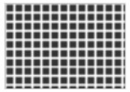
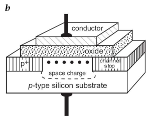
TSBB21, Lecture 2
Image Sensing

- Image sensing and the digital image sensor
- Two photo detectors, the photodiode and the MOS
- Thermal excitation and noise
- The readout problem, the CCD array and the CMOS camera
- Rolling or global shutter
- Blooming
- Fill factor and micro-lenses
- Noise sources
- Shot noise (or photon noise, or Poisson noise, or Quantum noise)
- The output voltage
  - Fixed pattern noise
  - Signal-to-noise ratio
  - Dynamic range
- Color cameras
  - 3 chip cameras
  - 1 chip camaras
  - Bayer filters
- mm
- 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.

Image sensing

- In the previous lecture we saw how light reflected (or emitted) from objects in a 3D scene is projected onto the image plane of a camera
- In this lecture we will see how this image is sensed to produce a digital image
- Detailed physics/electrical explanation about image sensors is in a lecture next week
- Main method
  - The image is spatially sampled and truncated 
  - Photons are converted to electric charge/voltage 

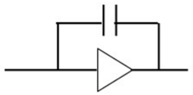
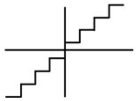
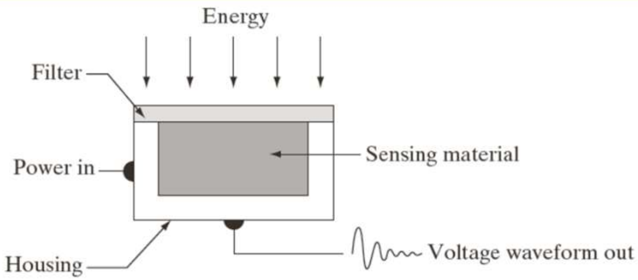
  - The charges are converted to voltage 
  - The voltage is quantized (A/D-converted) 

Figure from:
Gonzales & Woods

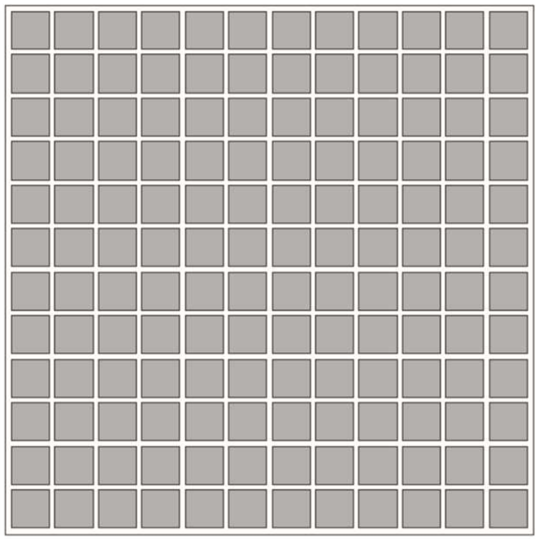
The digital image sensor

Single sensor element:



Line sensor:

Array sensor:



The digital image acquisition process

Figure from: Gonzales & Woods

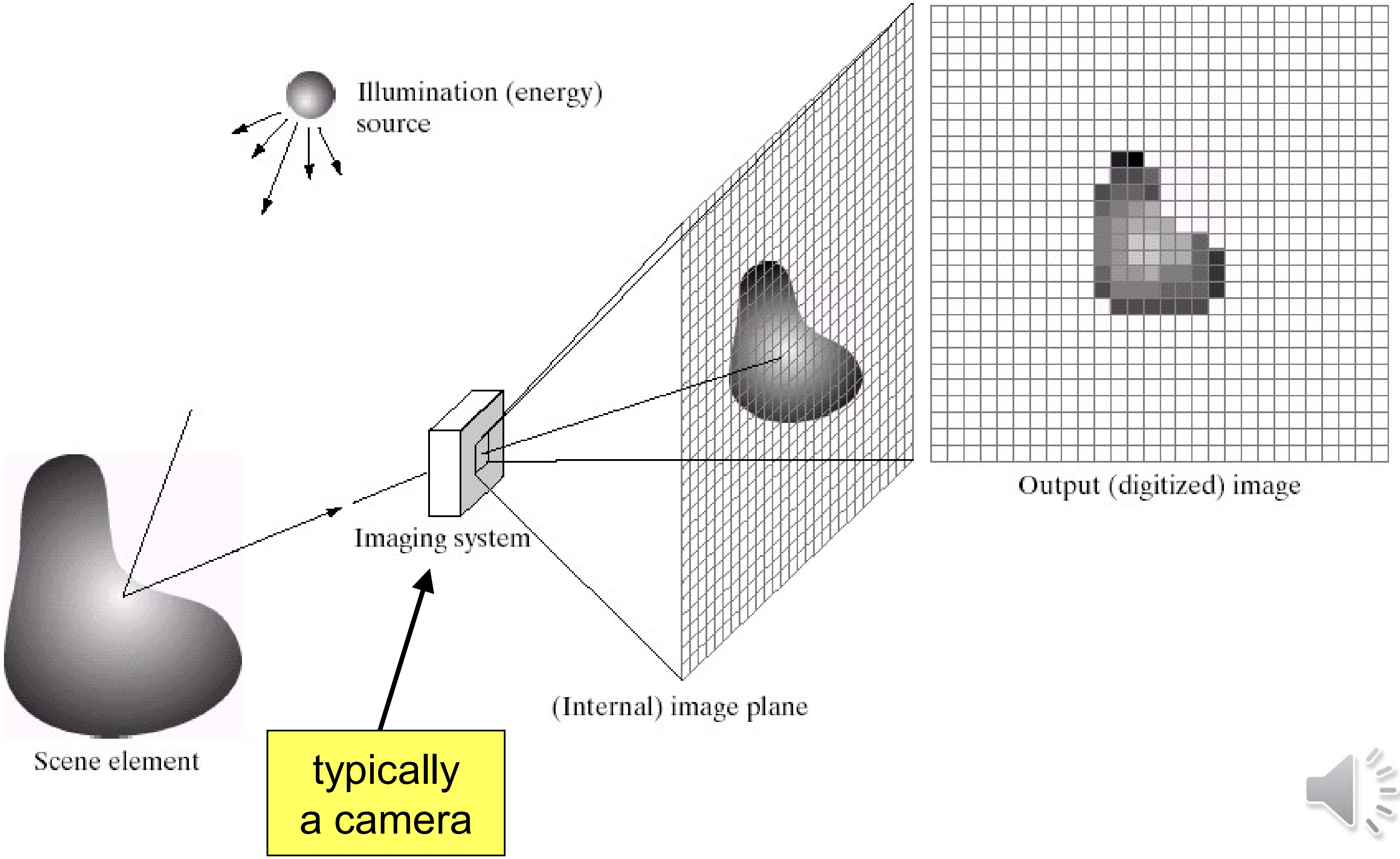
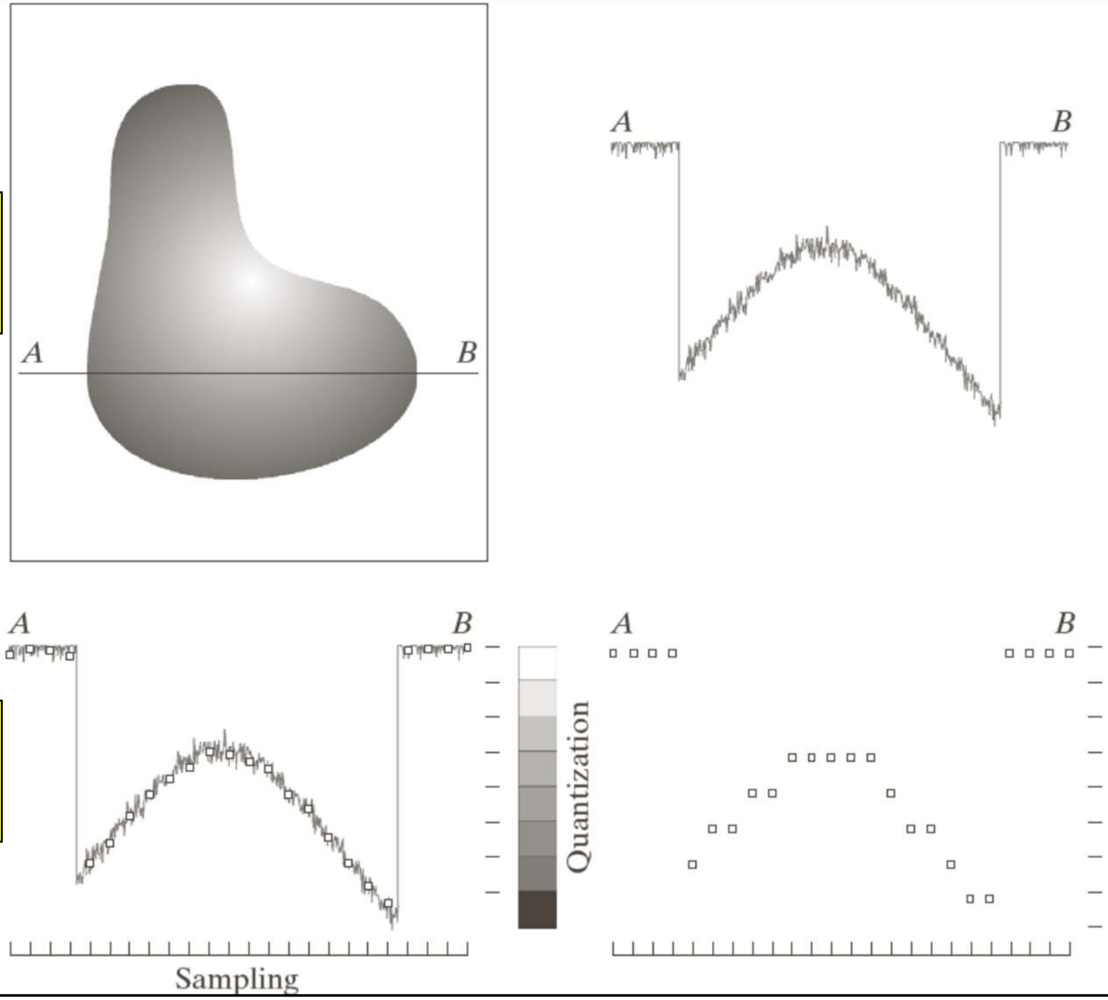


