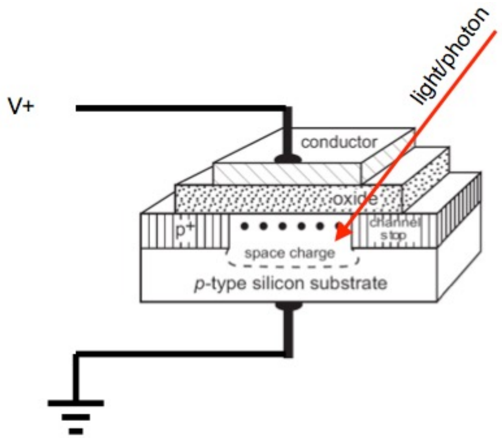
TSBB21 Image Sensors

Image Sensing

Part 2

Robert Forchheimer

Two different pixel types



MOS capacitor

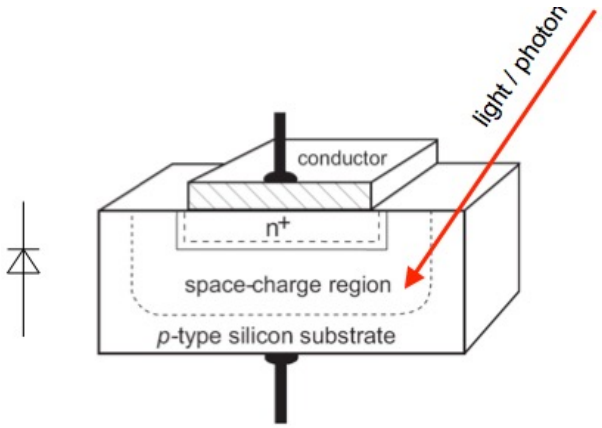


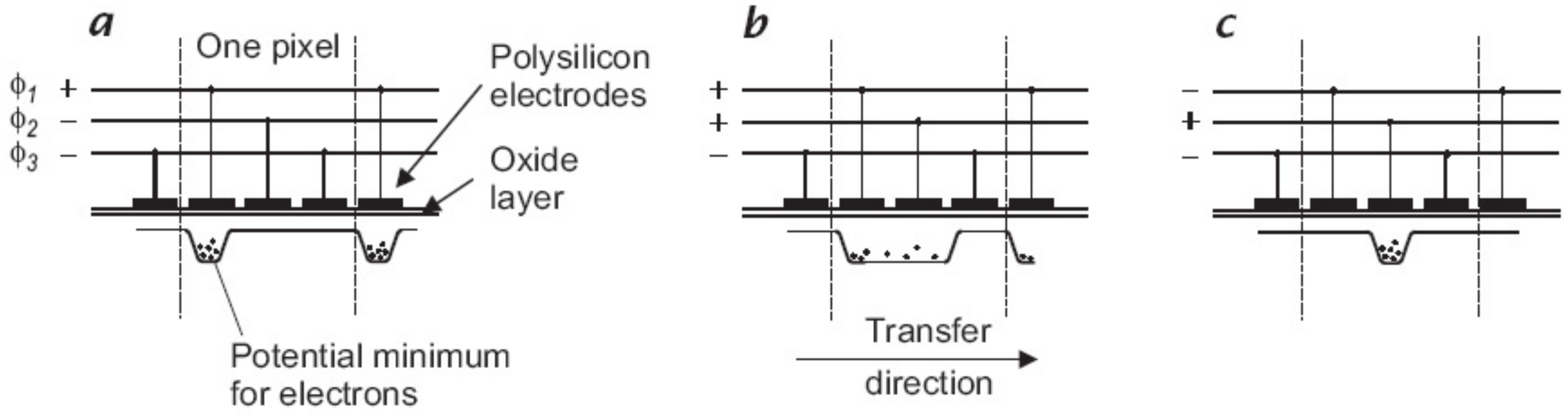
Photo diode

The read-out problem

- Light has caused a change in electric voltage or charge in a light detector element (photo diode 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) (traditional approach)
  - **Switches** to a common signal/video line
    (photo diode or MOS capacitor) (increasingly more common!)

The CCD array

CCD = Charge Coupled Device



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

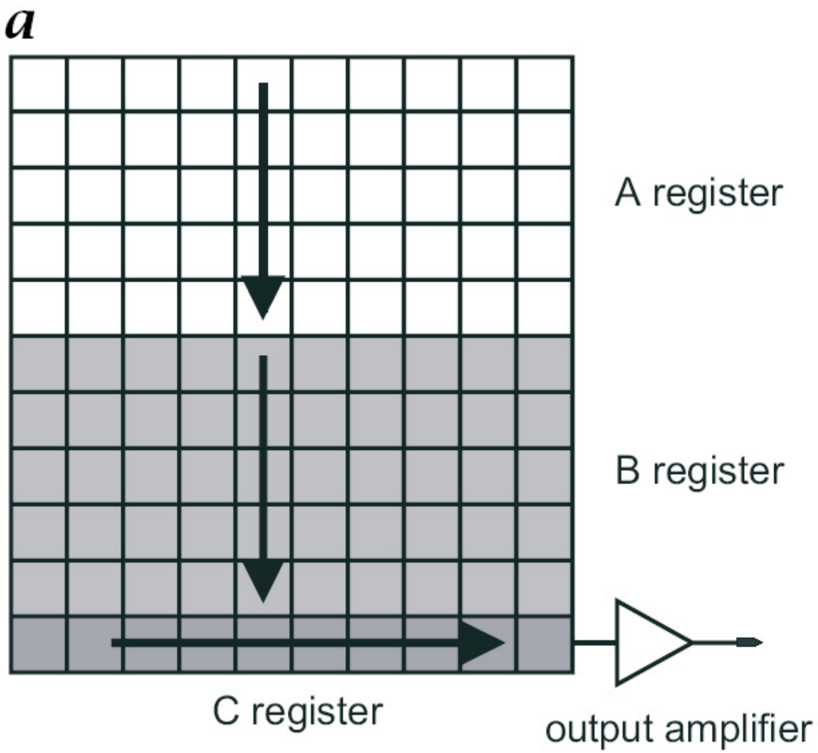
One pixel = 3 capacitors

At the end, a charge-to-voltage translation is done (*charge amplifier*)

