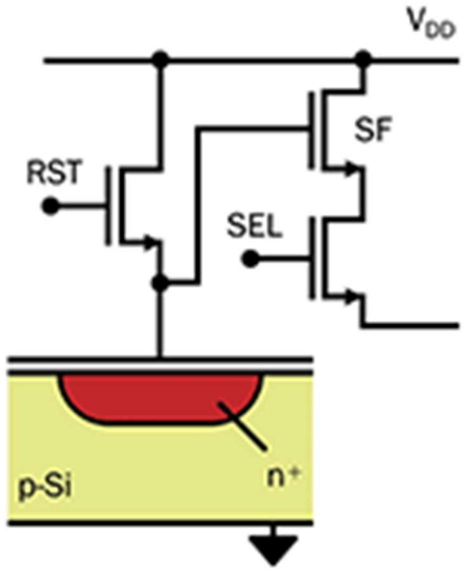
Active pixel sensors (APS)

- Developments in the early 1990's and onward have led to an improved CMOS sensor called *Active Pixel Sensor* (APS)
- Basic idea:
  - Add an amplifying transistor to the pixel:
    - Voltage readout instead of charge transport
    - The readout line becomes less sensitive to noise
  - With modern technology:
    - the extra transistors per pixel can be very small compared to the rest of the pixel area devoted to light sensing)

      => reasonable fill factor

Active Pixel with a Photo Diode (PD, 3T)

- The 3 transistor (3T) pixel
  - One for recharging the diode (RST)
  - One used as voltage buffer (SF)
  - One for connecting the voltage to the readout line (SEL)
- Less sensitive to noise compared to the passive pixel sensor (PPS)

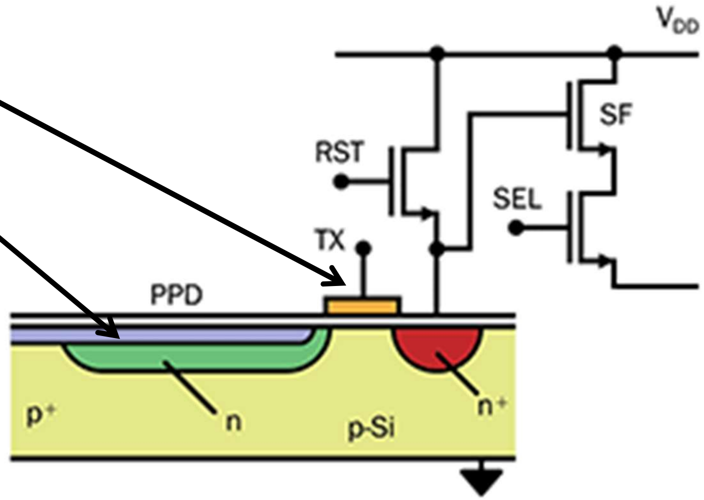


https://www.photonics.com/Articles/Advances_in_CMOS_Image_Sensors_Open_Doors_to_Many/a57683

Active Pixel with a Pinned Photo Diode (PPD, 4T)

- The *pinned photodiode* (PPD) incorporates a p+ implant above each pixel's light-sensitive structure. It permits the total transfer of charge onto the measurement node under the control of a 4th transistor, the transfer gate TX.
- The other three transistors operate as a standard 3T APS.

Extra doping layer and a "transfer gate" (TX) results in a high quality photo sensitive device



https://www.photonics.com/Articles/Advances_in_CMOS_Image_Sensors_Open_Doors_to_Many/a57683

The three pixel variants summarized

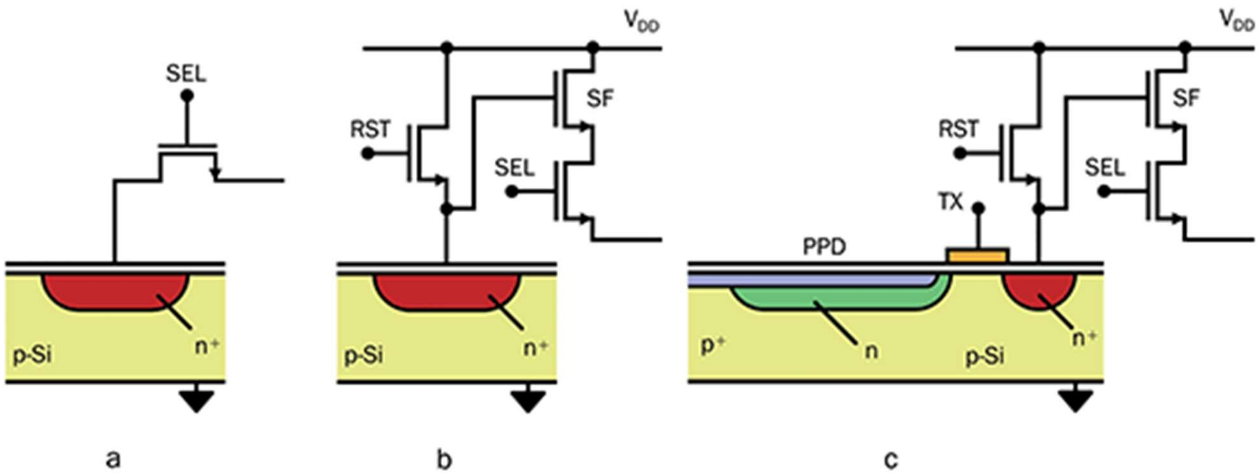


Figure: Structure of a …

a) passive pixel sensor element (PPS)

b) active pixel (PD, 3T)

c) pinned photodiode pixel (PPD, 4T)

SEL = addressing transistor. RST = reset transistor. SF = source follower. TX = transfer gate. VDD = power supply voltage. PPD = pinned photodiode

https://www.photonics.com/Articles/Advances_in_CMOS_Image_Sensors_Open_Doors_to_Many/a57683

The read-out problem

- Light has caused a change in electric voltage or charge in the pixel sensor and this change needs to be measured to produce an image
- Traditionally not measured per detector element
  - Would require many components per detector
  - Would give too small fill factor for 2D arrays
- The read-out problem:
  - The voltage/charge has to be transported out of the array and sensed outside
  - Often with a single sensing unit per sensor array or per column
- Two principles for solving the read-out problem
  - The *CCD array* (CCD = Charge Coupled Device)
  - *CMOS sensor array*
- CMOS makes up about 95% of the market today
  - CCD is used in some scientific applications, where low noise is a demand and high speed is not a necessity.

Figure from:
Gamal & Eltoukhy

The CCD array
(Now used more and more seldom)

- In a CCD, the charge is *shifted out* of the array via vertical and horizontal CCDs, converted into voltage via a simple follower amplifier, and then serially read out
- A depleted MOS structure was used as the photodetector in early CCD devices
- Nowadays, Pinned Photo Diodes (PPD) are used in nearly all CCD

Nobel Price in 2009!

