Photons and sensors

(with an interlude on the history of color photography)

CS 178, Spring 2009

Begun on Thursday, 4/16/09, finished on Tuesday, 4/21. Slide added (#24) on 4/21 to further explain 2D convolution.

Marc Levoy
Computer Science Department
Stanford University

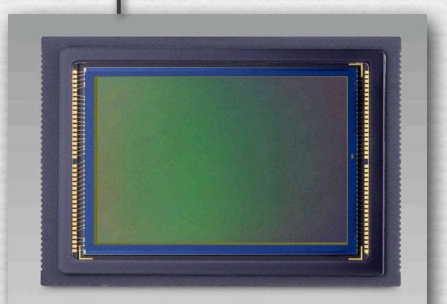
Camera pixel pipeline

sensor → analog to digital conversion (ADC) → processing: demosaicing, tone mapping & white balancing, denoising & sharpening, compression → storage

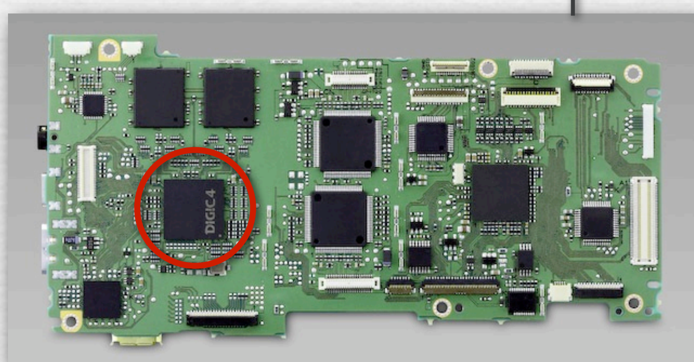
- every camera uses different algorithms
- the processing order may vary
- most of it is proprietary

Example pipeline

sensor → analog to digital conversion (ADC) → processing: demosaicing, tone mapping & white balancing, denoising & sharpening, compression → storage



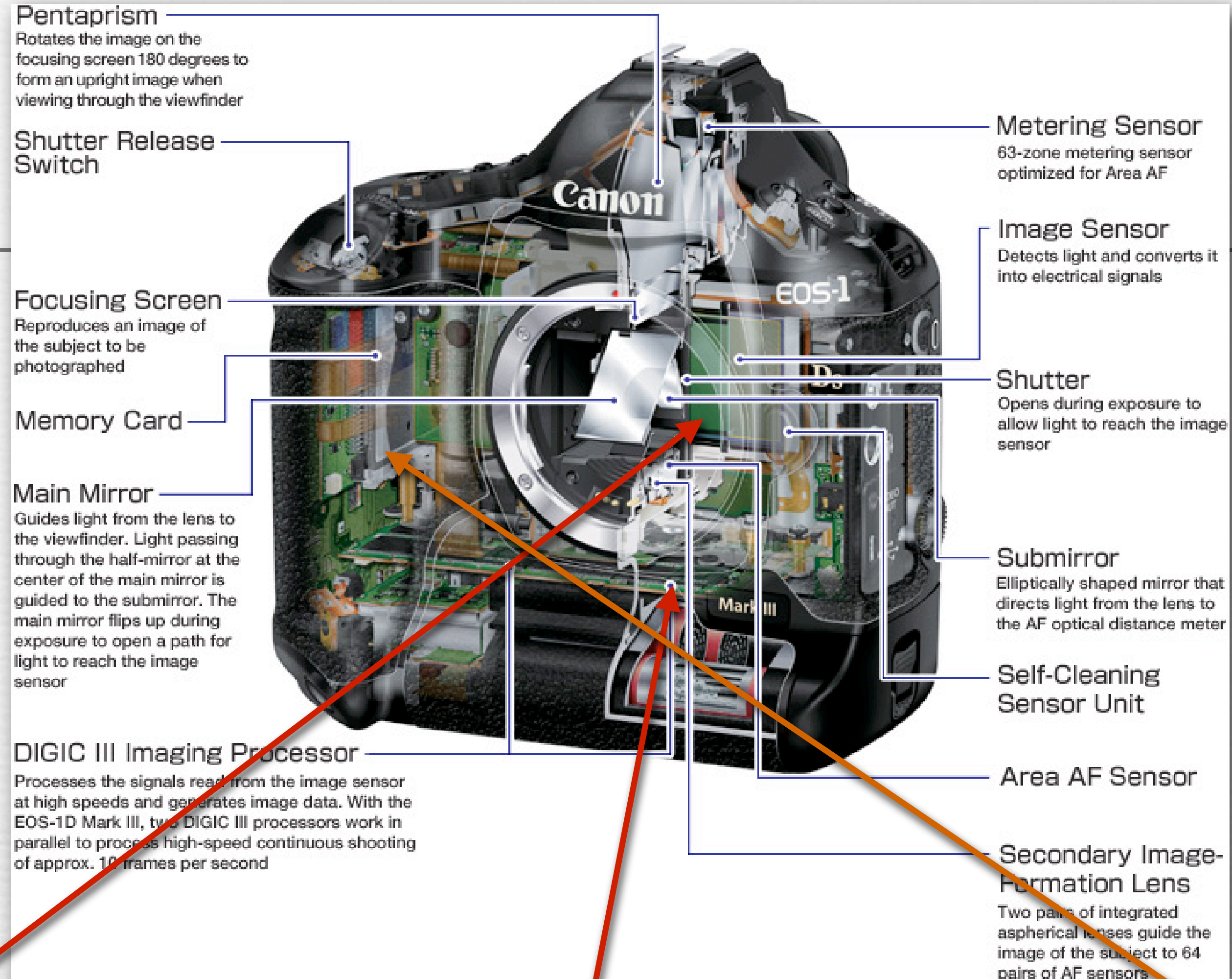
Canon 21 Mpix CMOS sensor

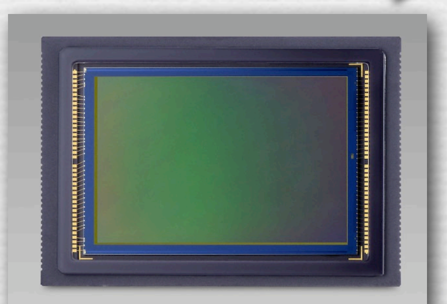


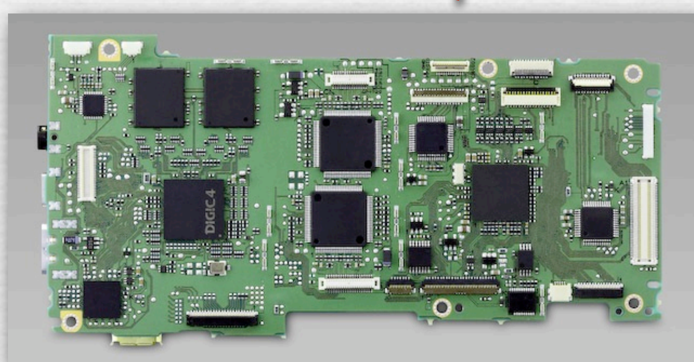
Canon DIGIC 4 processor

Compact Flash card

Example




Canon 21 Mpix CMOS sensor

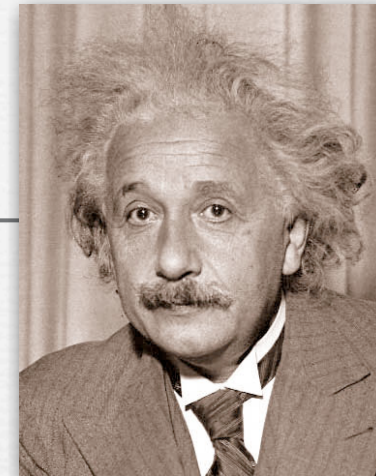

Canon DIGIC 4 processor

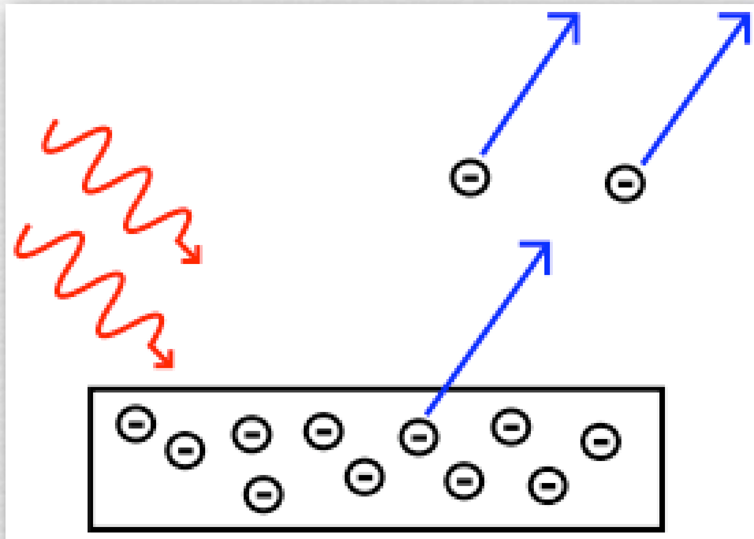
Compact Flash card

Outline

- converting photons to charge
- getting the charge off the sensor
  - CCD versus CMOS
  - analog to digital conversion (ADC)
- supporting technology
  - microlenses
  - antialiasing filters
  - shutters
- sensing color

The photoelectric effect

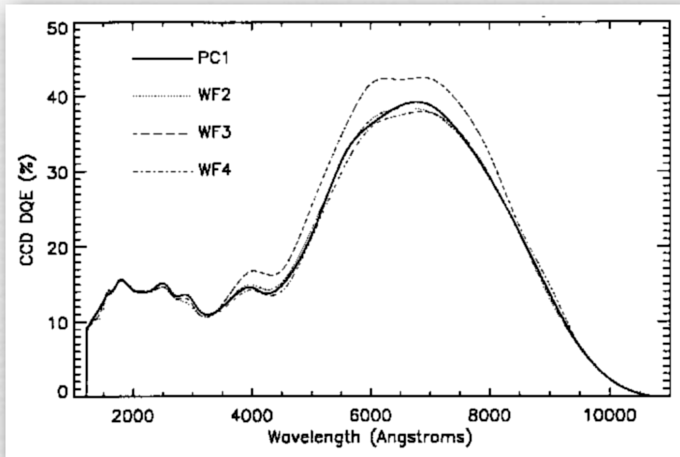

Albert Einstein


(wikipedia)