CCD architectures

- A CCD array can have different ways of implementing the readout of an entire image
- Limiting factors:
  - there is a maximal readout frequency from the entire array
    - this limits the readout speed from the individual pixel
  - the MOS-capacitors are sensitive to light exposure during the transport
    - Charges should be moved to light insensitive areas as quickly as possible

Frame-transfer CCD



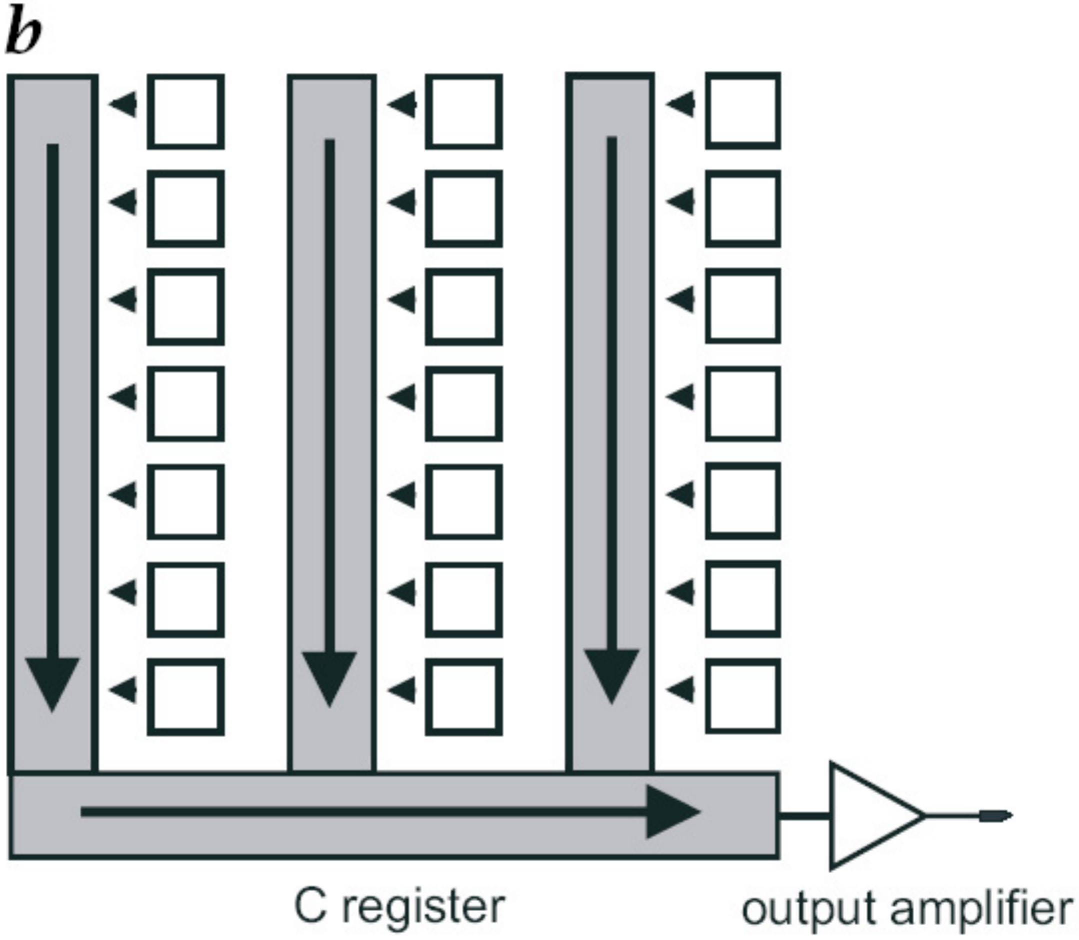
Active sensor area

Covered with an opaque metal shield

Frame-transfer CCD

- Advantages
  - The whole of area A is light sensitive, fill factor can be close to 100%
  - Simple to manufacture
- Disadvantages
  - It takes some time to shift the entire image from A to B, during this time area A is still sensitive to light, leading to *after-effects*
  - Mechanical shutters can be used to remove after-exposure