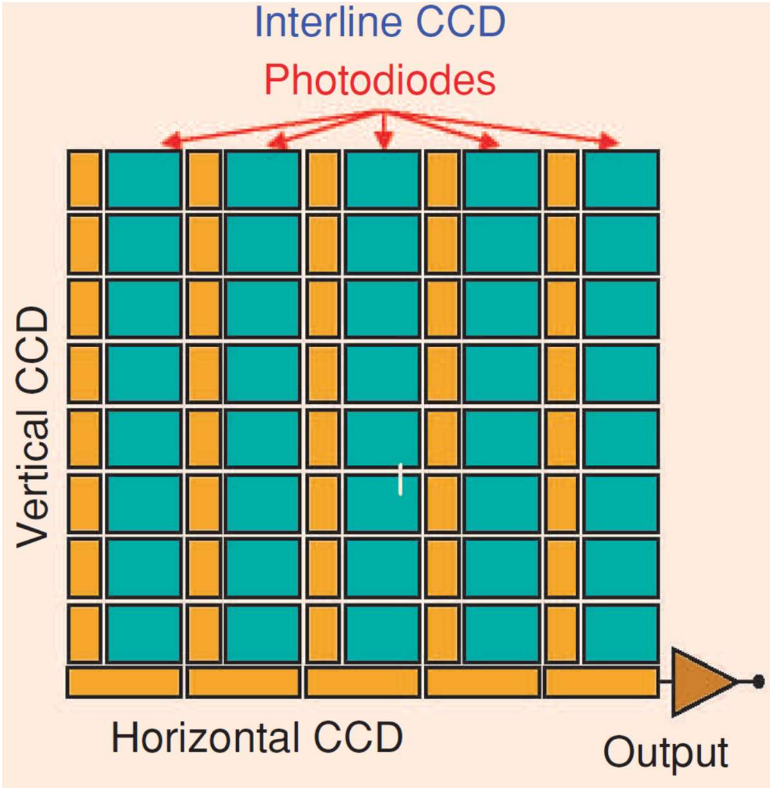
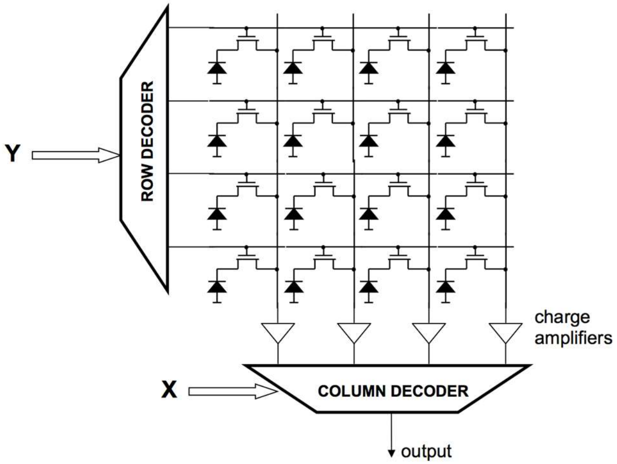


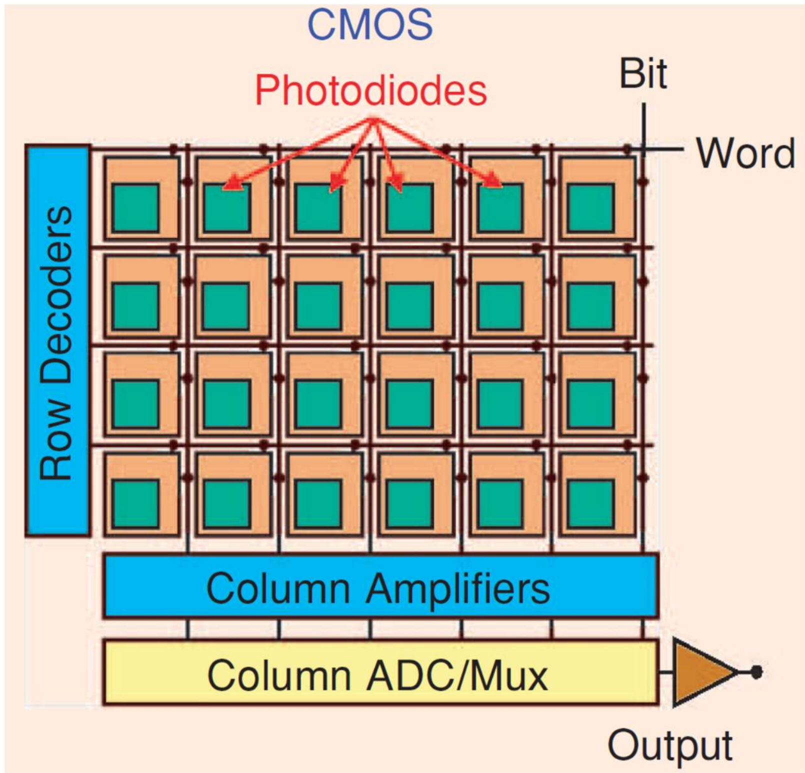
Figure from:
Gamal & Eltoukhy

CMOS sensor array

- In a CMOS image sensor, charge voltage signals are read out one row at a time in a manner similar to a random access memory using row and column select circuits
- Example below:

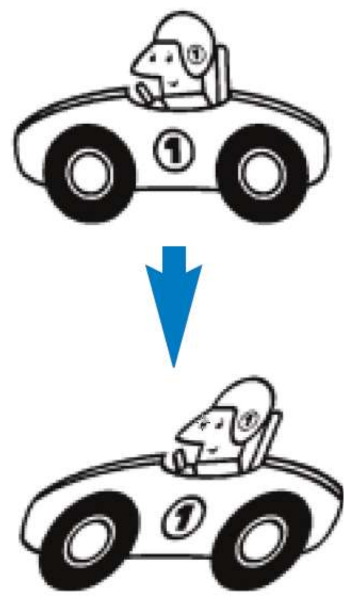
1-transistor (1T) pixel array. Column amplifiers convert charge to voltage





Rolling or global shutter

- For a CMOS sensor array, the simplest approach is to make the exposure and read-out happen line after line
- Each successive line is exposed at successive points in time.
- This is called: *rolling shutter*.
- For a stationary scene, rolling shutter is OK
- If the camera or the scene is moving, a rolling shutter may distort the image
- This is called:
  - Rolling shutter problem
  - Jello/jelly effect
  - CMOS distortion
- The other alternative is: *global shutter*
- CCD usually uses a global shutter



Jello effect / CMOS distortion example

- CMOS distortion image found on the Internet

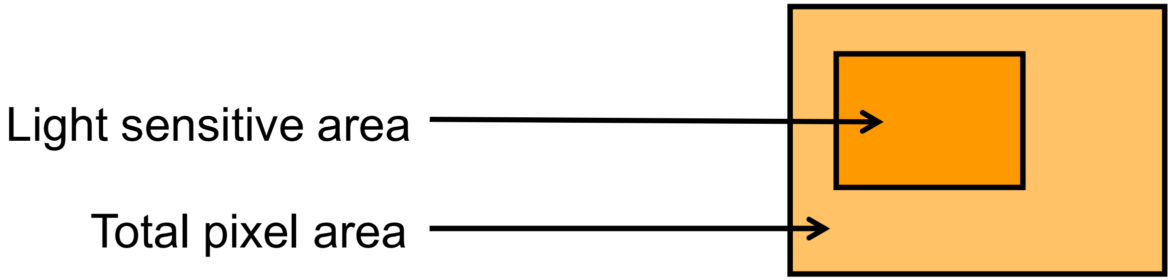
Blooming

- The photodiode “collect” electric charge in a small region, *a well*, corresponding to the conductor region
- When this region becomes saturated, the charge spills over to neighboring elements
- Saturation is when the maximum charge/signal has been reached.
- This is called *blooming*
- Barriers between the detectors can reduce this effect, but not eliminate it entirely