Image sampling and quantization

Figure from:
Gonzales & Woods

Continuous Image

Line from image



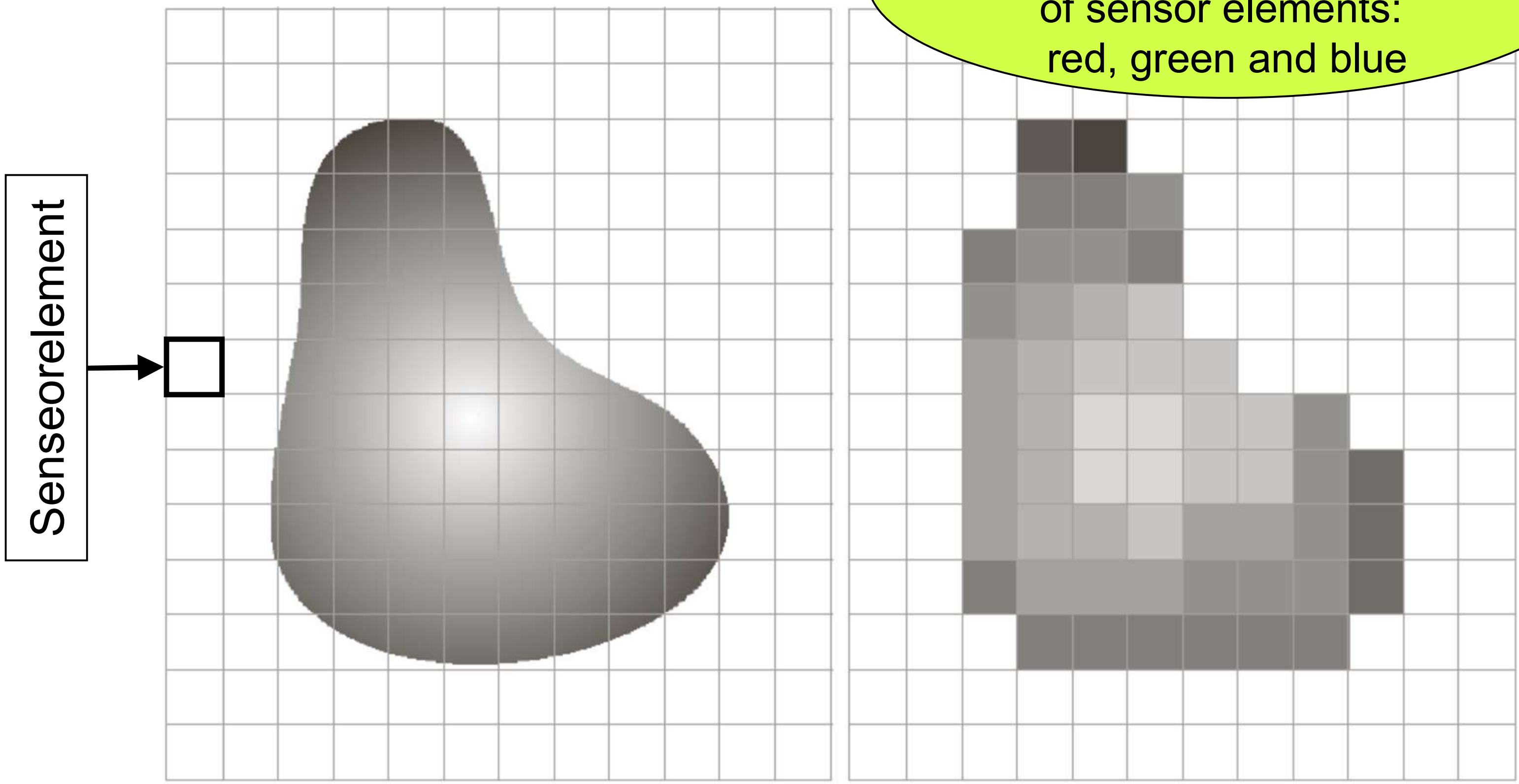
After sampling

After quantization

The digital image sensor

Figure from: Gonzales & Woods

For a color image, there are 3 different types of sensor elements: red, green and blue

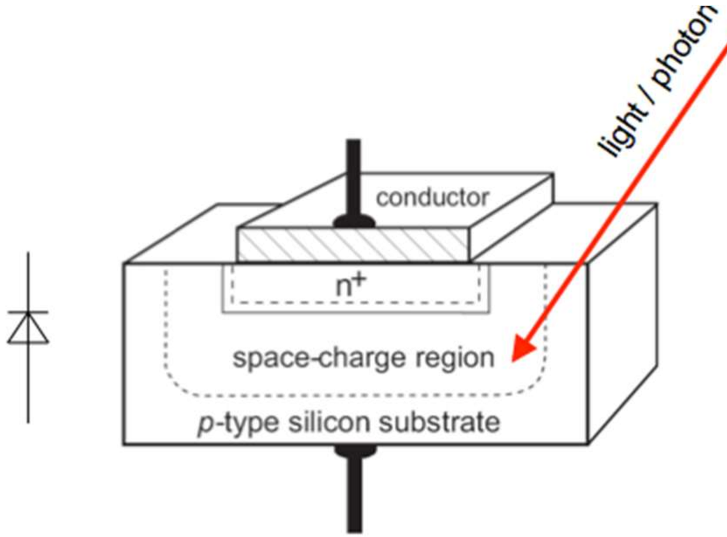


A continuous image projected onto the sensor array	Result after sampling and quantization

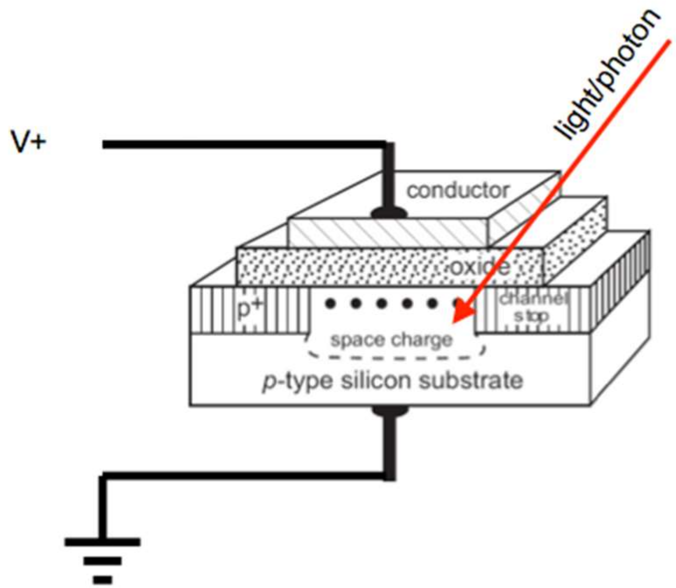
Two main types of photo detectors

- Most image sensors are based on either of two distinct types of photodetectors
  - The **photodiode** (the photovoltaic effect)
  - The **MOS capacitance** (intrinsic absorption)
  - MOS = metal-oxide semiconductor
- Both can be manufactured using standard semiconductor processes
- Both can be seen as a capacitor which is discharged/charged by means of I_{photo} (and I_{thermo}, see 2 slides ahead)

Two main types of photo detectors



Photodiode



MOS

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}

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

  - Q_g is the gap between the material's valence and conduction bands
  - $\mathcal{E}$ 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}
- We will return to the noise issue later