Interline-transfer CCD

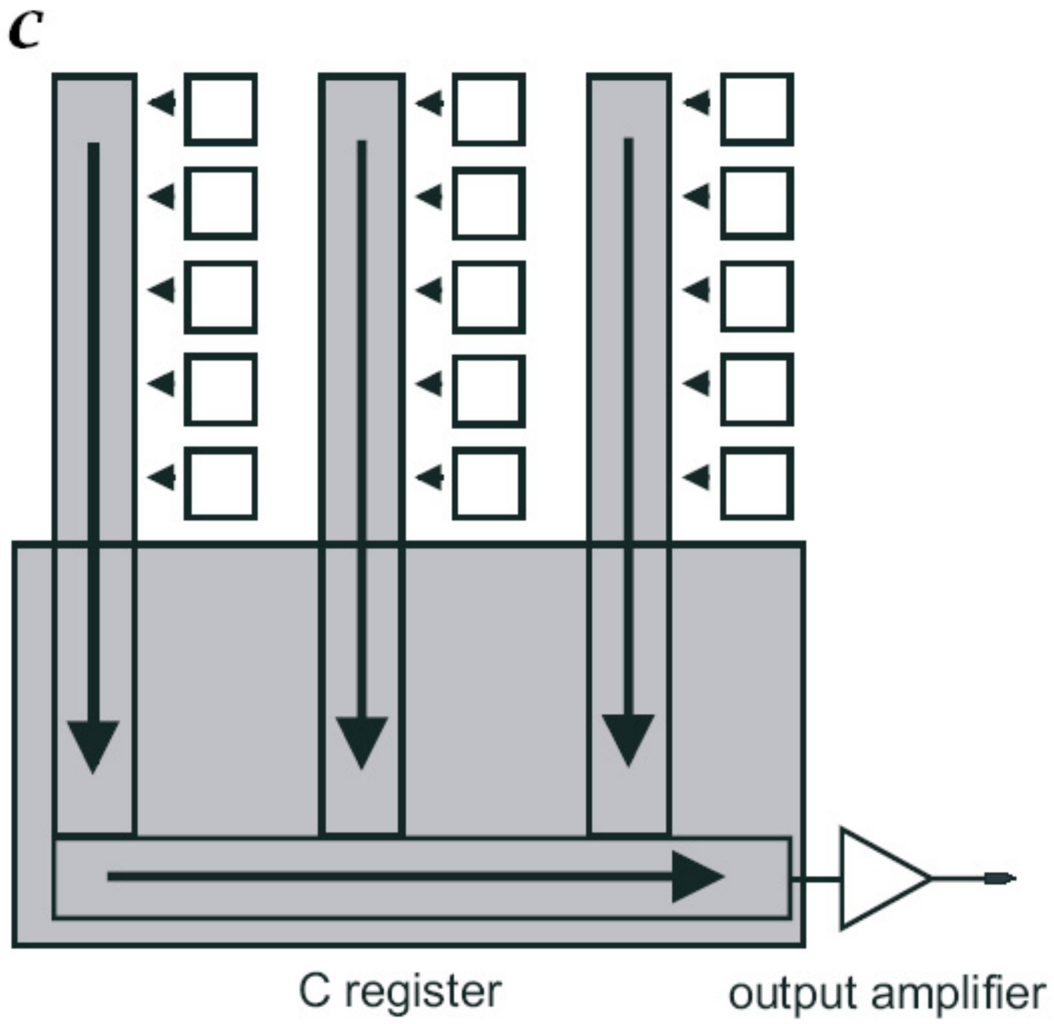


Gray areas are covered with an opaque metal shield

Interline-transfer CCD

- Advantages
  - Small after-effects
- Disadvantages
  - Reduced fill factor due to the vertical shift columns. Common fill factor ≈ 30%
- During shifting the transport column may still be affected by small amount of photo-charge.
- This approach is the most common in today’s CCD-based consumer cameras
- Not OK for studio cameras

Field-interline-transfer CCD



Gray areas are covered with an opaque metal shield

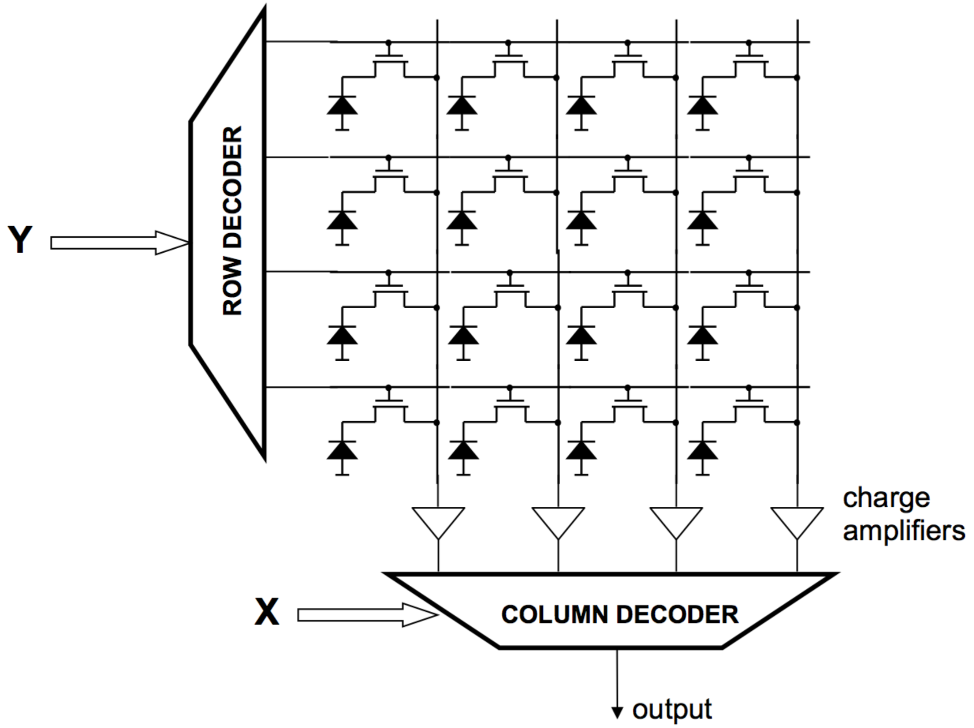
Field-interline-transfer CCD

- Each transport column is quickly shifted into the completely opaque area from which it is shifted out in slower pace via the C register
- Very small after-effects
- More complicated manufacturing
  = more expensive
- Mainly used for studio cameras