Image illustrating blooming found on the Internet:

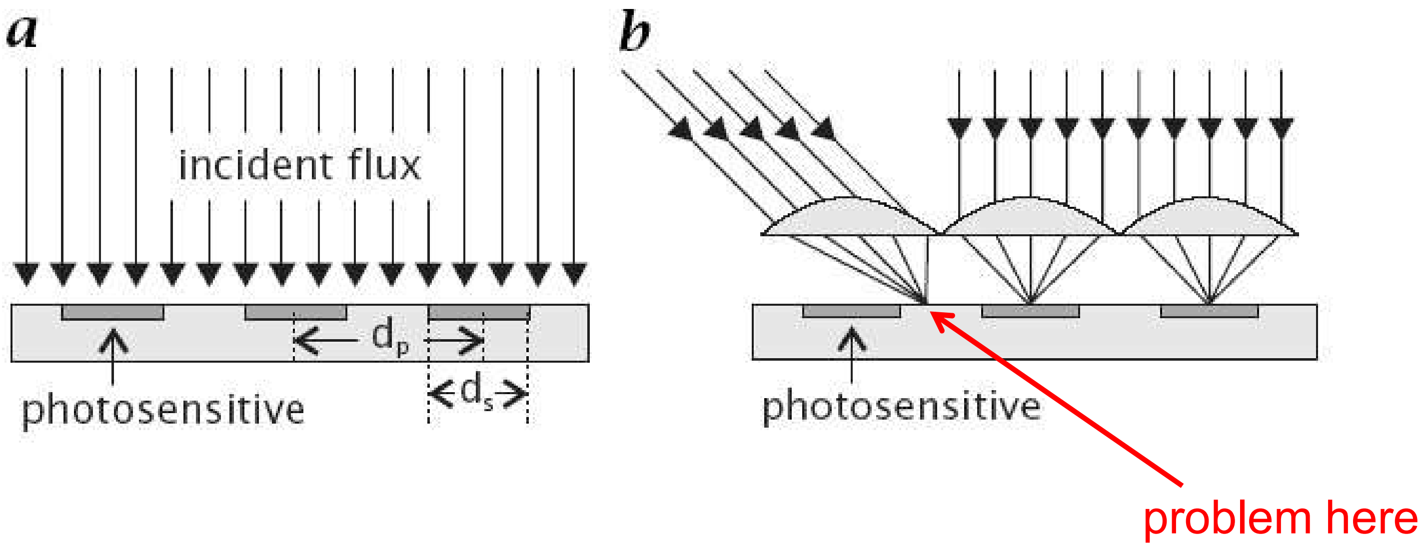
Fill factor

- In practice, the light sensitive area of an image sensor cannot fill the entire detector area.
- Electronic components and wiring reduce the light sensitive area
- The *fill factor* is the percentage of the total area which is light sensitive
- In the example here, the fill factor is 25%



Micro-lenses

- To overcome low fill factors, an array of micro-lenses in front of the sensor array can be used

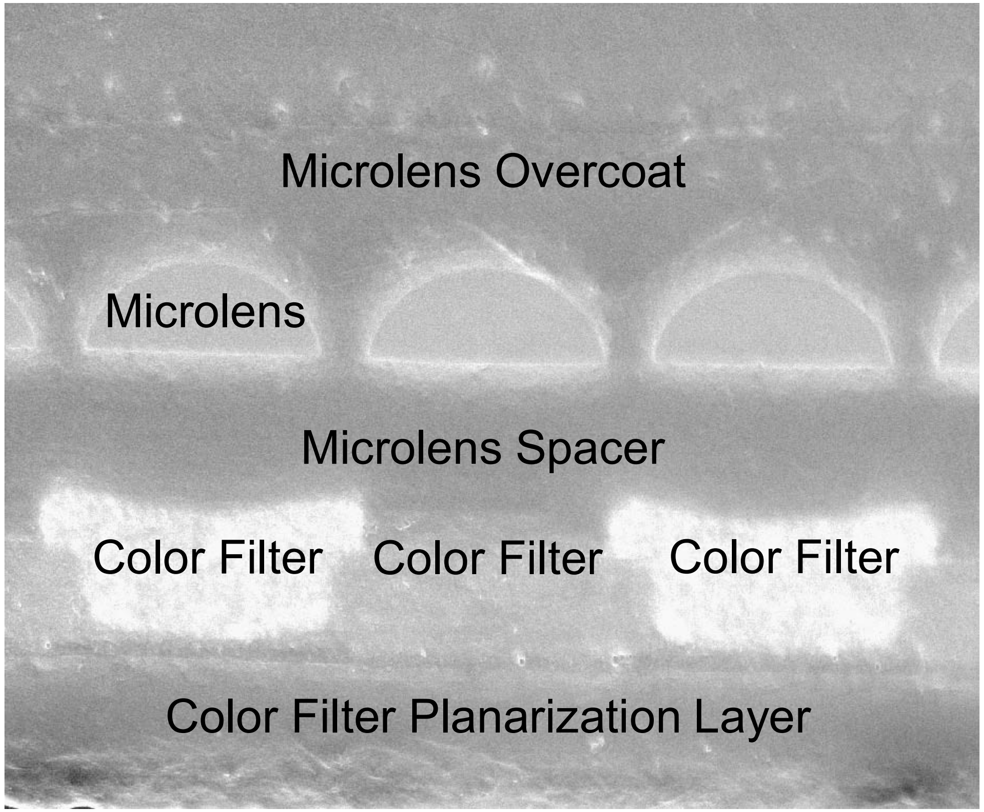


Micro-lenses

- Micro-lenses enhance the fill factor
- But
  - Due to the manufacturing process, the detector area can often have an inhomogeneous sensitivity
  - When light is focused onto a smaller spot in the sensor, the inhomogeneities become more noticeable as measurement noise
  - At high incident angles, this spot may miss the detector area, see the previous illustration

A cross-section SEM photograph of the micro-lenses

Figure from:
Gamal & Eltoukhy



SEM:
Scanning
Electron
Microscope

More on color filters later!

A cross-section SEM photograph of an image sensor showing the microlens and a color-filter array (CFA) deposited on top of the photodetectors.