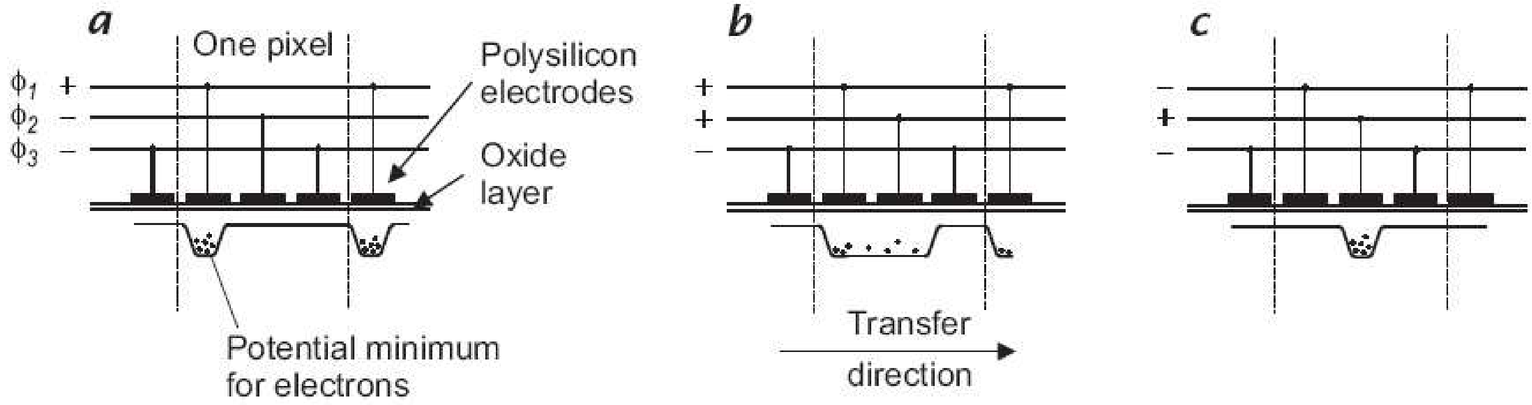
The read-out problem

- Light has caused a change in electric voltage or charge in a light detector element (photodiode or MOS capacitor), 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** (MOS capacitor only) (the traditional, older approach)
    CCD = Charge Coupled Device
  - **Switches** to a common signal/video line
    (photodiode or MOS capacitor) (increasingly more common!)

The CCD array

(Now used more and more seldom)

Nobel Price in 2009!

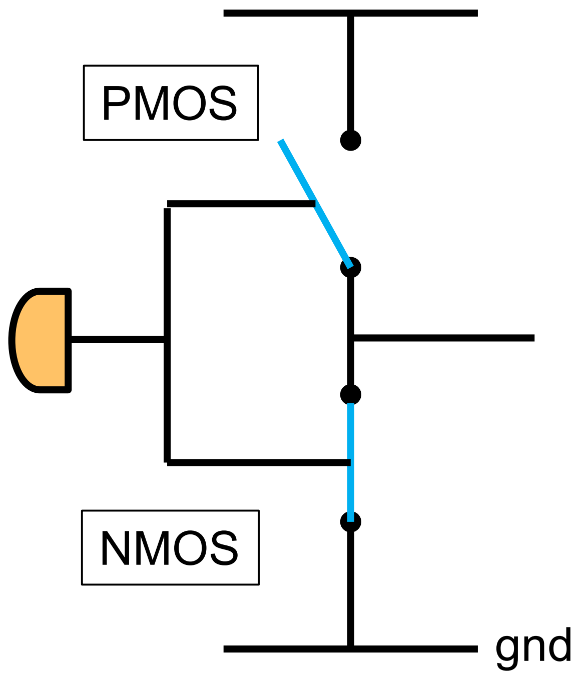


- A chain of MOS capacitors where the voltages change in the pattern shown above can “move” the charge
- This transport can take place along an entire row/column of a detector array
- One pixel = 3 capacitors
- At the end, a charge-to-voltage translation is done (charge amplifier)

CMOS

- CMOS = complementary MOS
- MOS= metal-oxide semiconductor
- Both PMOS (P-doped MOS) and NMOS (N-doped MOS) are included
- Closing/breaking contact
- Consumes no power

CMOS illustated

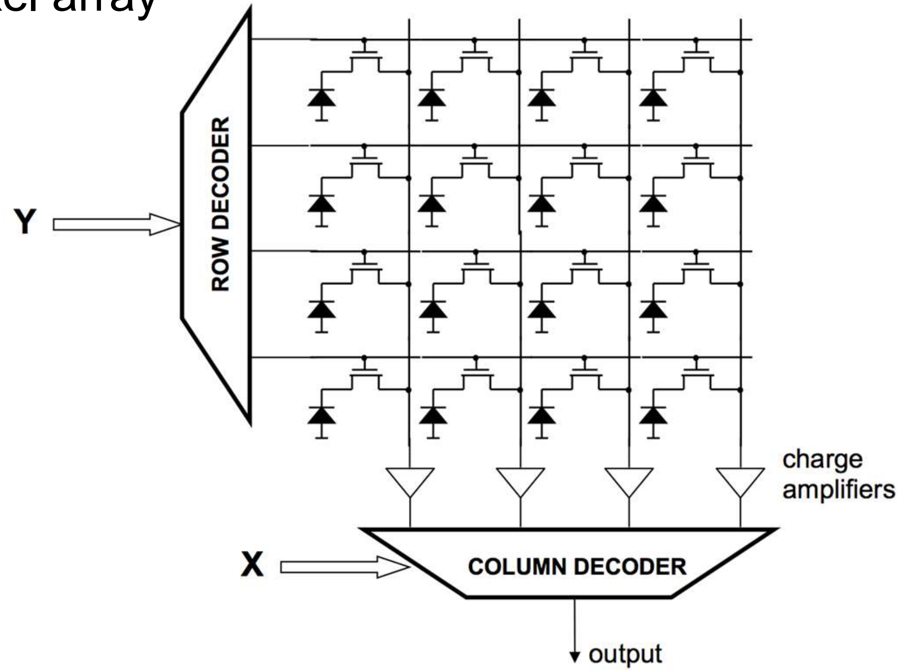


Photodiode arrays (CMOS cameras)

- The photo diode charges are read-out one row at a time via parallel column bus lines
- CMOS devices without individual pixel amplifiers are called *passive pixel sensor* (PPS)
- Example: 1-transistor (1T) pixel array
  - Column amplifiers convert charge to voltage

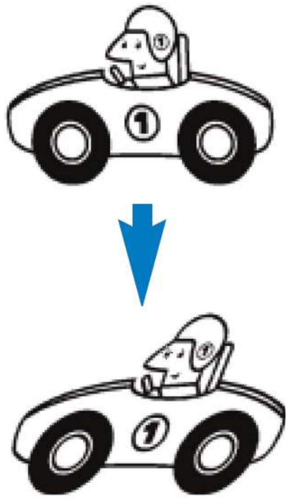
For each addressed row, the photodiode charges are translated into voltages using charge amplifiers.