Photo diode 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

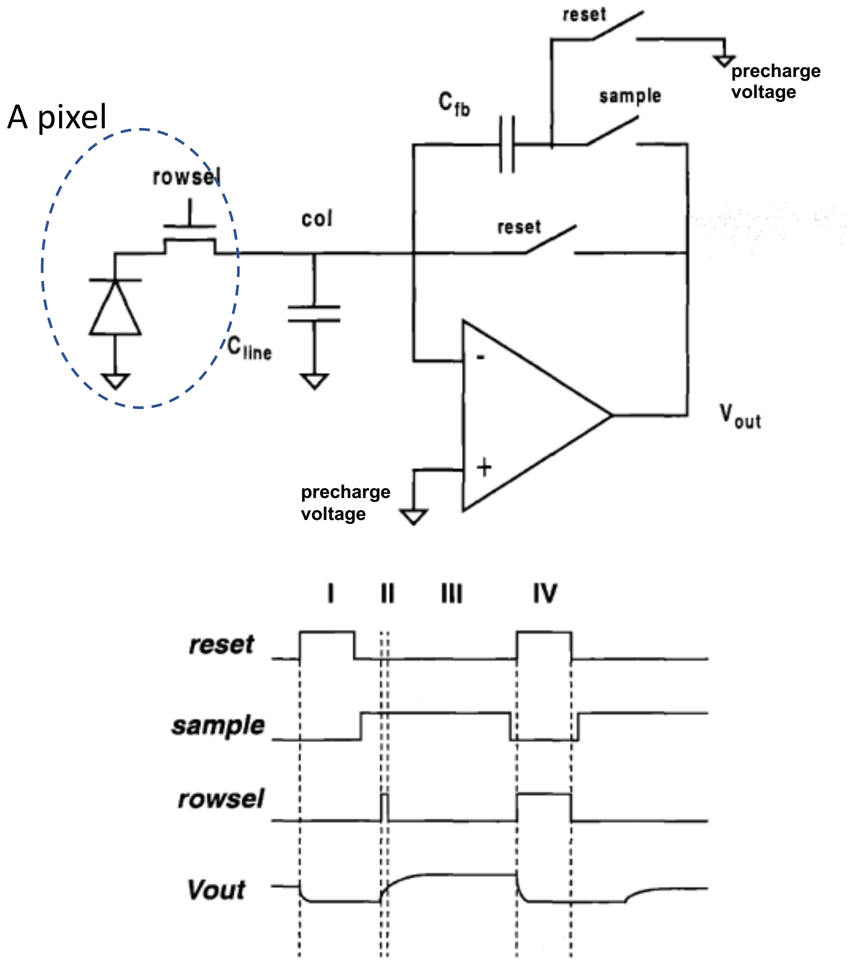
For each addressed row, the photo diode charges are translated into voltages using charge amplifiers. The column decoder outputs one column at a time.



Column amplifier for 1T pixel array

Proper control of the switches leads to:

- **The selected photo diode is charged via the output bus.**
- **Sensing is offset-compensated**
- **Column bus stays at a constant voltage which reduces the influence of its high capacitance (C_{line})**



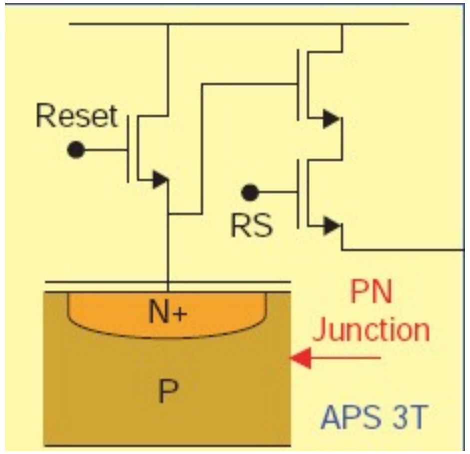
Active pixel sensors (APS)

- 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
    - This requires an additional pre-charge transistor
  - 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

3T APS

- The 3 transistor (3T) APS
  - One for recharging the diode
  - One used as voltage buffer
  - One for connecting the voltage to the readout line
- Less sensitive to noise compared to PPS



4T APS

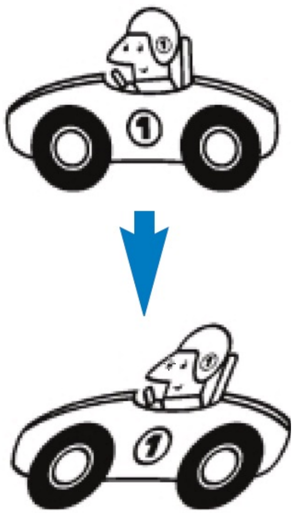
- Add a fourth transistor to each pixel (4T)
- This transistor acts as a “memory” during readout
    - All photo-charge is moved globally to the memory transistor after exposure
    - The other three transistors operate as a standard 3T APS
    - Can implement a global shutter read-out

CCD vs. CMOS

- Traditionally: more components and capacitance in the CMOS readout path => more noise
- CMOS: high-speed readout and support for window-of-interest. Low power. Can easily integrate with other electronics even on pixel scale.
- CCD: manufactured in specialized technologies
  - Easier to scale down pixel size
  - More difficult to integrate with standard electronics on the same ship
- CMOS has now reached a quality level where they can rival CCD even in professional applications.