Noise sources
(we focus more on the red marked)

- Temporal noise
  - pixel reset circuit noise
  - flicker noise
  - thermal noise
  - readout circuit noise due to readout device thermal and flicker noise
  - photodetector shot noise
  - dark current noise
  - quantization noise
- Fixed Pattern Noise
- Shot noise due to photons arriving on the image sensor

Pixel reset circuit noise and Flicker noise

- Pixel reset circuit noise
  - The measured voltage depends on the “fix” bias voltage over the photo diode
  - This voltage has always some amount of variation = noise
- Flicker or 1/f noise
  - Inhomogeneities and impurities in the materials produce low-frequency noise due to statistical fluctuations in various parameters which control the photon-to-voltage conversion
- These two factors may vary both across the array (spatially) and over time

Thermal excitation

- Because of heat in the material:
  - Electrons are always excited (moved from the valence band to the conduction band) due to thermal energy in the material
  - This induces an electric current I_{thermo}

$$\boxed{I_{thermo} \propto E e^{-\frac{Q_g}{kT}}}$$

  - Q_g is the gap between the material's valence and conduction bands
  - E is the electric field
  - T is the absolute temperature
  - k is Boltzmanns constant

Thermal noise

- I_{thermo} is not a constant current, it is rather a noise signal with a mean given by the last expression
- We have to treat it as a random signal added onto the desired signal I_{photo}
- Thermal noise can be reduced by cooling

Dark current

- The space-charge region is not a perfect isolator => there is a small leakage current
- It is called *dark current* since it discharges the capacitor even when no photons are absorbed
- Dark current in standard submicron CMOS processes is orders of magnitude higher than in a CCD and several process modifications are used to reduce it

The output voltage

- In the end, the output voltage of the sensor array, per detector element, is:

$$\boxed{V = g \cdot I + o + \Delta V}$$

I = incident radiant flux

g = gain factor

o = offset voltage

ΔV = noise voltage

These two are determined by the material and the design

Also determined by the material and the design but also temperature, number of photons, etc.

Fixed pattern noise (FPN)

- In the ideal world: *gain (g)* and *offset (o)* are constant over the sensor array
- In the real world: both g and o vary over the sensor array
  - Small variations in standard camera chips
  - Larger variations in many IR-sensors
- May even vary over time (for IR sensors)
- *Hot pixels*: strong local variation in g or o
- *Dead pixels*: $g \approx 0$

Dynamic range

- The dynamic range is the SNR of the largest detectable signal V_{max}

$$\text{DR} = 20 \cdot {}^{10}\log \frac{V_{max}}{\Delta V}$$

- Typical values
  - CMOS: 40-60 dB (may have increased last years)
  - CCD: 60-70 dB
  - Human eye: > 90 dB

Quantization noise

Ex) Often b=8 bits per pixel is used. Then $2^8=256$ different levels can be represented, e.g. 0,1,2,3,…,255.

- If b bits are used to represent a voltage up to V_{max}

$$\boxed{b = \frac{\mathrm{DR}}{20 \cdot {}^{10}\log 2}}$$

- gives a quantization noise of the same magnitude as the image noise
- Often, we want a few more bits than this to represent the image signal accurately

Shot noise

due to photons arriving on the image sensor

- Also known as: photon noise, Poisson noise, or Quantum noise
- Photons arriving on an image sensor carry a statistical variation of fluctuations in the photon arrival rate at a given point. This phenomenon is known as “photon noise” and follows the Poisson distribution.
- Additionally, there are inherent shot noise sources within the sensor.
  - Ex) shot noise due to dark current is also Poisson distributed.

Poisson noise

- Probability mass function:

$$\Pr[N = k] = \frac{a^k}{k!} e^{-a}, \quad \text{för } k = 0,1,\ldots$$

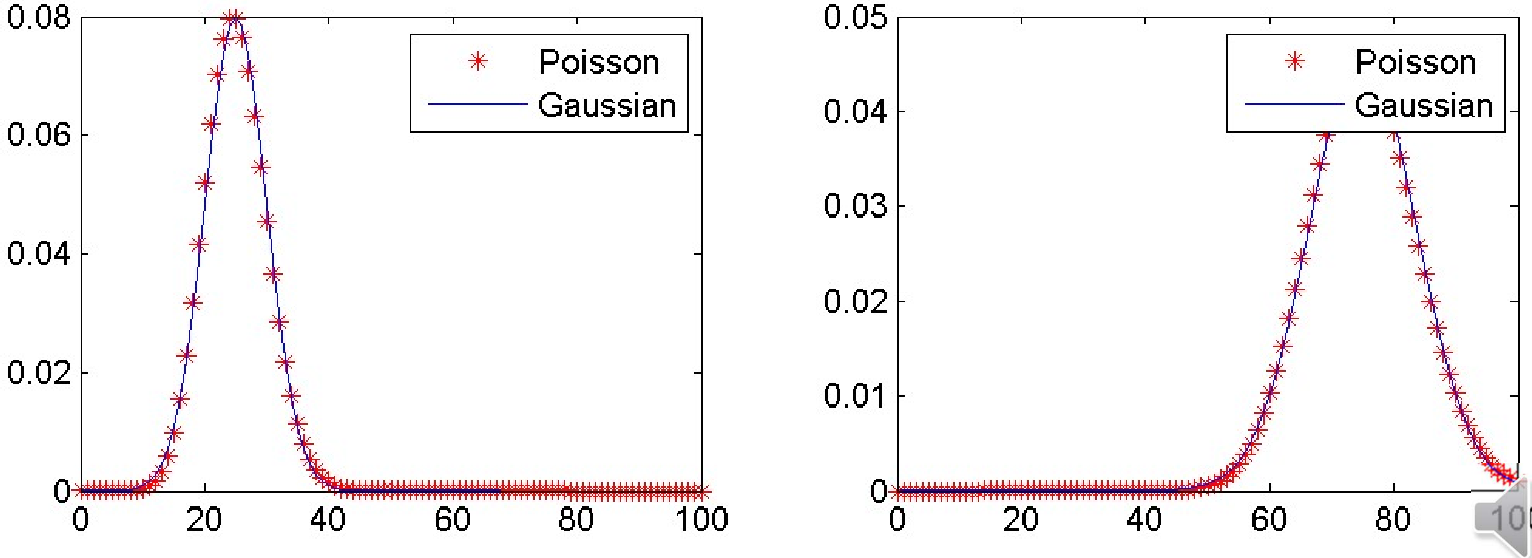
- Mean and variance:

$$\mu_N = a \quad \text{and} \quad \sigma_N^2 = a$$

- Is used to statistically characterize the distribution of photons per unit of area.
- Important for normal camera images, and medical image systems using X-rays and gamma-rays.