- when a photon strikes a material, an electron may be emitted
  - depends on wavelength, not intensity

$$E_{photon} = \frac{h \times c}{\lambda}$$

Quantum efficiency

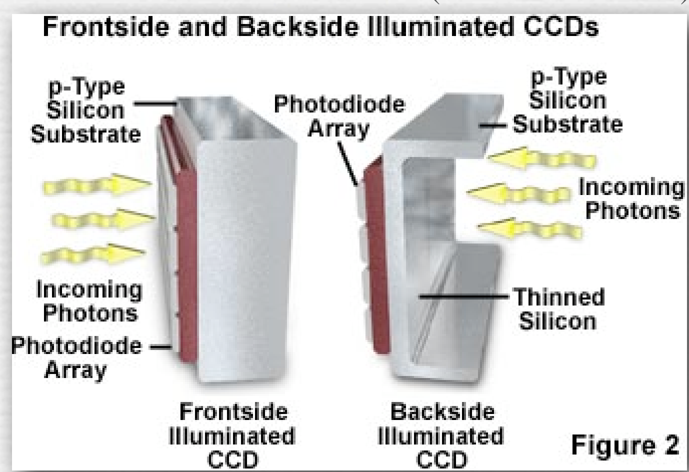


Hubble Space Telescope Camera 2

- not all photons will produce an electron
  - depends on quantum efficiency of the device

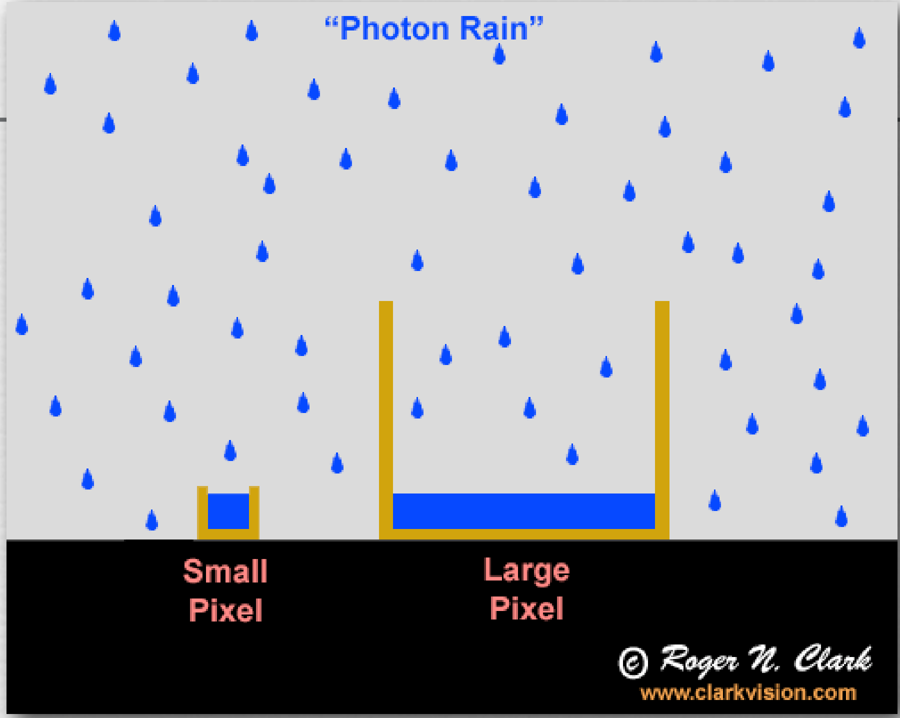
$$QE = \frac{\#\,electrons}{\#\,photons}$$

(hamamatsu.com)



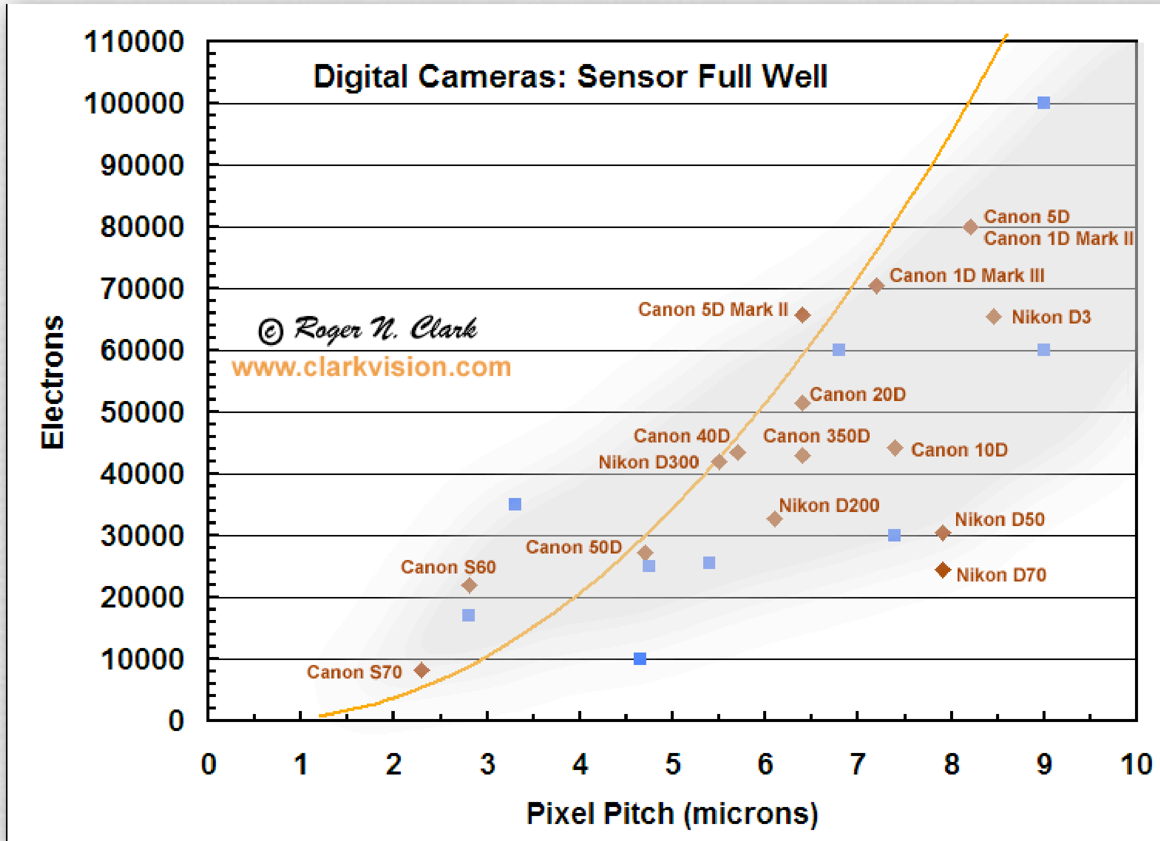
  - human vision: ~15%
  - typical digital camera: < 50%
  - best back-thinned CCD: > 90%

Sensor pixel



- ✦ resulting current is small (10-100 fA)
  - so integrate over space and time (pixel area × exposure time)
  - larger pixel × longer exposure means more accurate measure
- ✦ typical pixel sizes
  - casio EX-F1: $2.5\mu \times 2.5\mu = 6\mu^2$
  - Canon 5D II: $6.4\mu \times 6.4\mu = 41\mu^2$

Full well capacity



(clarkvision.com)

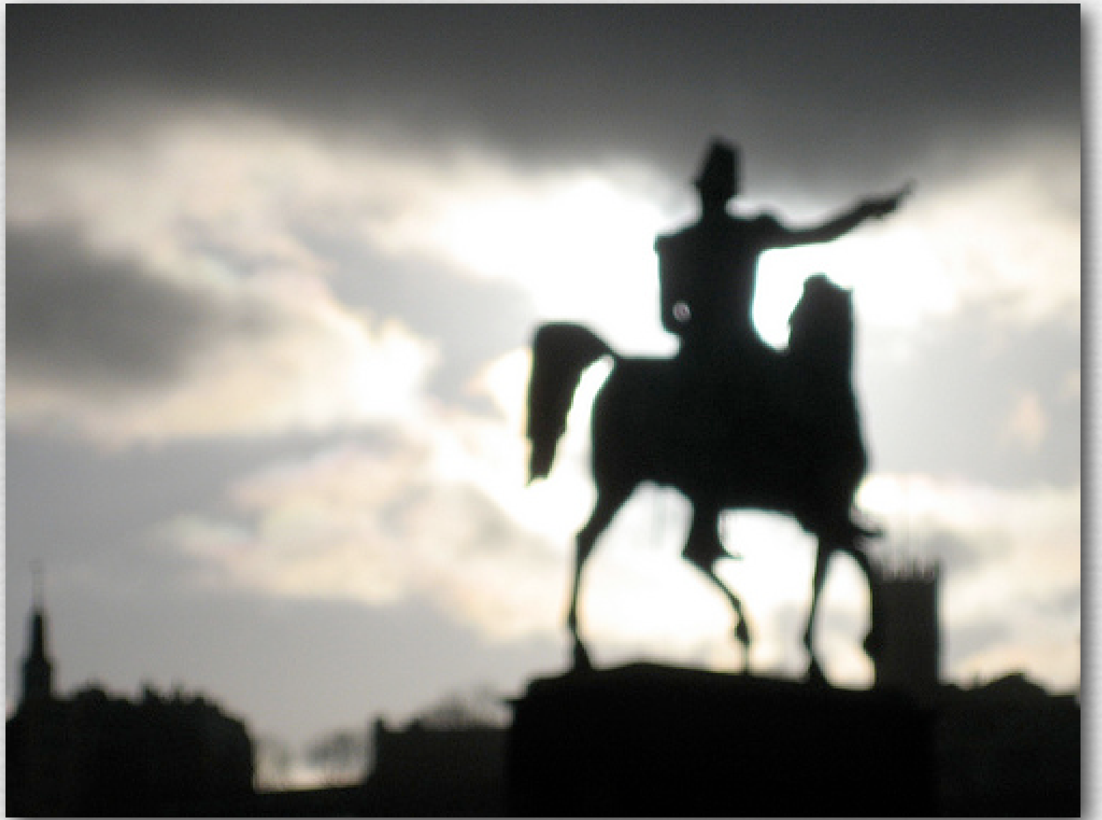
- too many photons causes *saturation*
  - larger capacity leads to higher *dynamic range*
  - but the *noise floor* is also a factor, as we'll see

Blooming

(ccd-sensor.de)