Rolling or global shutter

- For a CMOS camera, the simplest approach is to make the exposure and read-out line by 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 video

- Jello effect video on YouTube:

  http://www.youtube.com/watch?v=UEaDrS-yzIE

CMOS distortion image found on the Internet

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 discharge the capacitor even when no photons are absorbed
- Thermal noise
  - Can be reduced by cooling
- Design noise effect: blooming, after-effects

Shot noise

- Even if a constant number of photons hit the photo detector, the absorption process is a probabilistic phenomena:
  - 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*

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*: $g \approx 0$

The output voltage

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

$$V = g \cdot I + o + \Delta V$$

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

Signal to noise ratio

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

$$\boxed{\mathrm{SNR} = 20 \cdot {}^{10}\log \frac{V}{\Delta V}} \qquad \text{(dB)}$$

- Means: darker images have a lower SNR than brighter images
(assuming constant average noise)

Dynamic range

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

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

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

Image Sensing

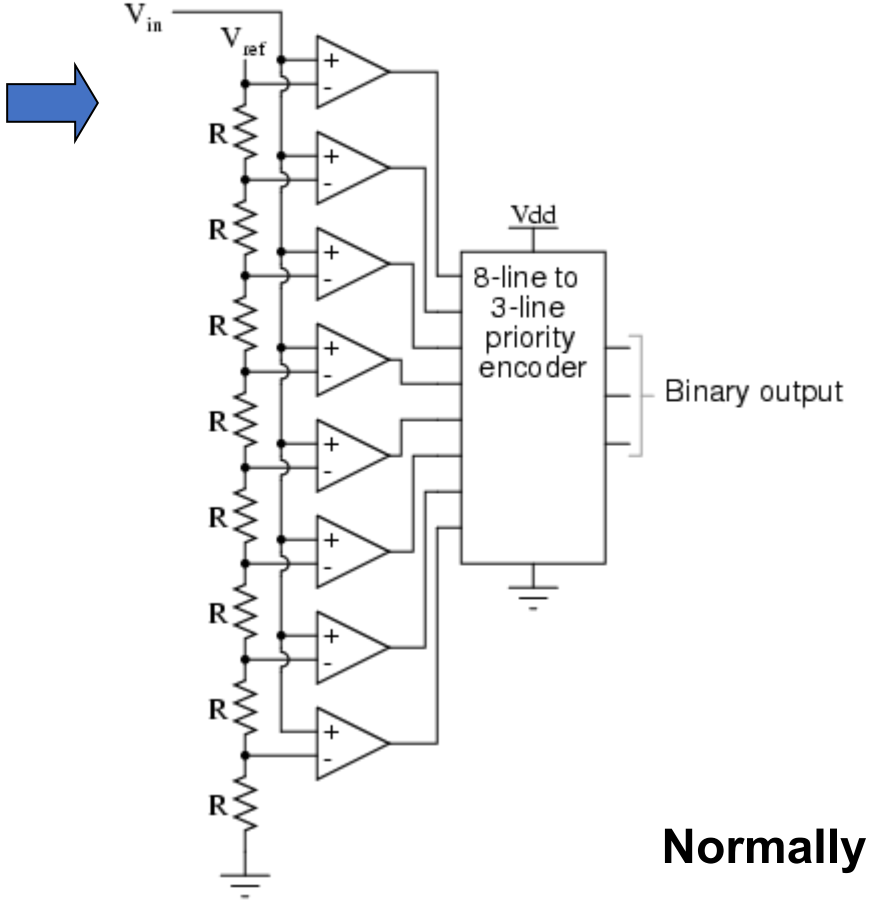
Part 3

Digitalization

- The analogue voltage signal is normally transformed to a digital representation by means of an analog-to-digital converter (ADC)
- Two common principles:
  - Flash ADC (up to 8 bits)
  - Successive approximation (>8 bits)
- Other methods exist

Flash ADC

Example:
3-bit converter



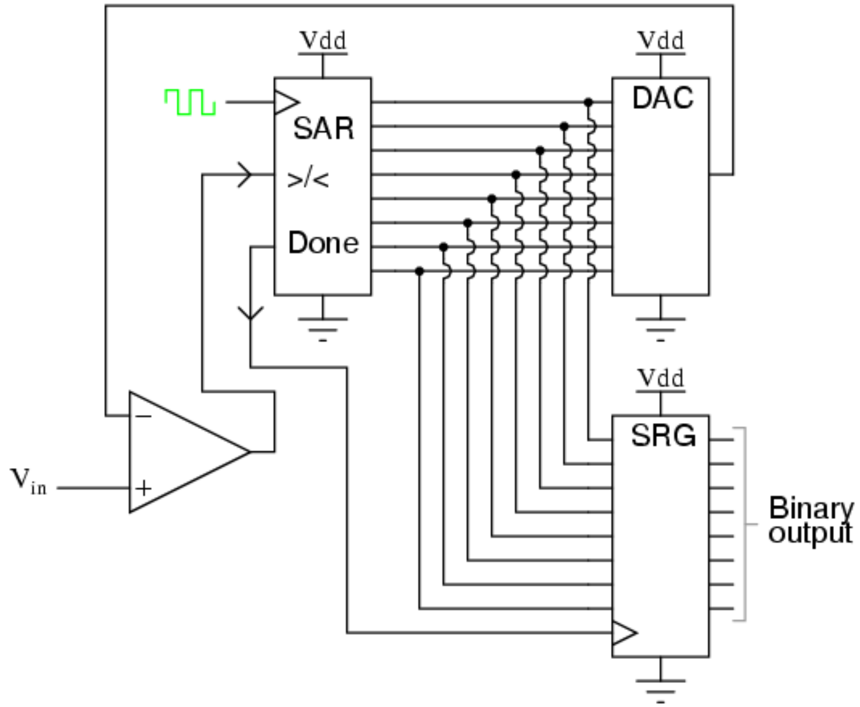
Fast and simple

Normally used up to 8-bits

Successive approximation ADC

- The values of the resulting digital output are determined successively, from MSB to LSB