The column decoder outputs one column at a time.



Rolling or global shutter

- For a CMOS camera, 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.
- Called: *rolling shutter*. Alternative: *global shutter*
- For a stationary scene, rolling shutter is OK
- If the camera or the scene is moving, rolling shutter may distort the image
- This is called:
  - Rolling shutter problem
  - Jello/jelly effect
  - CMOS distortion
- CCD usually use a global shutter



Jello effect/ CMOS distortion example

- CMOS distortion image found on the Internet

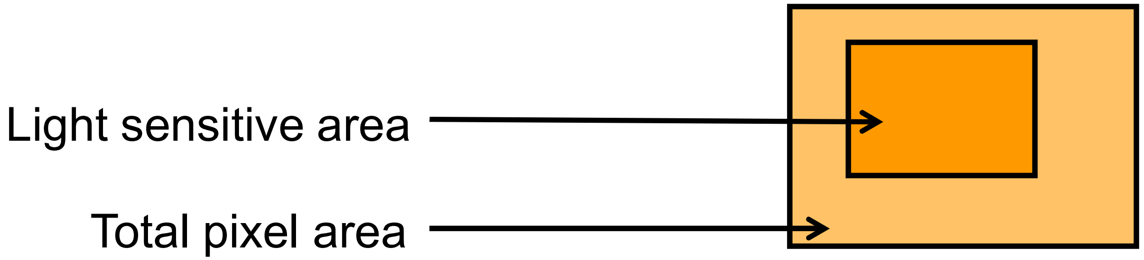
Blooming

- Both the photodiode and the MOS capacitor "collect" electric charge in a small region corresponding to the conductor region
- When this region becomes saturated, the charge spills over to neighboring elements
- This is called *blooming*
- Barriers between the detectors can reduce this effect, but not eliminate it entirely

Image illustrating blooming found on the Internet:

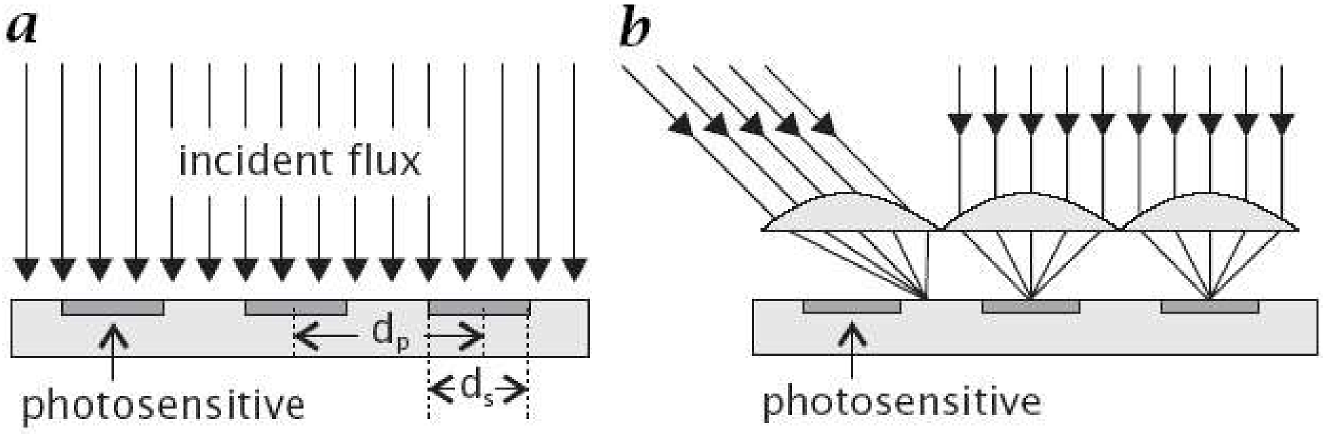
Fill factor

- 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



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

Noise sources

- Reset noise
  - The measured voltage depends on the “fix” bias voltage over the photo diode or MOS capacitor
  - 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

Noise sources

- The space-charge region is not a perfect isolator => there is a small leakage current
  - Called *dark current* since it discharges the capacitor even when no photons are absorbed
- Thermal noise (explained in two earlier slides)
  - Can be reduced by cooling
- Design noise effect: blooming, after-effects

Shot noise (or photon noise, or Poisson noise, or Quantum noise)

- Even if a constant number of photons hit the photodetector, the absorption process is a probabilistic phenomenon:
  - Each time we observe/measure the voltage/charge difference at the detector, there will a small variation in the result
  - This variation is larger the shorter the exposure time is, and vice versa
  - This noise has approximately a Poisson distribution
  - Known as *shot noise, photon noise, or Poisson noise or Quantum noise.*

Poisson noise

- Probability mass function:

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

- Mean and variance:

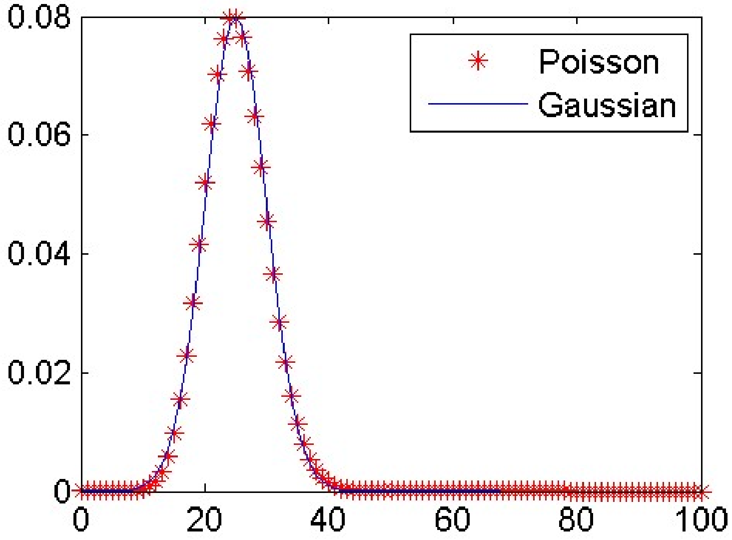
$$\mu_N = a \quad \text{and} \quad \sigma_N^2 = a$$