- charge spilling over to nearby pixels
  - can happen on CCD <u>and</u> CMOS sensors
  - don't confuse with glare or other image artifacts

Image artifacts can be hard to diagnose

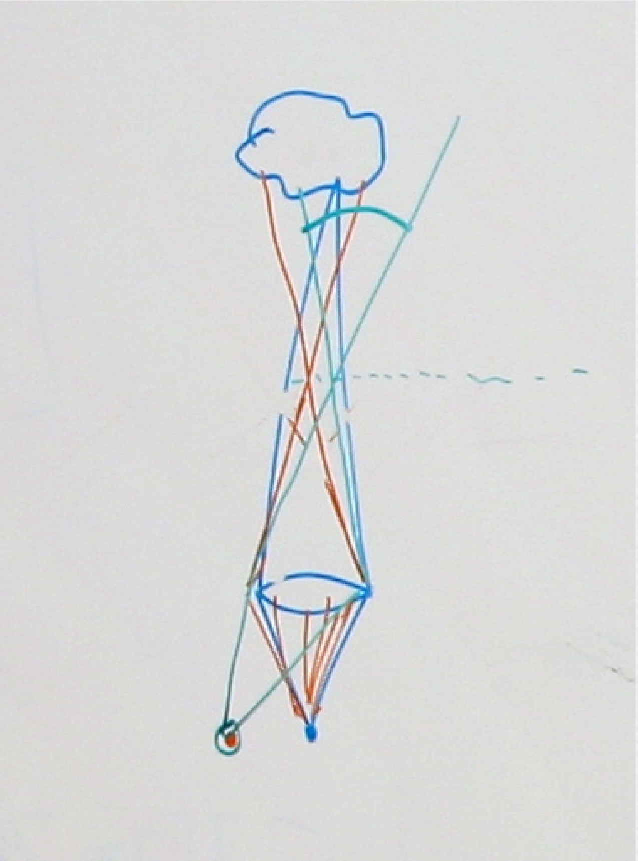


(http://farm3.static.flickr.com/2102/2248725961_540be5f9af.jpg?v=0)

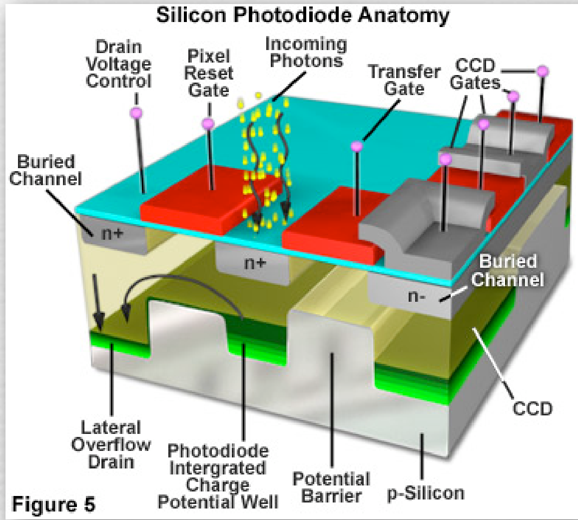
Q. Is this blooming?

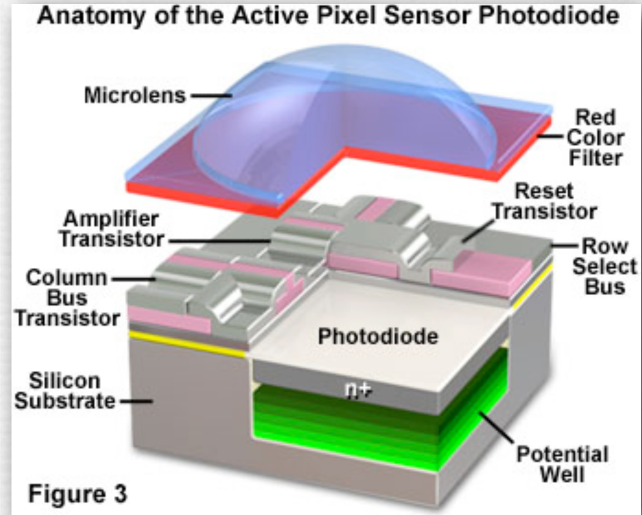
Explanation of preceding image

- there may be blooming in the sky, but the shrinkage of the horse's leg can be explained purely as a byproduct of misfocus
  - in the accompanying plan view diagram, the horse's leg is shown at top (in cross section)
  - the red bundle of rays, corresponding to one sensor pixel, crossed before the leg (was misfocused), then spread out again, but saw only more leg, so its color would be dark
  - the green bundle of rays, corresponding to a nearby pixel, crossed at the same depth but to the side of the red bundle, then spread out again, seeing partly leg and partly sky; its color would be lighter than the leg
  - this lightening would look like the sky was "blooming" across the leg, but it's just a natural effect produced by misfocus