- **More complex circuitry**
- **Slower than flash ADC**
- **Can manage more output bits**



Digitalization

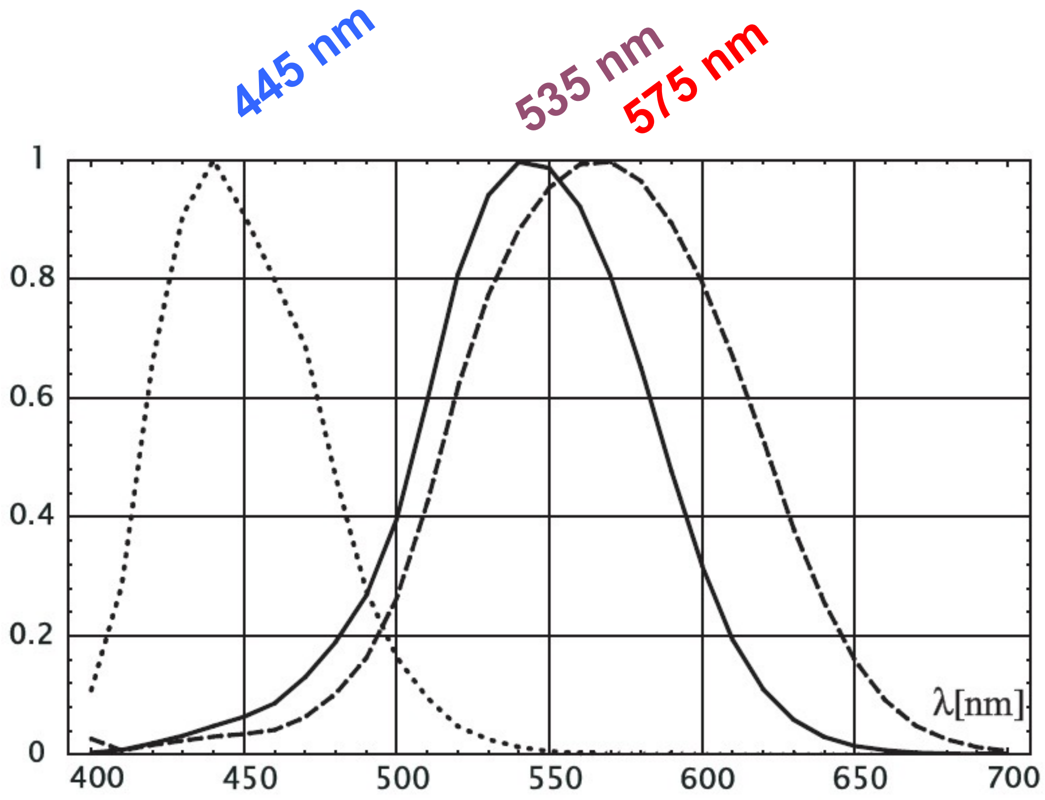
- Independent of method: *quantization noise*
- If b bits are used to represent voltage up to V_{max}:

$$\boxed{b = \log_2 \frac{V_{max}}{\Delta V} = \frac{DR}{20 \cdot \log_2 2}}$$

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

Color vision

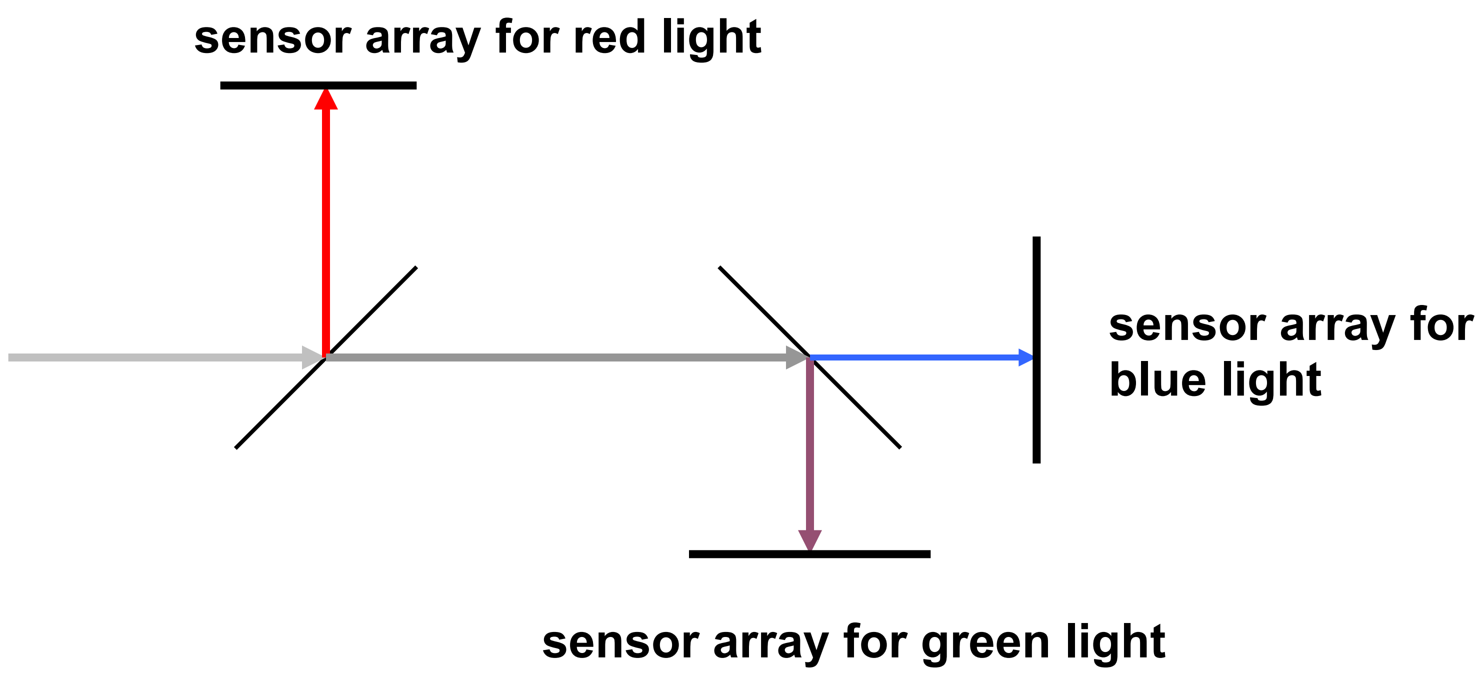
- The human eye has cones which are sensitive to different wavelength bands