Poisson noise

- Previous slide gave that Poisson noise has equal mean and variance: $\mu_N = \sigma_N^2 = a$
- It can be approximated with Gaussian noise with with: $\mu_N = \sigma_N^2 = a$



Poisson noise

- The standard deviation

  $$\sigma_N = \sqrt{a}$$

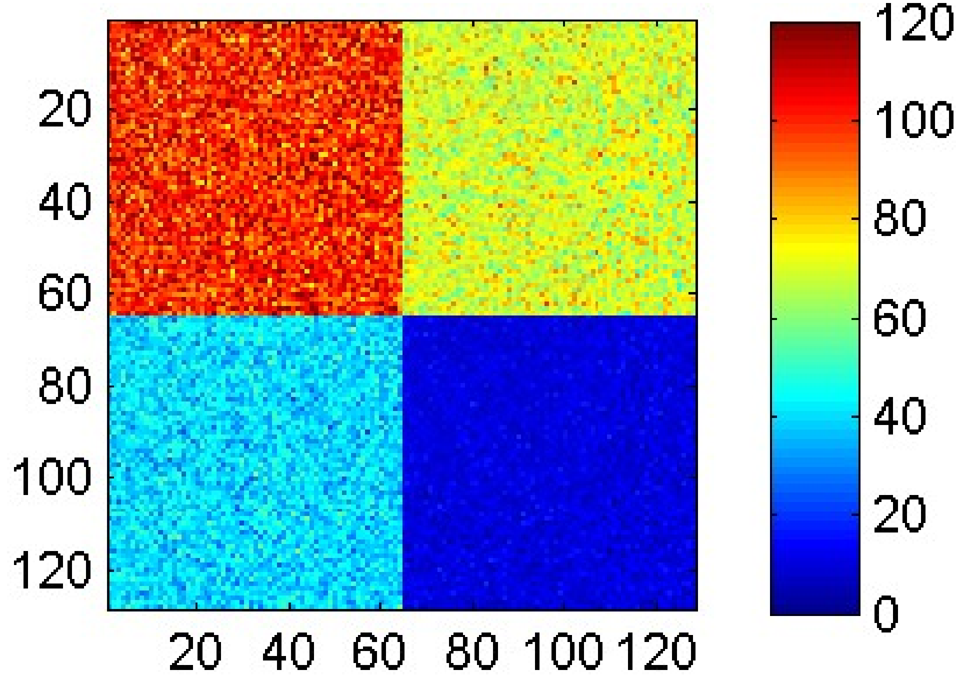
- increase with the signal value

  $$\mu_N = a$$

- But the signal to noise ratio (SNR)

  $$\mu_N/\sigma_N = a/\sqrt{a} = \sqrt{a}$$

  improve!



SNR

- Signal-to-noise ratio (SNR) describes the quality of a measurement. In image sensors, SNR refers to the relative magnitude of the signal compared to the uncertainty in that signal on a per-pixel basis.
- Specifically, it is the ratio of the measured signal to the overall measured noise (frame-to-frame) at that pixel.
- High SNR is particularly important in applications requiring precise light measurement.

SNR variants

$$\mathrm{SNR} = \frac{P_{signal}}{P_{noise}} = \left(\frac{A_{signal}}{A_{noise}}\right)^2$$

$$\mathrm{SNR_{dB}} = 10\log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) = 20\log_{10}\left(\frac{A_{signal}}{A_{noise}}\right)$$

- Psignal is the power of the signal.
- Asignal is the amplitude of the signal.
- For the noise effect, Pnoise = “the noise variance” is often used.

- Alternative definitions:

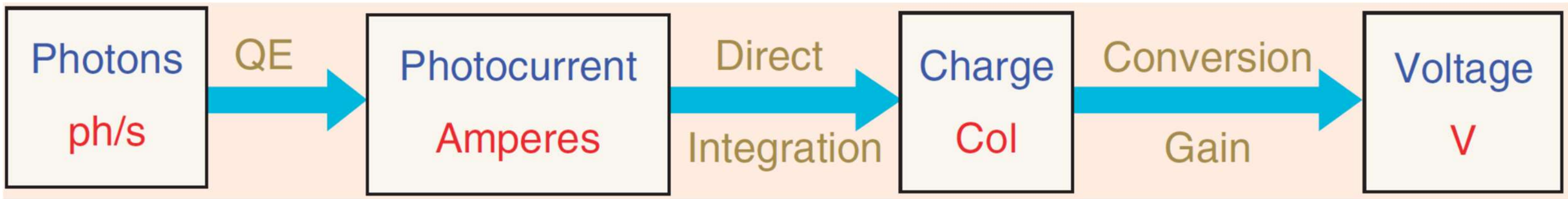
$$\mathrm{SNR} = \frac{\mu^2}{\sigma^2} \qquad \mathrm{SNR} = \frac{\mu}{\sigma}$$

- where μ is the mean of the signal (”its amplitude”) and σ is the standard-deviation of the noise

Figure from:
Gamal & Eltoukhy

Quantum efficiency (QE)

- The main imaging characteristics of a photodiode are external quantum efficiency (QE) and dark current.
- External QE is the fraction of incident photon flux contributing to the photocurrent.
- External QE can be expressed as the product of internal QE and optical efficiency (OE).
  - Internal QE is the fraction of photons incident on the photodetector surface that contributes to the photocurrent
  - OE is the photon-to-photon efficiency from the pixel surface to the photodetector's surface



Signal-to-Noise Ratio for sensors

Ex) an equation for a CCD array

- According to:
- https://www.photometrics.com/learn/imaging-topics/signal-to-noise-ratio
- The SNR for a CCD camera can be calculated from the following equation:

$$SNR = \frac{I\ QE\ t}{\sqrt{I\ QE\ t + Nd\ t + {N_r}^2}}$$

SNR definition used here: $\mathrm{SNR} = \frac{\mu}{\sigma}$

  - I = Photon flux (photons/pixel/second)
  - QE = Quantum efficiency
  - t = Integration time (seconds)
  - Nd = Dark current (electrons/pixel/sec)
  - Nr = Read noise (electrons)