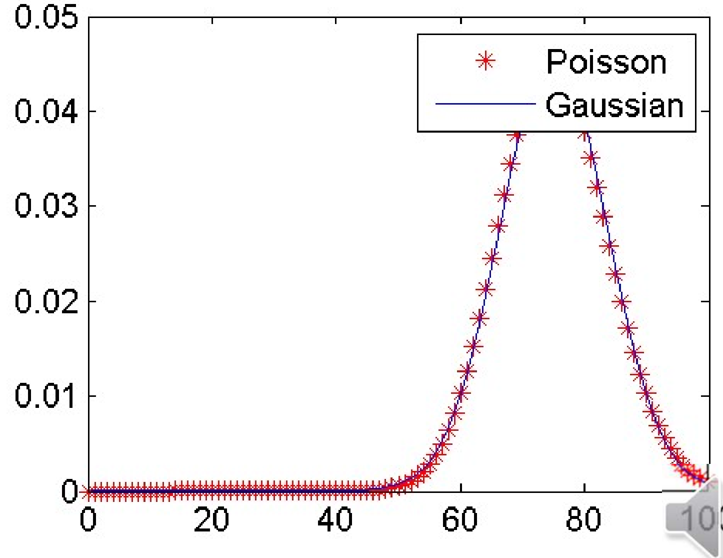
- Is used to statistically characterize the distribution of photons per unit of area.
- Important for normal camera images, and medical image systems using X-rays and gamma-rays.

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

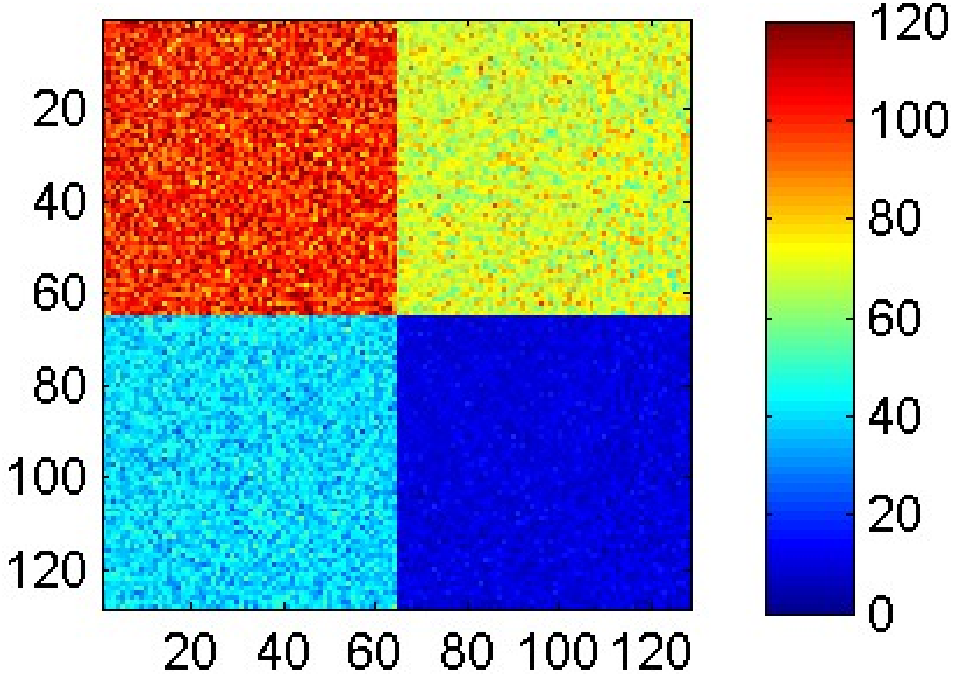
$$\mu_N = \sigma_N^2 = a$$





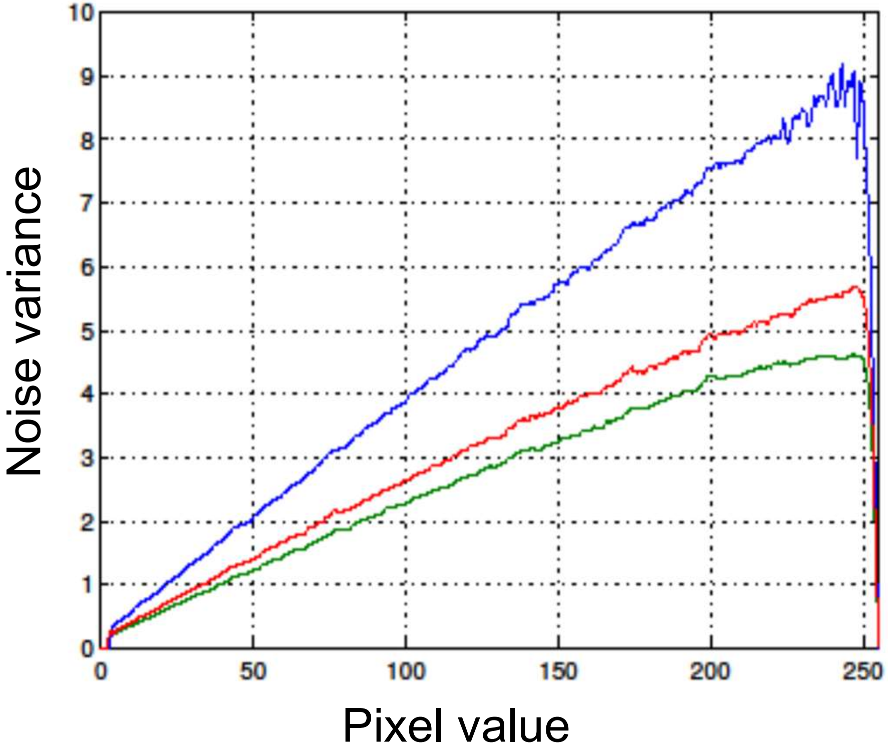
Poisson noise

- The standard deviation
  $$\sigma_N = \sqrt{a}$$
- increase with the signal value
  $$\mu_N = a$$
- But the signal to noise ratio
  $$a/\sqrt{a} = \sqrt{a}$$
  improve!