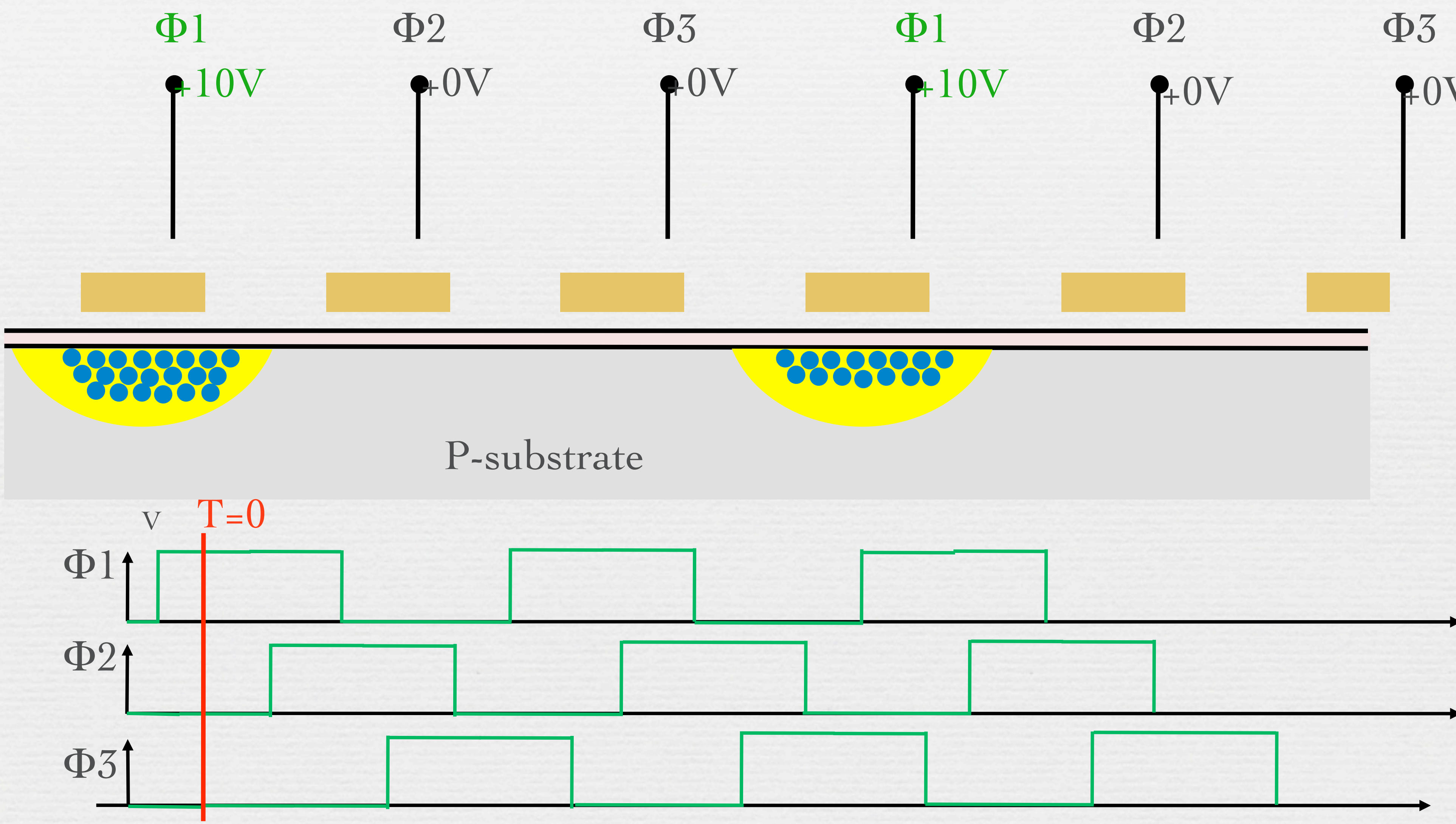
CCD versus CMOS sensors



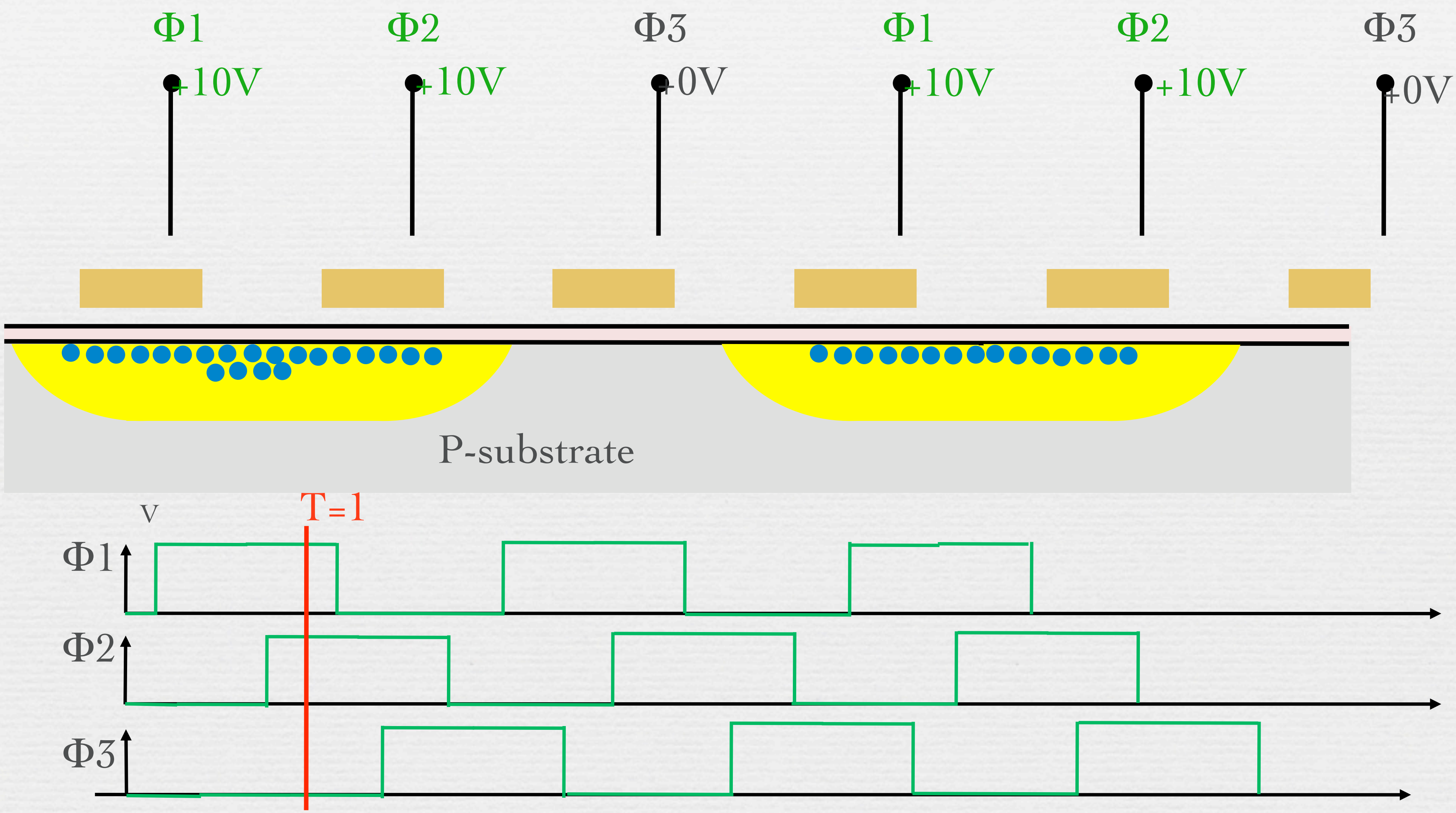


- CCD = charge-coupled device
  - charge shifted along columns to an output amplifier
  - oldest solid-state image sensor technology
  - highest quality, but not as flexible or cheap as CMOS
- CMOS = complementary metal-oxide semiconductor
  - an amplifier per pixel converts charge to voltage
  - low power, but noisier (but getting better)

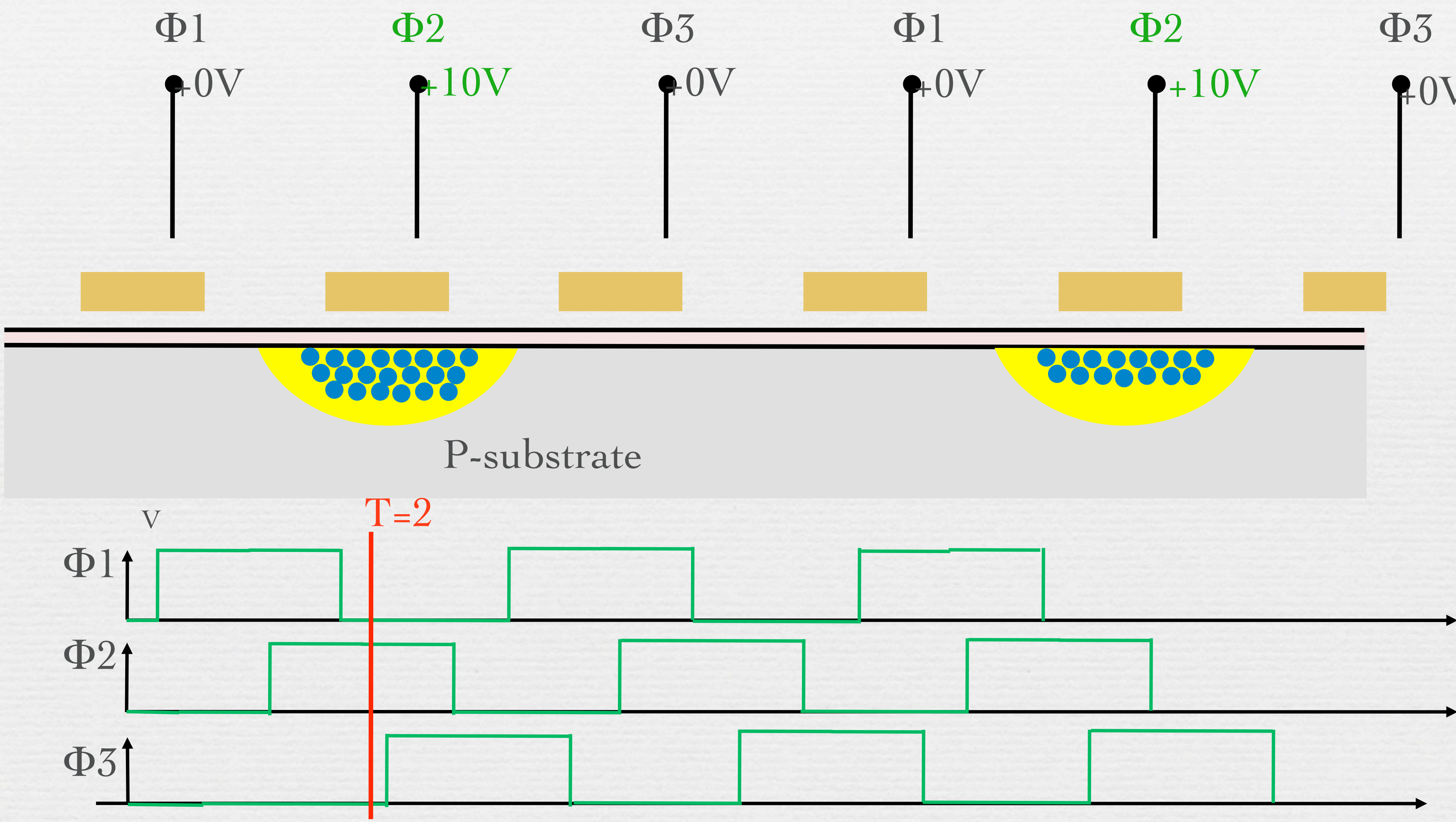
Gratuitous animation showing a CCD “bucket brigade” readout



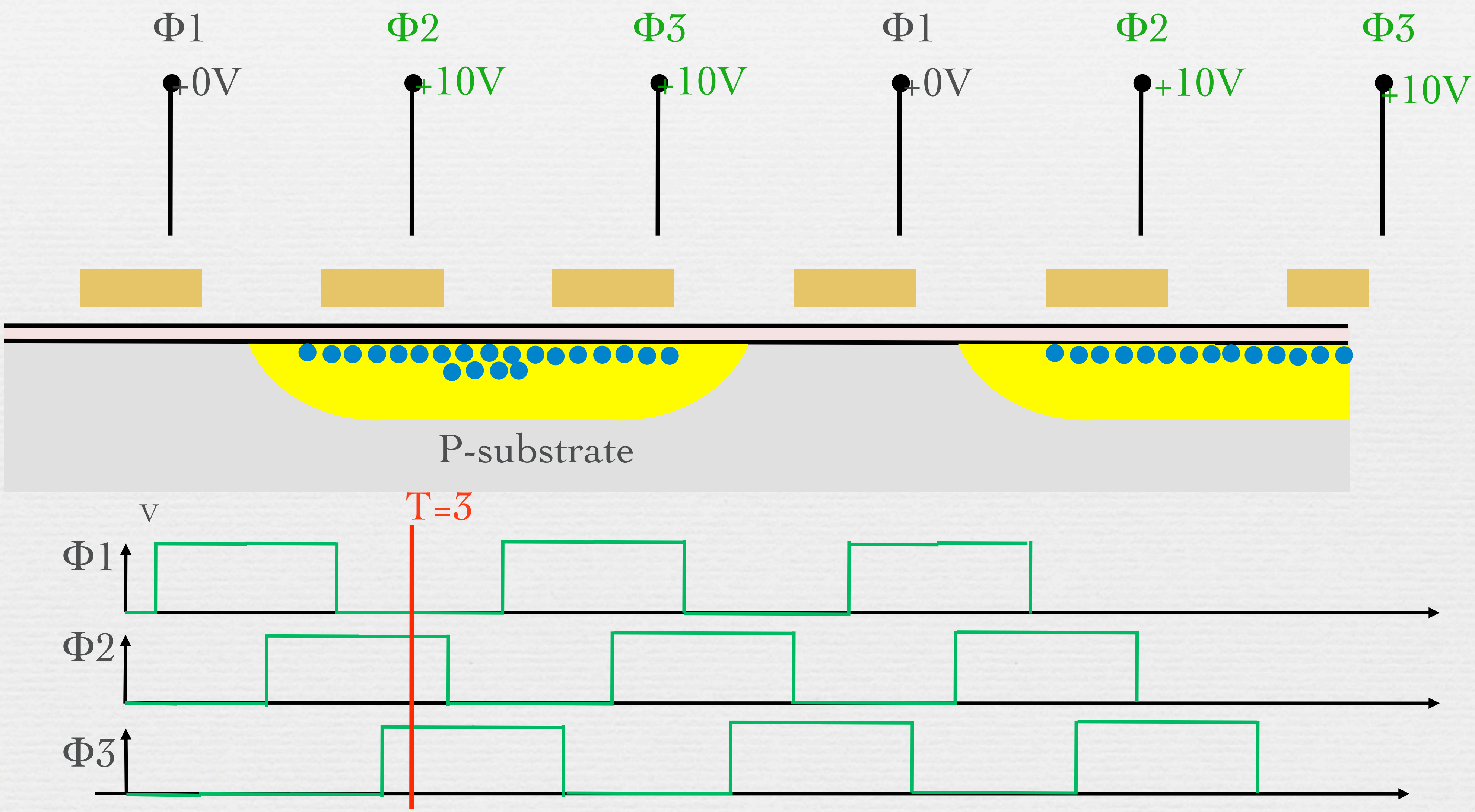
Gratuitous animation showing a CCD "bucket brigade" readout