Grassman’s law

- It is (in principle) sufficient to measure the light spectrum in three distinct wavelength bands to represent any perceivable color
- We do not have to use a sensor with sensitivity curves identical to the human eye
- The 3 color channels are often called “red”, “green”, “blue”, even though they don’t correspond to the eye’s curves

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

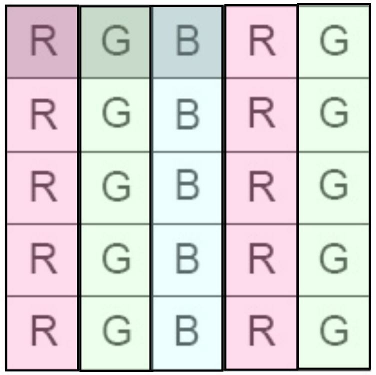
- To reduce cost:
  - Use one sensor array
  - Place a color filter on top of each detector element
  - Each detector element is now sensitive to only a specific wavelength range
  - Reduces the fill factor for each range
  - The colors are not measured at the same places
    - color-aliasing

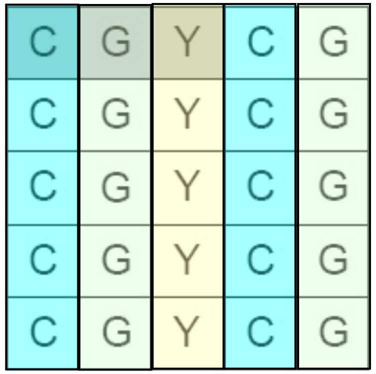
1 chip color cameras

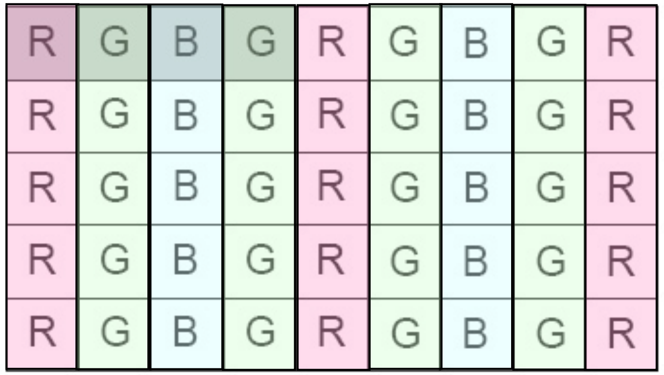
- Standard RGB-filters
  - Each color channel is rather narrow
    $\Rightarrow$ blocks more photons $\Rightarrow$ less effective
- Cyan-Yellow-Magenta (white) filters
  - Magenta = red + blue, Cyan = blue + green
  - Yellow = red + green, White = red+green+blue
  - Each color channel is wider than standard 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

Stripe filters

Examples:







Extra green

Dark area is a cell that represents one ”pixel”

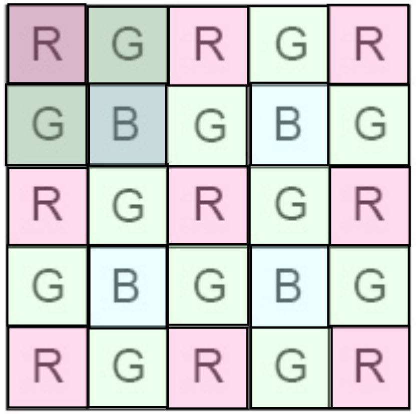
Easy to implement

May produce moiré effects due to color aliasing

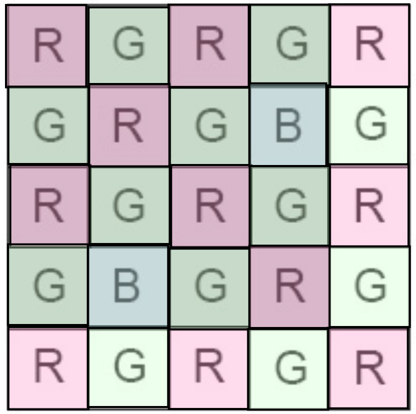
Bayer filters

Examples:

Reduces aliasing effects



R+2G+B
per cell



6R+8G+2B
per cell

12W+12G+8C
per cell

Color post-processing

- We can see the image detected by the sensor as a mono-chrome 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 (*de-mosaicking*)
- 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