Poisson noise

- The noise variance depends linearly on the image intensity (lab exercise)



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, blooming, after-exposure, etc

Fixed pattern noise

- 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*: $g \approx 0$

Signal to noise ratio

- ΔV = The overall noise voltage measured at the output
- V = the actual output voltage

$$\mathrm{SNR} = 20 \cdot {}^{10}\log \frac{V}{\Delta V}$$

- Darker images have a lower SNR than brighter images

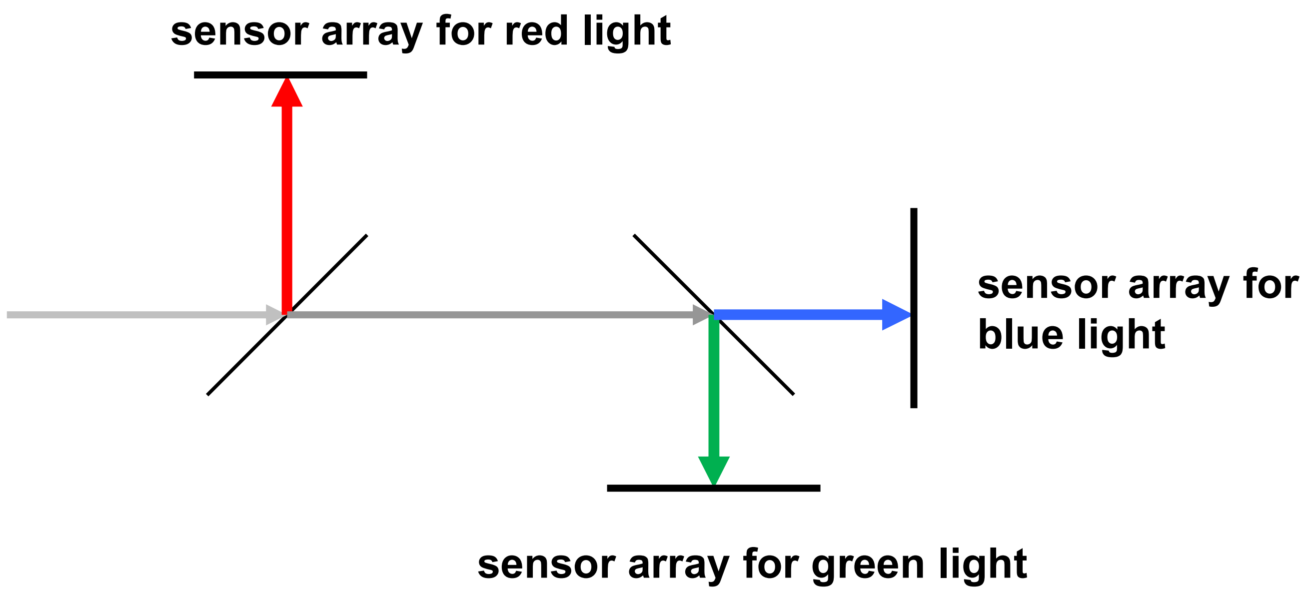
Dynamic range

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

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

- Typical values
  - CMOS: 40-60 dB
  - CCD: 60-70 dB
  - Human eye: > 90 dB

3 chip color cameras



3 identical standard chips

2 semi-transparent mirrors that refract different wavelengths

3 chip color cameras

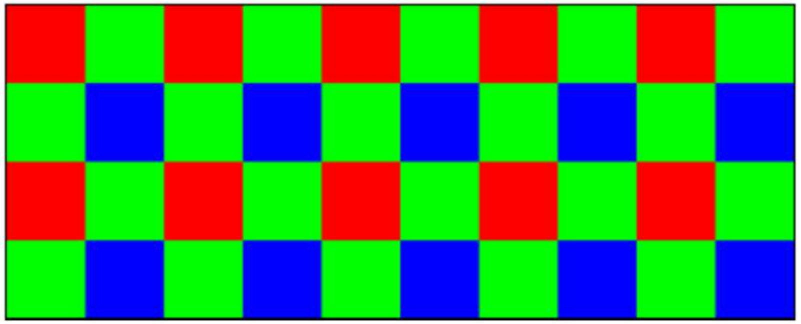
- Based on standard “black-and-white” sensor chips (3 identical sensor chips)
- The 3 sensor arrays need to be aligned with tolerances smaller than the inter-pixel distance
- Gives good performance
  - Is expensive
  - Only used in professional cameras

1 chip color cameras

- To reduce cost:
  - 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
    - May give color-aliasing

Bayer filters

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



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

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).

$$w = \frac{1}{4}\begin{array}{|c|}\hline 1 \\ \hline \mathbf{2} \\ \hline 1 \\ \hline\end{array} * \frac{1}{4}\begin{array}{|c|c|c|}\hline 1 & \mathbf{2} & 1 \\ \hline\end{array} = \frac{1}{16}\begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 2 & \mathbf{4} & 2 \\ \hline 1 & 2 & 1 \\ \hline\end{array}$$

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

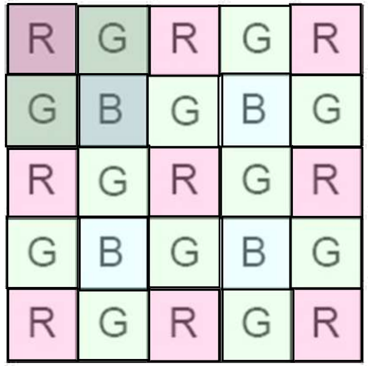
Gimage

/8

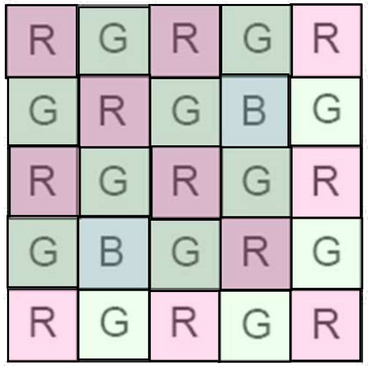
38

Bayer filters, other variants

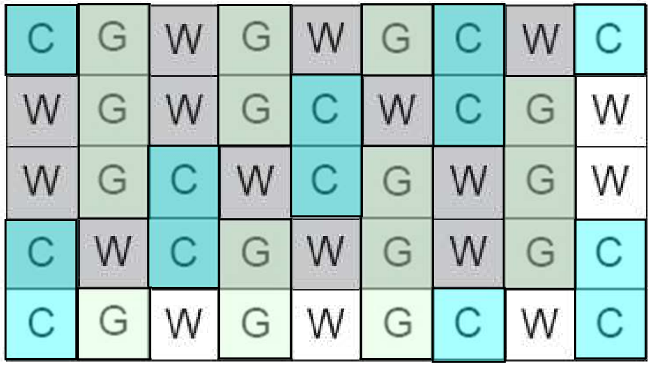
Dark area is a cell that represents one ”pixel”