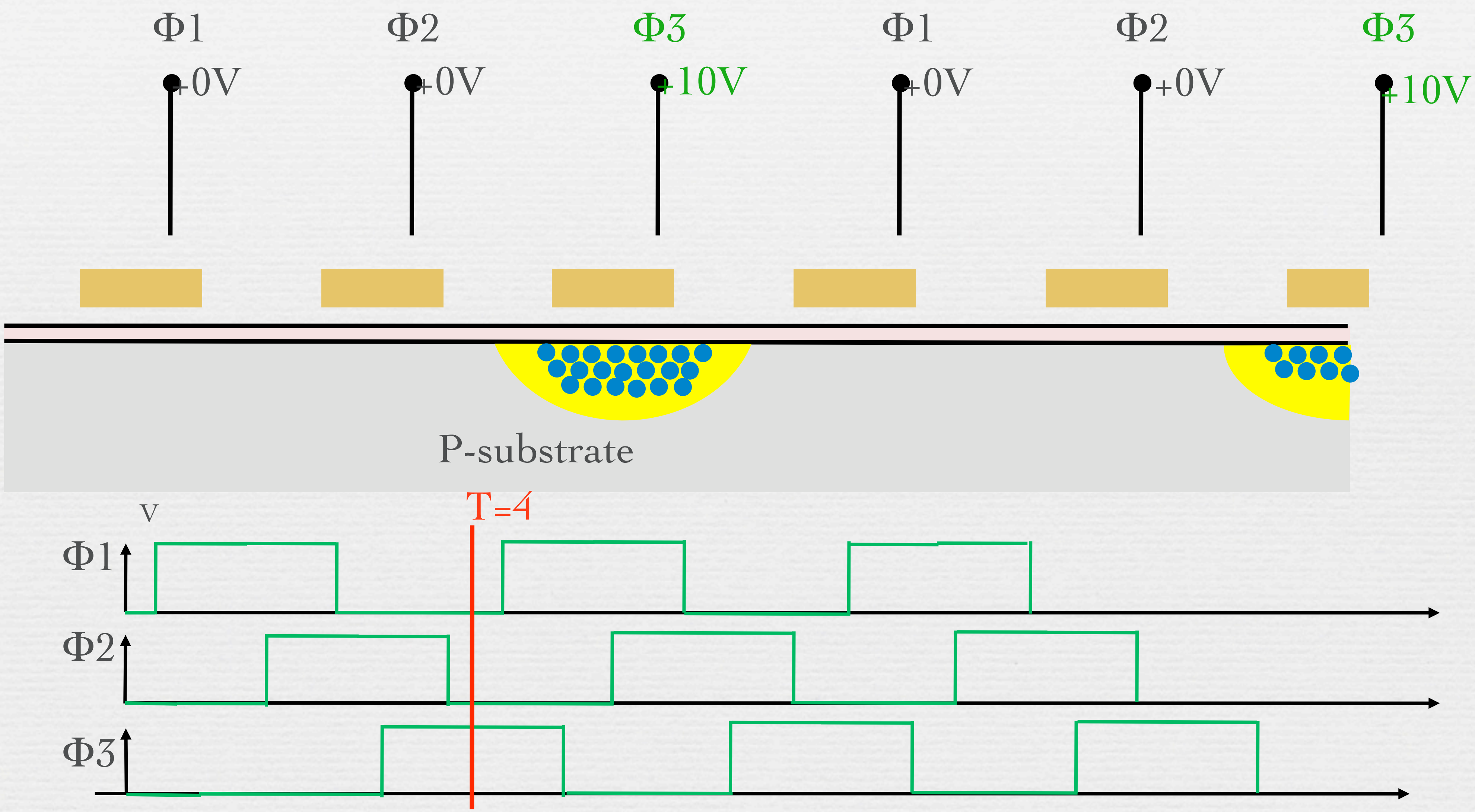
Gratuitous animation showing a CCD "bucket brigade" readout



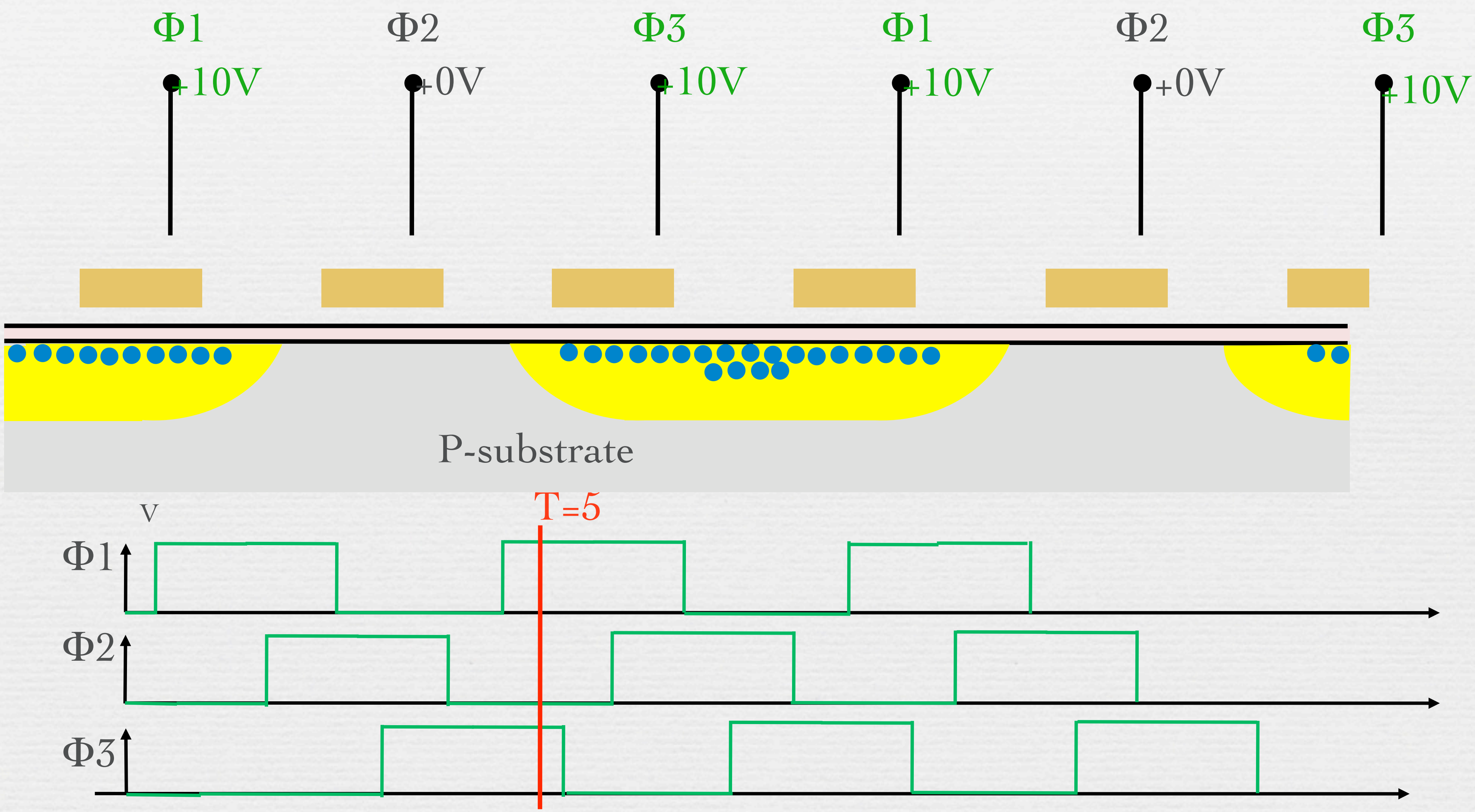
Gratuitous animation showing a CCD “bucket brigade” readout



Gratuitous animation showing a CCD "bucket brigade" readout



Gratuitous animation showing a CCD “bucket brigade” readout



Smearing

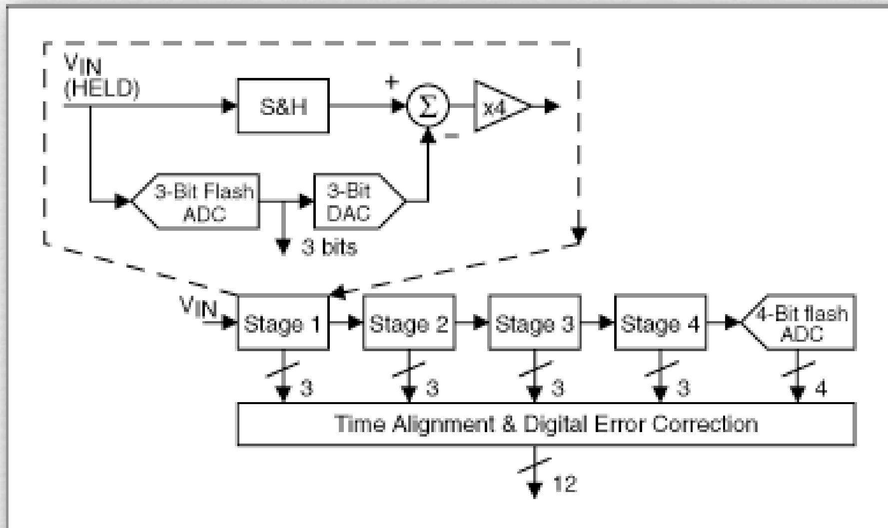
(dvxuser.com)

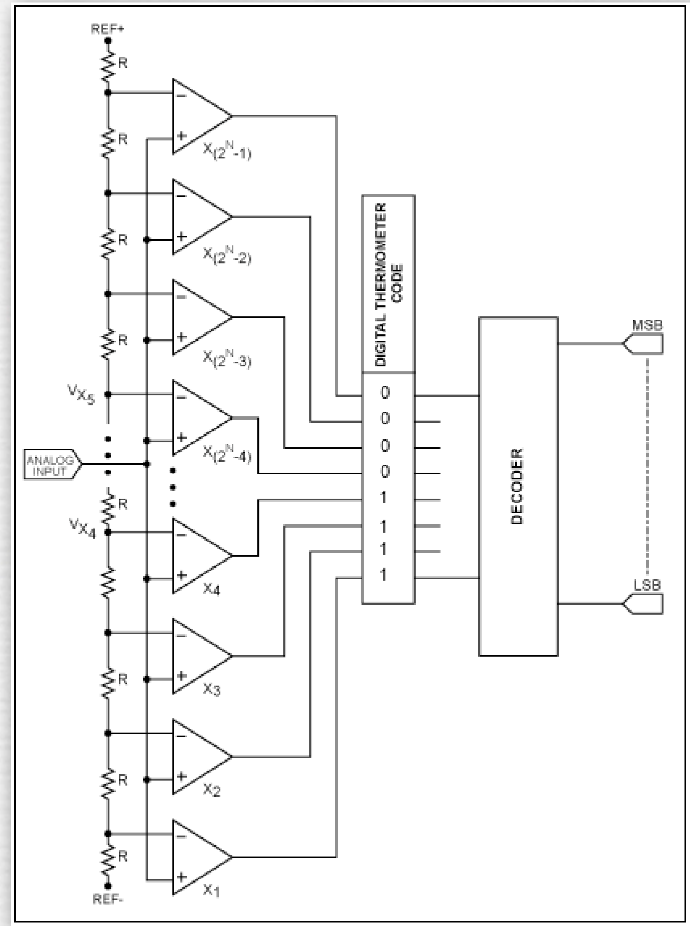
CCD

CMOS

- side effect of bucket-brigade readout on CCD sensors
  - only happens if pixels saturate
  - doesn't happen on CMOS sensors