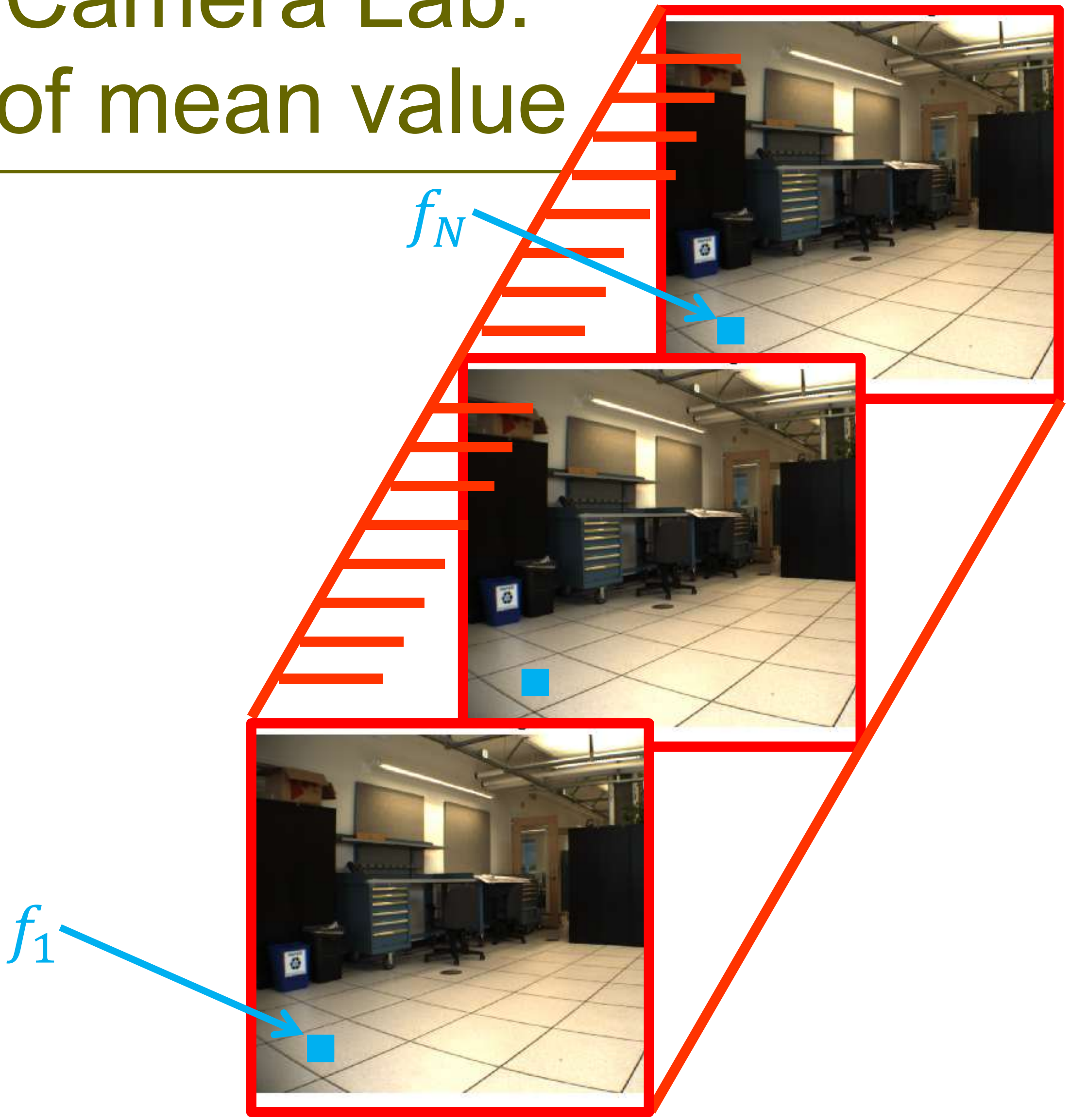
Ex) Assume a high photon flux *I*. Then SNR is proportional to the square root of the photon flux.

- Read noise refers to the uncertainty introduced during the process of quantifying the electronic signal on the CCD.

It is unclear how thermal noise is included in the equation. (?)

The Digital Camera Lab: Estimation of mean value

$$m = \frac{1}{N}\sum_{n=1}^{N} f_n$$



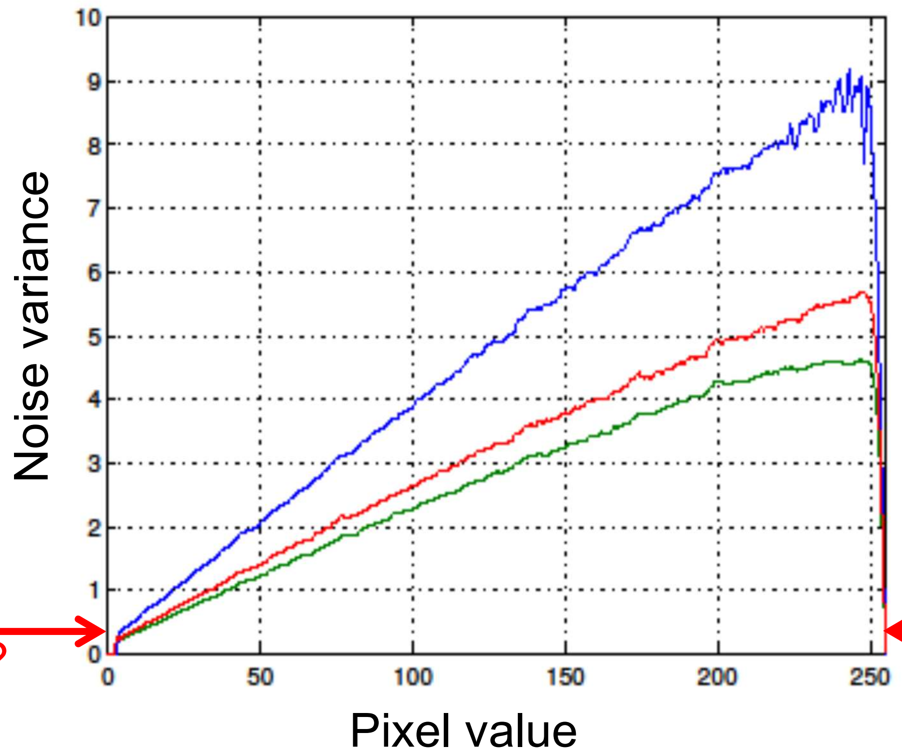
The Digital Camera Lab: Estimation of variance

$$s^2 = \frac{1}{N-1}\sum_{n=1}^{N}(f_n - m)^2$$

$$s^2 = \frac{N}{N-1}\left(\left(\frac{1}{N}\sum_{n=1}^{N} f_n^2\right) - m^2\right) \approx \left(\frac{1}{N}\sum_{n=1}^{N} f_n^2\right) - m^2$$

The Digital Camera Lab: Poisson noise

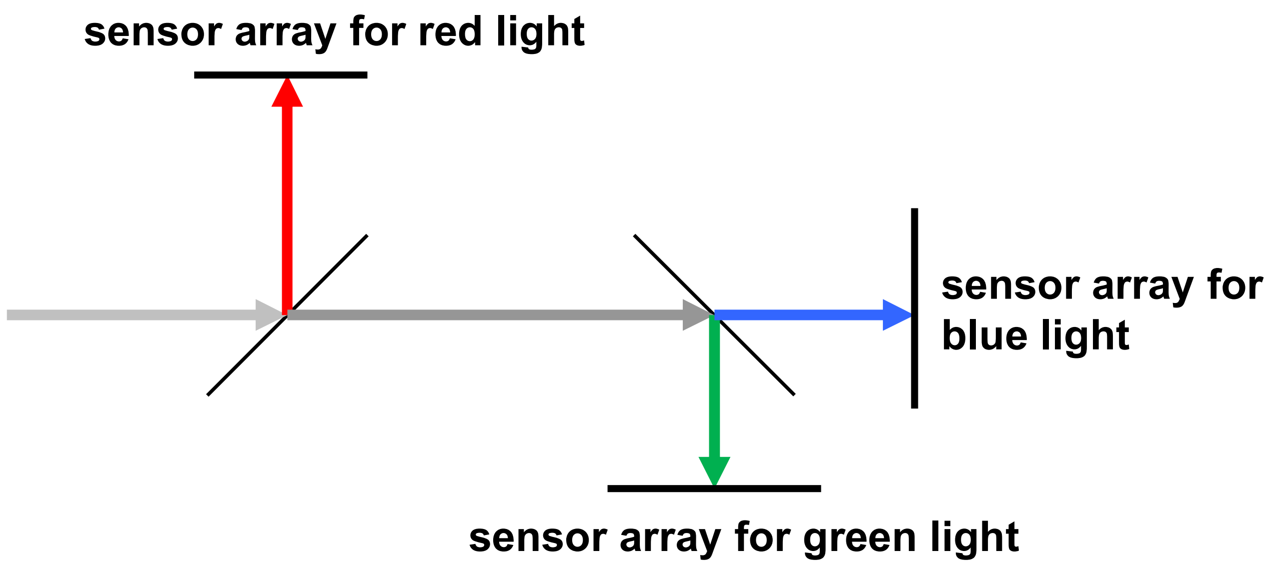
- The noise variance depends linearly on the image intensity (mean of the image pixel). This is in agreement with the Poisson distribution. (Lab exercise)