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

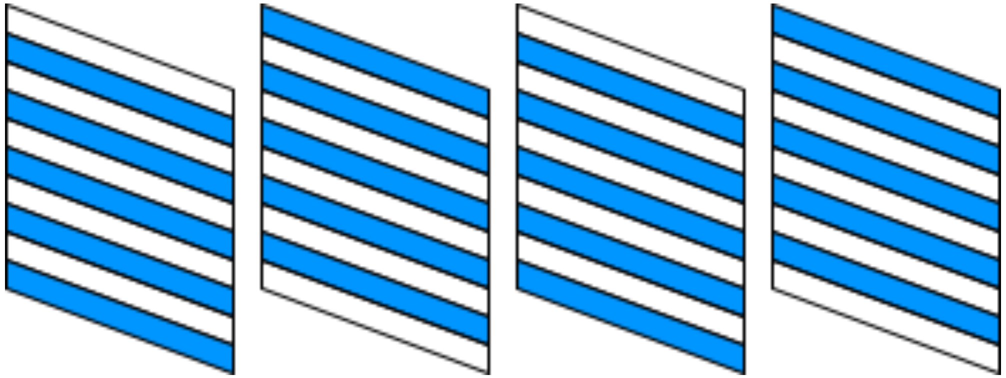
- The perception of color is complex
  - Humans tend to perceive color independent of illumination
  - A color camera makes a measurement of physical quantities: very dependent on illumination
- White balancing:
  - Transforms the color measurements to make what we perceive as a white surface produce 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 is becoming more and more common
  - Used in many modern video standards

Interlaced vs. progressive scan

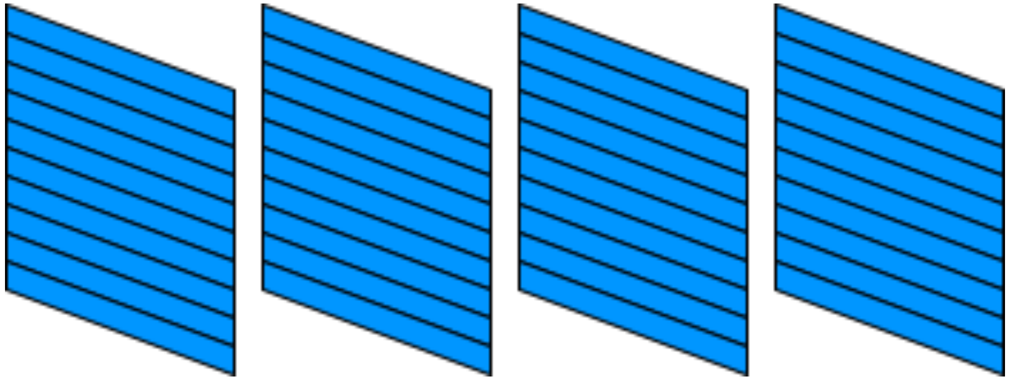


Interlaced scan

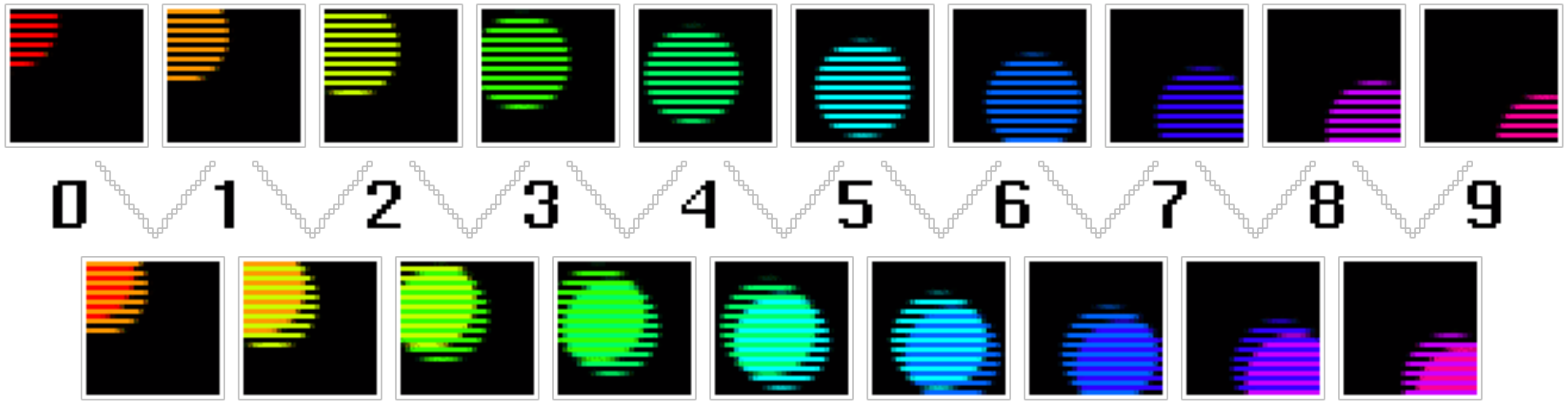
E.g., one “half image”
at 50 Hz $\Rightarrow$
one “full image” at 25 Hz

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 where 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 (lens limitation)
  - 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

A final word

- The technology related to image sensors is in rapid development
    - The components are constantly becoming smaller (Moore's law)
    - New solutions to various problems appear at high pace
    - More and more functionality is being integrated with the image sensor
    - Image sensors are being integrated with other functionalities
- Keep an eye on what is happening!