Analog to digital conversion (ADC)

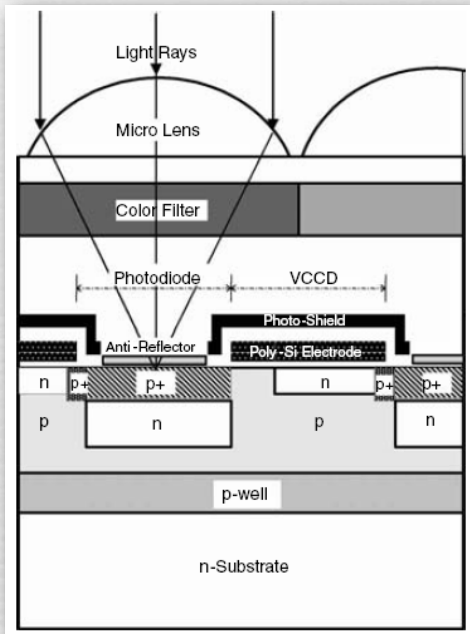




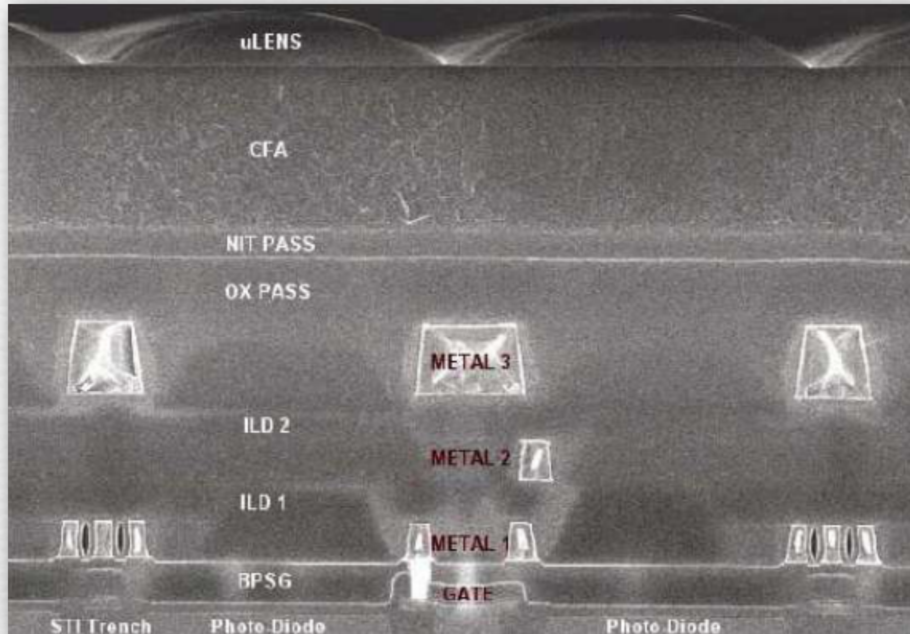
(maxim-ic.com)

- flash ADC
  - voltage divider → comparators → decoder
  - for n bits requires 2^n comparators
- pipelined ADC
  - 3-bit ADC → 3-bit DAC → compute residual → 4× → repeat
  - longer latency, but high throughput
  - some new sensors use an ADC per column

Fill factor



on a CCD sensor



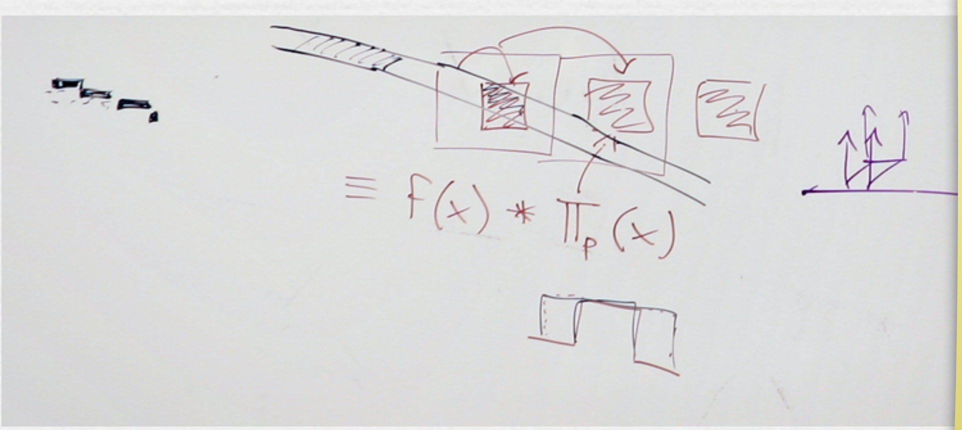
on a CMOS sensor

- ✦ fraction of sensor surface available to collect photons
  - can be improved using per-pixel microlenses

Q. An image sensor performs 2D sampling.
What is the prefilter, with and without microlenses?

What per-pixel microlenses do

- integrating light over a pixel serves two functions: capturing more photons, and convolving the focused image with a 2D Rect prefilter, to avoid aliasing
  - if only a portion of each pixel site is photo-sensitive, this Rect doesn't span the spacing between pixels, so the prefilter is poor, and aliasing can result (like the short exposures in old movie cameras)
- microlenses both gather more light and improve the prefilter
  - with microlenses, prefilter width roughly equals pixel spacing

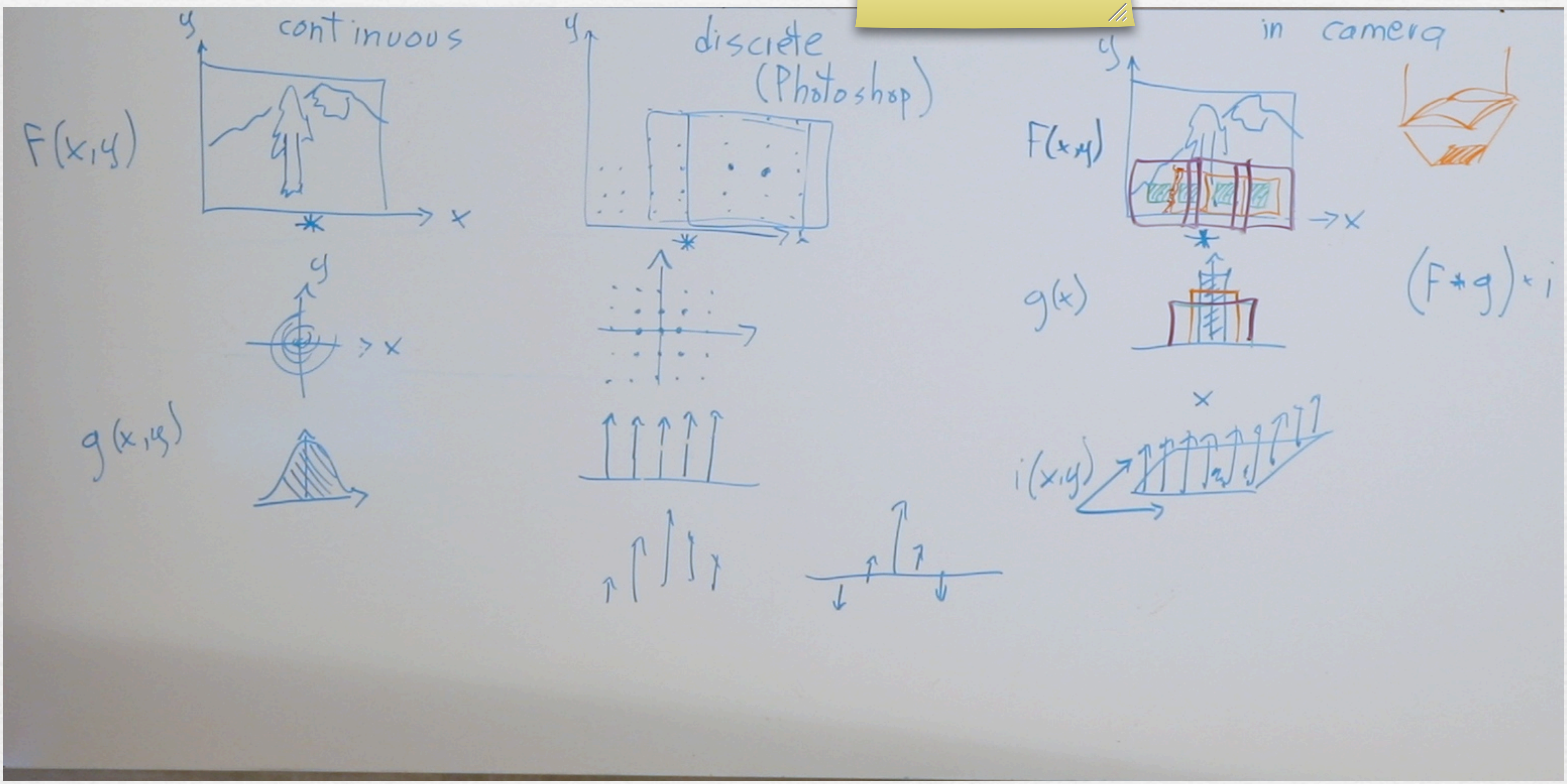


In lecture I described the signal f(x) and filters g(x) and r(x) as being one-dimensional, and I defined convolution is a single integral over the single variable "tau". I should have made clear that the theory extends naturally to two dimensions. In particular, the focused image arriving on a sensor can be treated as a 2D function f(x,y). Prefiltering consists of convolving this with a 2D filter function g(x,y). This 2D convolution is the double integral of f(x-tau, y-phi) x g(tau,phi) dtau dphi. For the special case of a square pixel that is equally sensitive at all points on its surface, g(x,y) is a 2D Rect function, i.e. a box from -1/2 to 1/2 in x and y and of unit height. See next slide, added 4/21/09. If you want to include exposure time, then the arriving time-varying image becomes the 3D function f(x,y,t), the prefilter g(x,y,t) is a pixel area considered over time, and convolution becomes a triple integral. Assuming the shutter opens and closes instantly, g(x,y,t) is a 3D Rect, i.e. a rectangular parallelepiped, still of unit volume. The same extensions apply to postfiltering, i.e. to reconstruction; it consists of convolving the discretely sampled function fs(x,y), i.e. the array of pixels, with a 2D filter function r(x,y). For images displayed on a laptop, r(x,y) is another 2D Rect function, equal is extent to a pixel on the laptop's screen. This convolution is done for you naturally when you display the image.