Does not cross the origin. Why? Lab question

Goes down to zero due to? Lab question

3 chip color cameras



- 3 identical standard sensor chips
- 2 semi-transparent mirrors that refract different wavelengths

3 chip color cameras

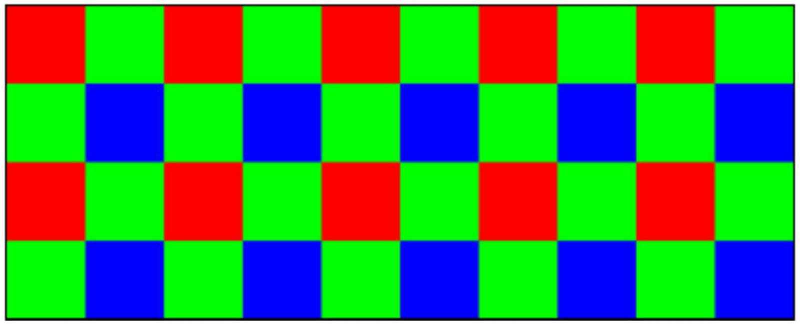
- Based on standard “black-and-white” sensor chips (3 identical sensor chips)
- The 3 sensor arrays need to be aligned with tolerances smaller than the inter-pixel distance
- Gives good performance, but …
  - is expensive
  - Is only used in professional cameras

1 chip color cameras

- To reduce cost, the 1 chip color camera:
  - Use one sensor array
  - Place a *color filter* on top of each detector element
  - Each detector element is now sensitive to only a specific wavelength range
  - Reduces the fill factor for each range
  - The colors are not measured at the same places. This may give color-aliasing.
- Most consumer cameras output only the interpolated image (typically compressed using JPEG)
- In more advanced cameras, the raw un-interpolated image can be read out from the camera and processed externally by the user.
  - We will use such images in the lab!

A color filter: The Bayer filter

- A Bayer filter is an *optical* filter placed over the sensor
- A common Bayer filter is shown below



1 chip color cameras: Color post-processing

- We can see the image detected by the sensor as a monochrome (grayscale) signal (the ”raw” image)
- An RGB signal (3 components per pixel) is then produced by interpolation from the raw image, using a set of space-varying filters for each of the three components (*demosaicking*)
- Note: two types of filtering!
  - An *optical filter* on the light before the sensor
  - An *interpolation filter* on the image signal to produce RGB signal
- In the simplest case the latter filters are linear
  - May produce color aliasing
- More advanced cameras have non-linear filters to reduce color aliasing

The Digital Camera Lab Help: Bayer filters Exercise

A small part of a Bayer image is shown below, left, with a corresponding Bayer pattern, right. Compute the numerical values of *Gimage*, the resulting green (G) color plane, for the small part of the image.
Assume that pixels outside the small part of the image are zero.
Assume that the Bayer pattern repeats itself outside the small part of the image.

0	0	0	0	0
0	0	1	0	0
0	1	2	1	0
1	2	3	2	1
2	3	3	3	2

Bayer image

R	G			
G	B			

Bayer pattern

Gimage ?

Use normalized averaging with the interpolation kernel w shown below (center is marked in boldface).

$$w = \frac{1}{4}\begin{bmatrix} 1 \\ \mathbf{2} \\ 1 \end{bmatrix} * \frac{1}{4}\begin{bmatrix} 1 & \mathbf{2} & 1 \end{bmatrix} = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

The Digital Camera Lab Help: Bayer filters Solution

	0		0	
0		1		0
	1		1	
1		3		1
	3		3	

Gmaskimage

0	1	2	1	0
1	4	6	4	1
4	9	12	9	4
8	16	20	16	8
8	16	18	16	8

*Gmaskimage*w* /16

	1		1	
1		1		1
	1		1	
1		1		1
	1		1	

Gmask

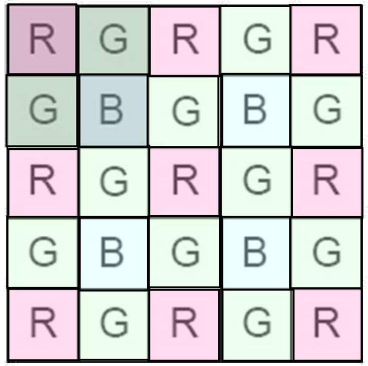
0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5
0.5	0.5	0.5	0.5	0.5

*Gmask*w*

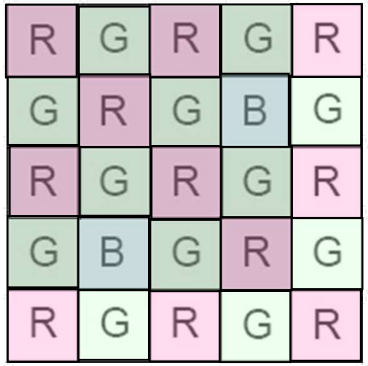
0	1	2	1	0
1	4	6	4	1
4	9	12	9	4
8	16	20	16	8
8	16	18	16	8

Gimage /8

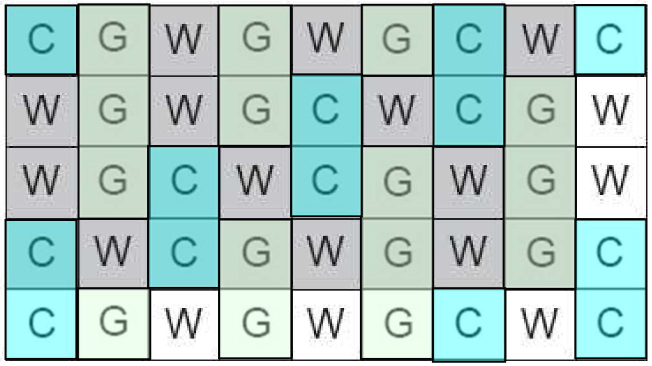
Bayer filters, other variants



R+2G+B
per cell



6R+8G+2B
per cell



12W+12G+8C
per cell

The dark area represents a
cell with a repeating pattern.

Another color filter type: Stripe filters

R	G	B	R	G
R	G	B	R	G
R	G	B	R	G
R	G	B	R	G
R	G	B	R	G

C	G	Y	C	G
C	G	Y	C	G
C	G	Y	C	G
C	G	Y	C	G
C	G	Y	C	G

R	G	B	G	R	G	B	G	R
R	G	B	G	R	G	B	G	R
R	G	B	G	R	G	B	G	R
R	G	B	G	R	G	B	G	R
R	G	B	G	R	G	B	G	R

Extra green

Easy to implement

May produce moiré effects due to color aliasing

Color filters, cont.