R+2G+B per cell

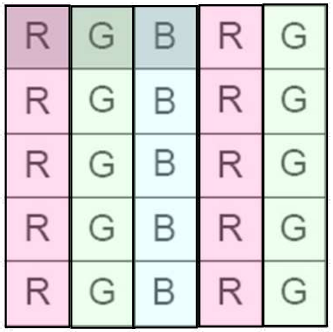


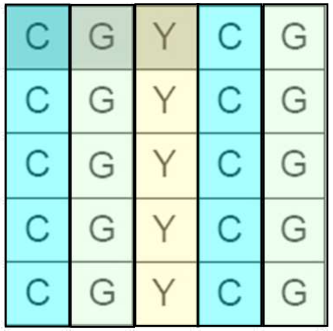
6R+8G+2B per cell

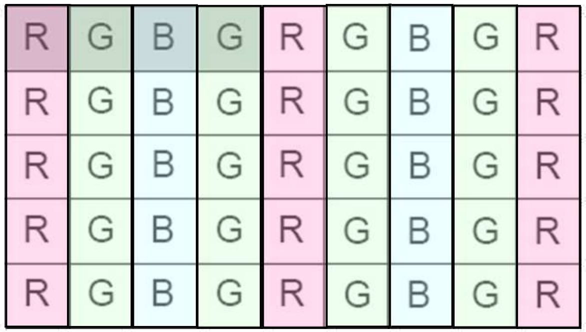


12W+12G+8C per cell

Stripe filters, an alternative to Bayer filters







Extra green

Easy to implement

May produce moiré effects due to color aliasing

1 chip color cameras

- 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, White = red+green+blue
  - Each color channel is wider than 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

1 chip color camera

- 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

Color processing

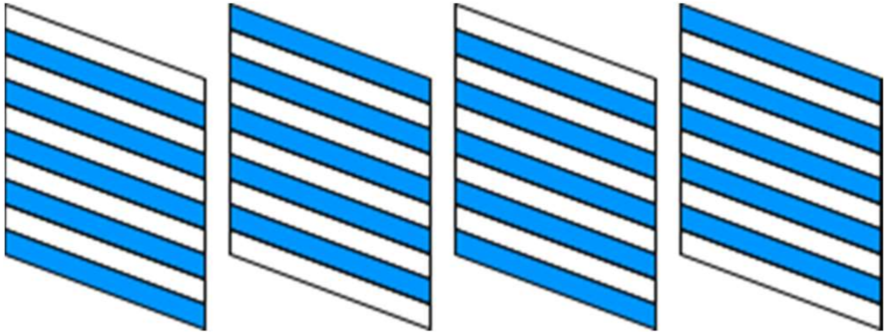
- The perception of color is complex
  - Humans tend to perceive color independent of illumination
  - A color camera makes a measurement of physical quantities: very dependent on illumination
- White balancing:
  - Transforms the color measurements to make what we perceive as white have equal RGB-values
  - 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)
- Legacy 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

Interlaced vs. progressive scan

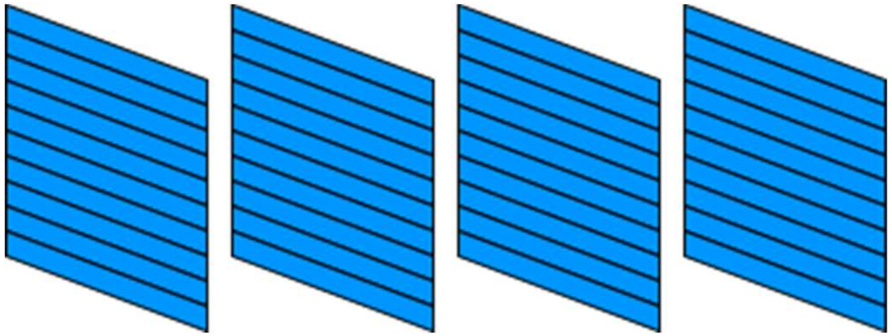


Interlaced scan

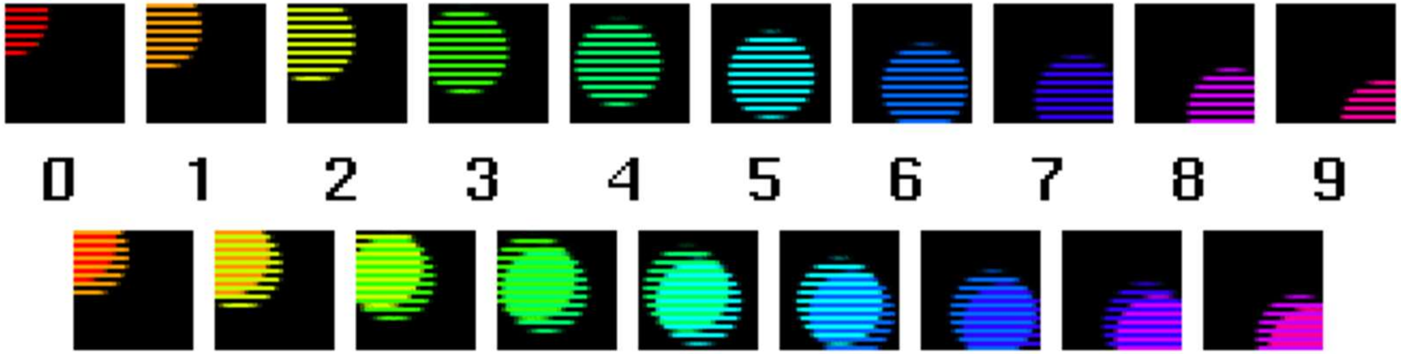
E.g., one “half image” at 50 Hz ⇒ one “full image” at 25 Hz

Progressive scan

E.g., one full image at 25 Hz



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
$\Rightarrow$ 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 a lecture next week)
  - Fixed pattern noise
  - Gradients in the illumination of the scene

Shading correction

- 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