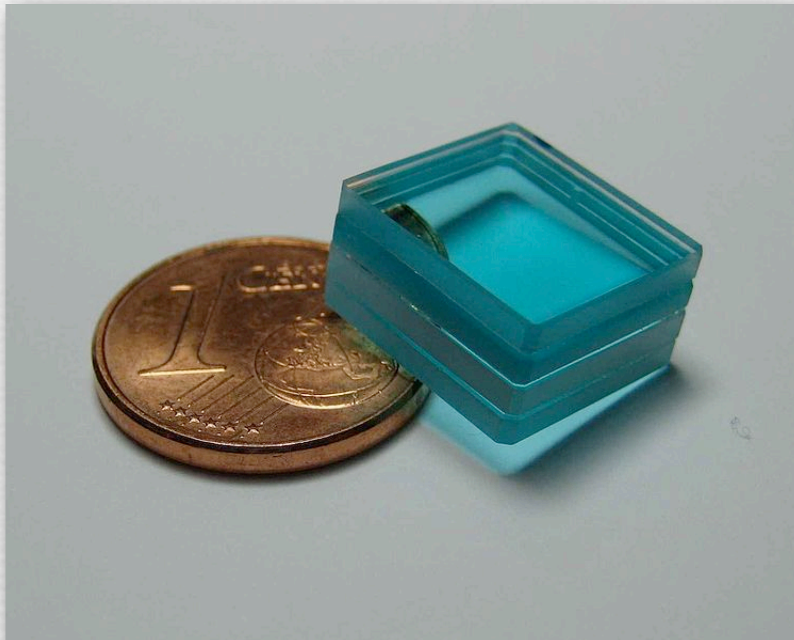
Spatial convolution in 2D

- re-iterating the points made in the yellow comment post-it on the previous slide

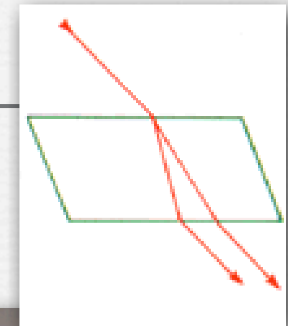
Slide added 4/21/09



Antialiasing filters



antialiasing filter

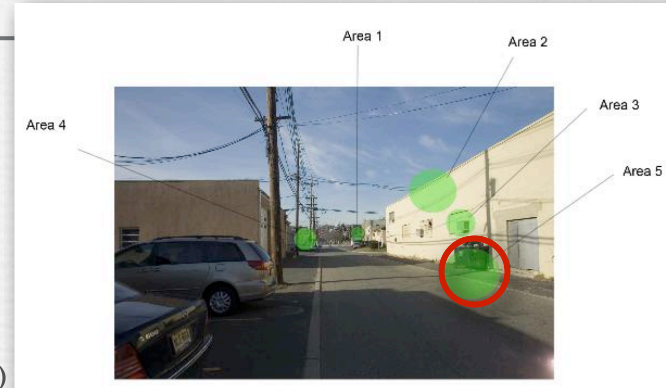


birefringence in a calcite crystal

- improves on non-ideal prefilter, even with microlenses
- typically two layers of birefringent material
  - splits 1 ray into 4 rays
  - operates like a 4-tap discrete convolution filter kernel!

Removing the antialiasing filter

- "hot rodding" your digital camera
  - $450 + shipping



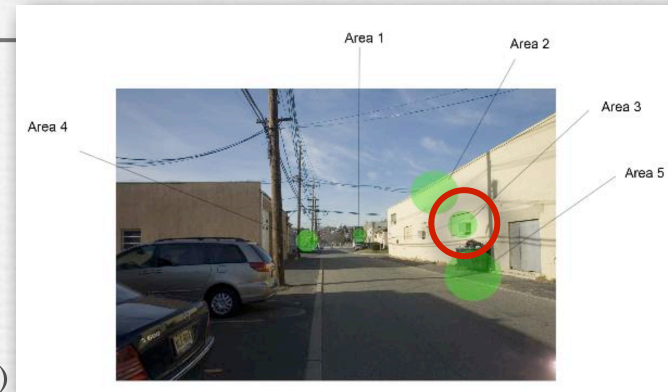
(maxmax.com)

anti-aliasing filter removed

normal

Removing the antialiasing filter

- “hot rodding” your digital camera
    - $450 + shipping

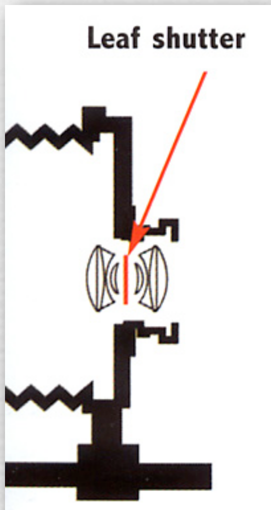


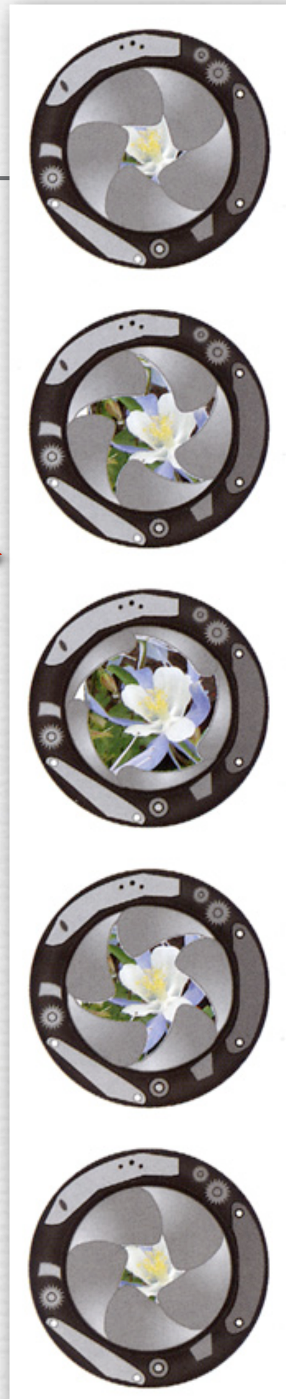
(maxmax.com)