- Standard RGB-filters
  - Each color channel is rather narrow $\Rightarrow$ blocks more photons $\Rightarrow$ less effective
- Cyan-Yellow-Magenta-White filters
  - Magenta = red + blue
  - Cyan = blue + green
  - Yellow = red + green,
  - Magenta = red + blue (This may not be practical, as the wavelengths of red and blue light are far apart.)
  - White = red + green + blue
  - Each color channel is wider than the standard RGB $\Rightarrow$ blocks fewer photons $\Rightarrow$ more effective
  - Post-processing needed to convert to RGB
- The eye is more sensitive to green light and less to blue light
  - It makes sense to have more green detectors and fewer blue detectors

Color processing

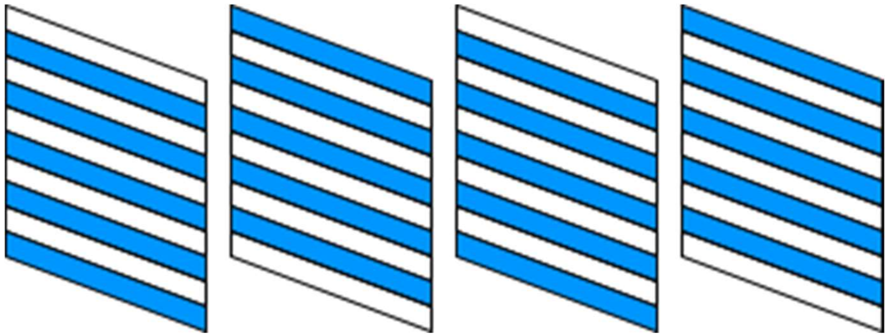
- The perception of color is complex
  - Humans tend to perceive color independent of illumination
  - A color camera measures physical quantities:
    very dependent on illumination
- White balancing:
  - Transforms the color measurements to make what we perceive as white have equal RGB-values
  - Automatic or manual
- The color information may also be converted to some other *color space* than RGB
  (e.g. HIS or XYZ)

The video camera

- Basic idea: take one image after another in sequence (temporal sampling)
- Older television standards (PAL, NTCS,…) require *interlaced video*
  - Take one half-image with all odd rows and then another half-image with all even rows, odd, even, etc

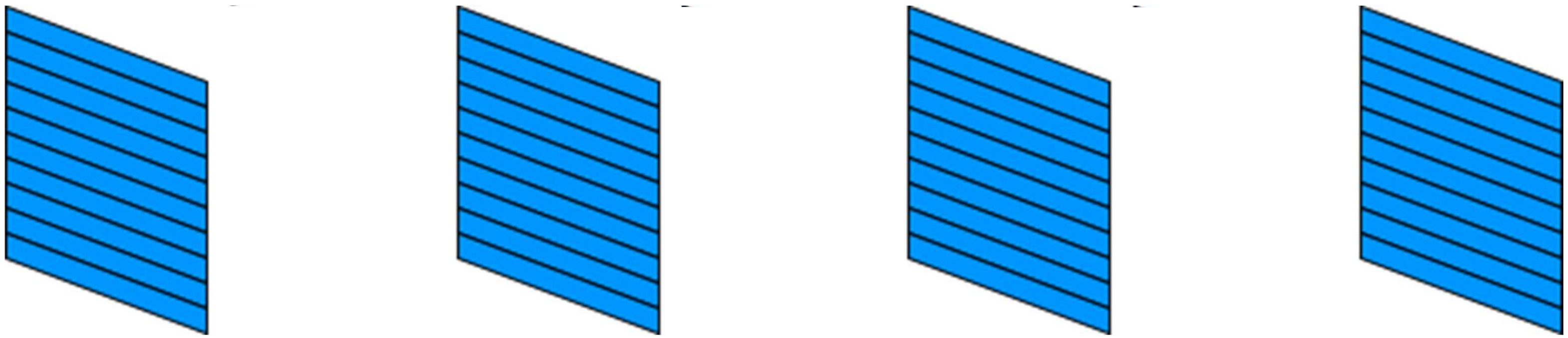
    => Odd and even rows are exposed at different times
  - Motivation: better bandwidth usage in broadcasted TV
- Today, *progressive scan* (or non-interlaced) video has replaced *interlaced video* to a large extent

The video camera: Interlaced vs. progressive scan



Interlaced scan

E.g., one “half image” at 50 Hz ⇒ one “full image” at 25 Hz

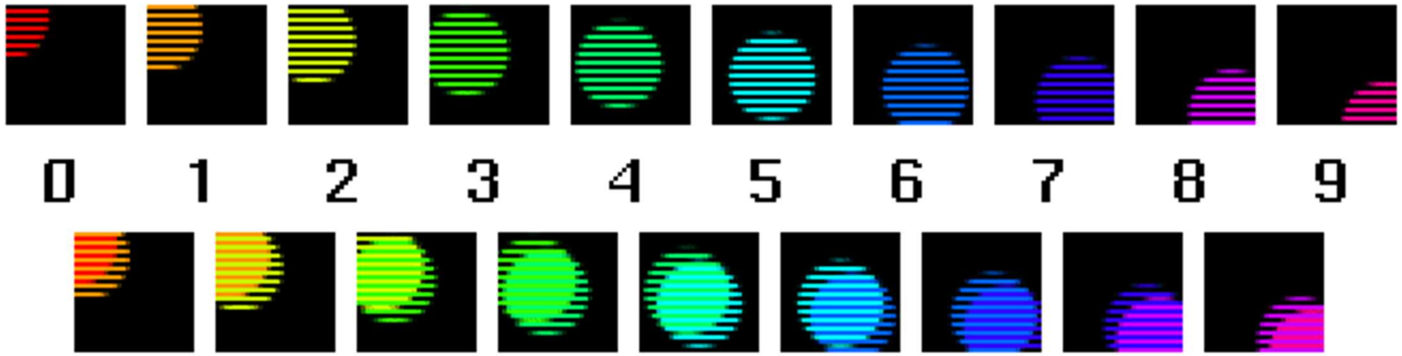


Progressive scan

E.g., one full image at 25 Hz

This may produce a flickering image.

The video camera: Interlaced vs. progressive scan



- Sometimes interlaced video (top) is represented as a sequence of ”complete” images, but the even and odd lines are taken at different time points (bottom)
- De-interlacing can be made by interpolation both spatially and over time
  - loss of spatial resolution

Shading correction

- Depending on the application and the sensor we may want to adjust the gain "*g*" and offset "*o*" of each pixel to assure that the resulting image is constant for a constant illumination. Takes care of
  - Vignetting (see next lecture)
  - Fixed pattern noise
  - Gradients in the illumination of the scene
- By projecting two different and constant illuminations into the camera, we can measure the individual deviations from a constant image in all pixels and compute adjustments of each pixel's gain and offset
- The shading correction is then made externally as part of the post-processing

Modern consumer cameras

- The effects described here relate to any type of light measuring digital camera
- Modern cameras (e.g., in mobile phones), however, include increasingly more and more sophisticated processing of the image and control of the camera
  - Automatic exposure time control
  - Automatic focus
  - Red-eye removal
  - Color balancing
  - Motion compensation
  - …
- These processes are not covered in this course