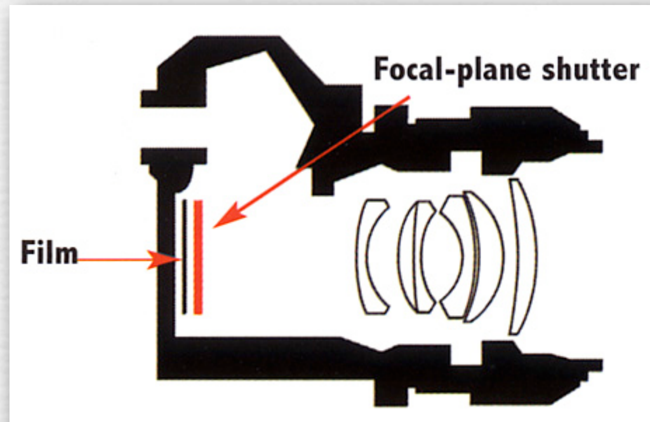
anti-aliasing filter removed

normal

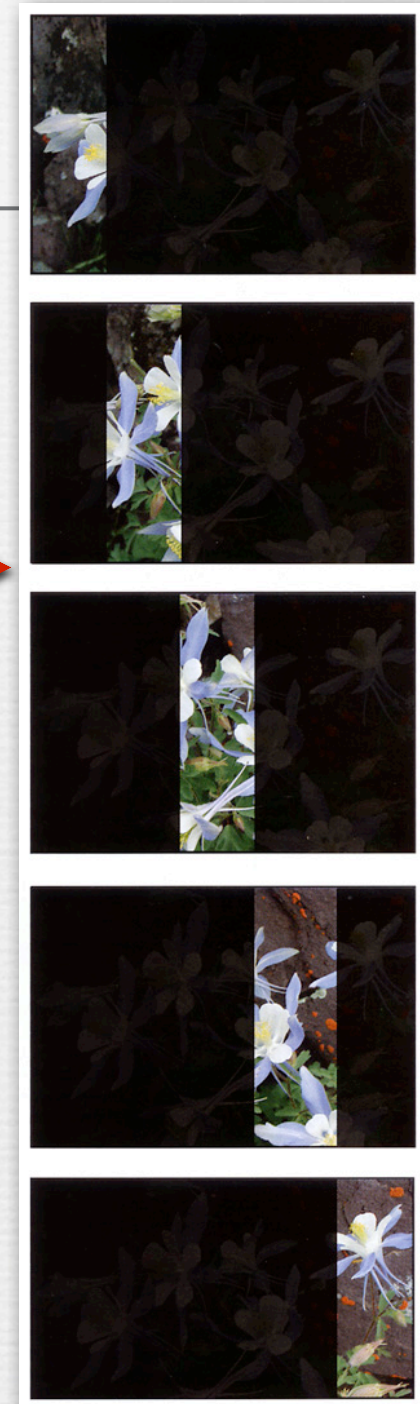
Shutters





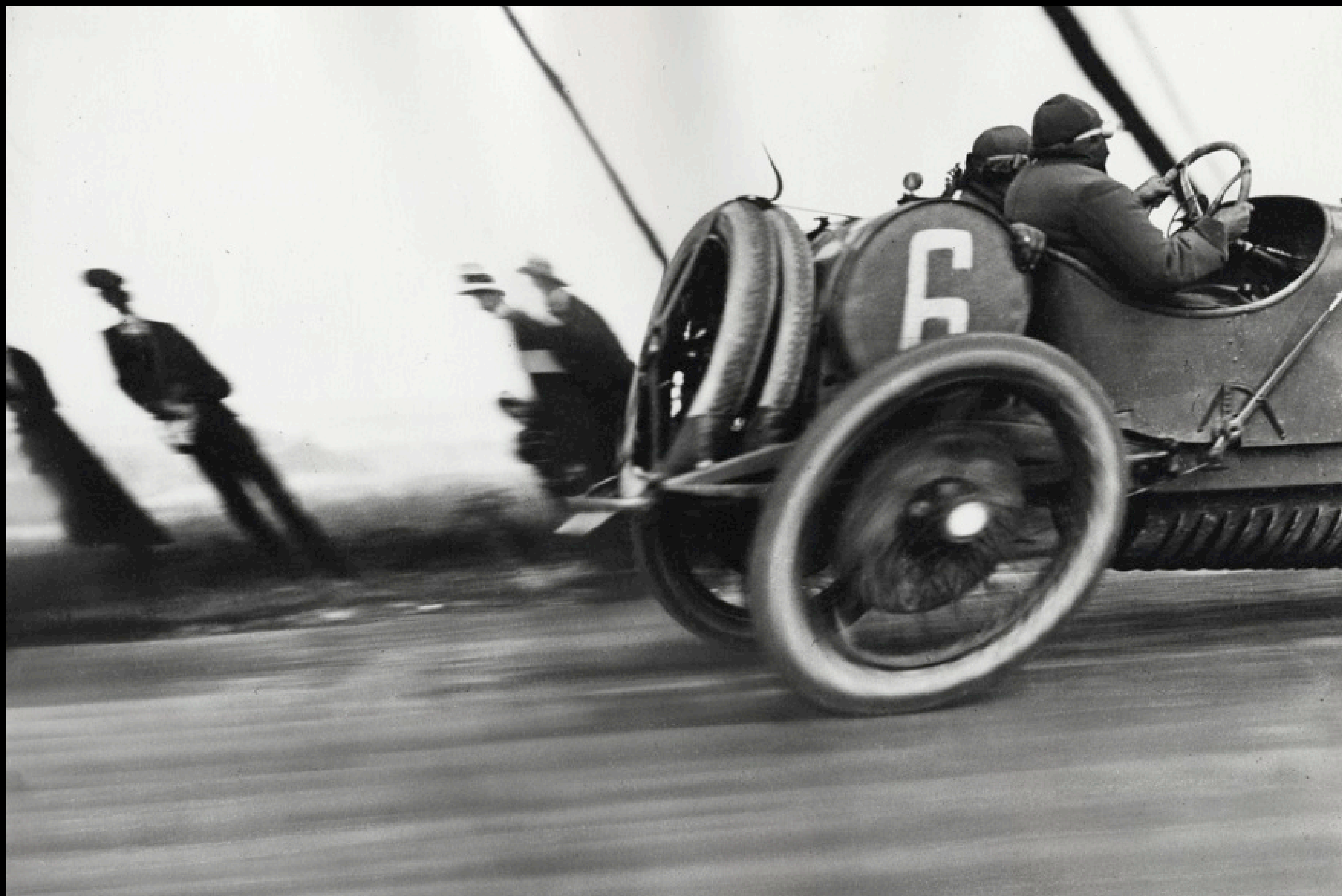


(London)



- quiet
- slow
  (max 1/500s)
- need one
  per lens

- loud
- fast
  (max 1/4000)
- distorts motion

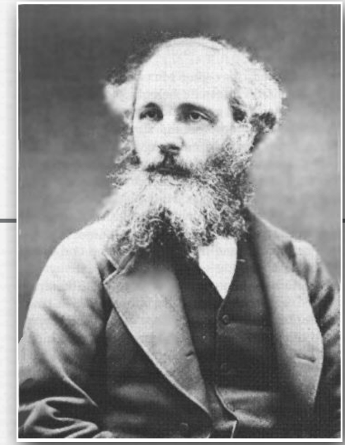


Jacques-Henri Lartigue, Grand Prix (1912)

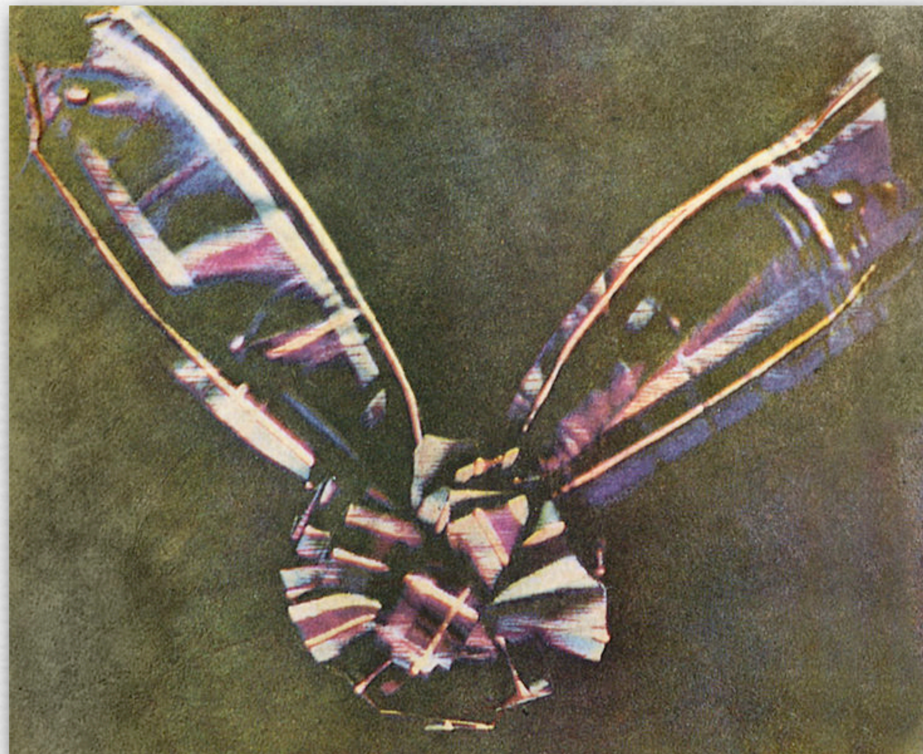
Color sensing technologies

- field-sequential
- 3-sensor
- spatial mosaic
- vertically stacked

Historical interlude



Q. Who made the first color photograph?



(wikipedia)

- James Clerk Maxwell, 1861
  - of Maxwell's equations
  - 3 images, shot through filters, then simultaneously projected

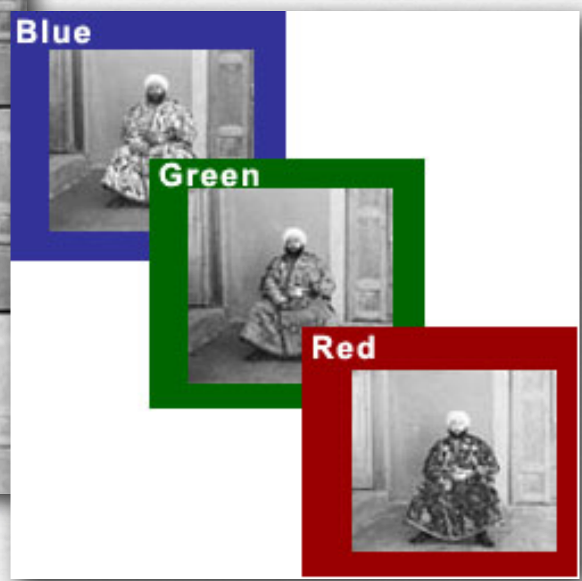
Historical interlude

Q. Who made the first color print?

(wikipedia)

- Louis Arthur Ducos du Hauron, 1877
  - 3 images, shot through filters, printed with color inks
  - he experimented with RGB and CMY

Sergey Prokudin-Gorsky



- simultaneous projection provided good saturation, but available printing technology did not
- digital restoration lets us see them in full glory...

Sergey Prokudin-Gorsky, Alim Khan, emir of Bukhara (1911)

Sergey Prokudin-Gorsky,
Pinkhus Karlinskii, Supervisor of the Chernigov Floodgate (1919)

First color movie technology?

(wikipedia)

A Visit to the Seaside (1908)

- George Albert Smith's Kinemacolor, 1906
    - alternating red and green filters, total of 32 fps
    - projected through alternating red and green filters

Technicolor

Toll of the Sea (1922)

Phantom of the Opera (1925)

- beam splitter leading through 2 filters to two cameras
- 2 strips of film, cemented together for projection

Technicolor

Disney's Flowers and Trees (1932)

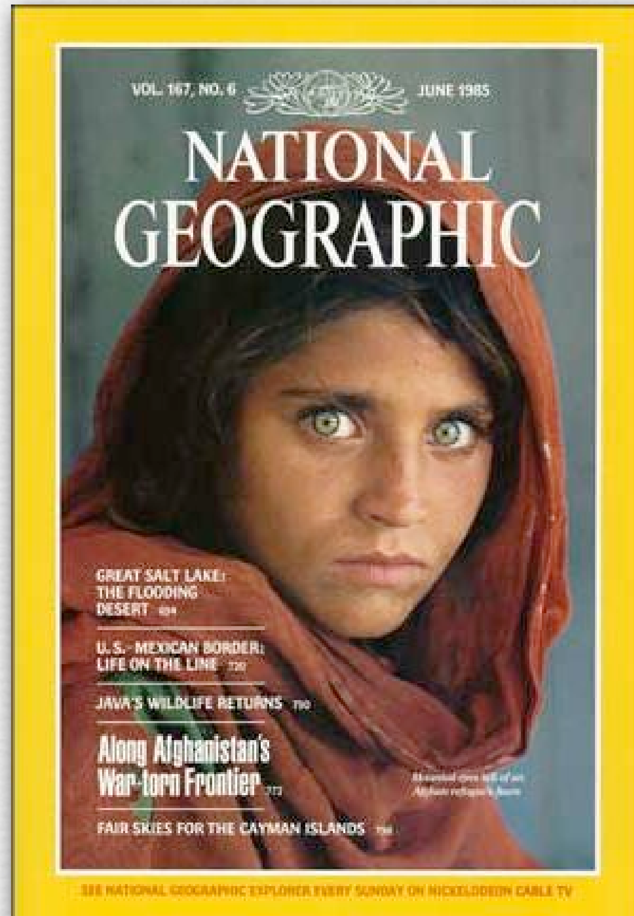

Wizard of Oz (1939)

- 3 filters, 3 cameras, 3 strips of film
- better preserved than single-strip color movies of 1960s!

First consumer color film?

(wikipedia)

Picadilly Circus, 1949



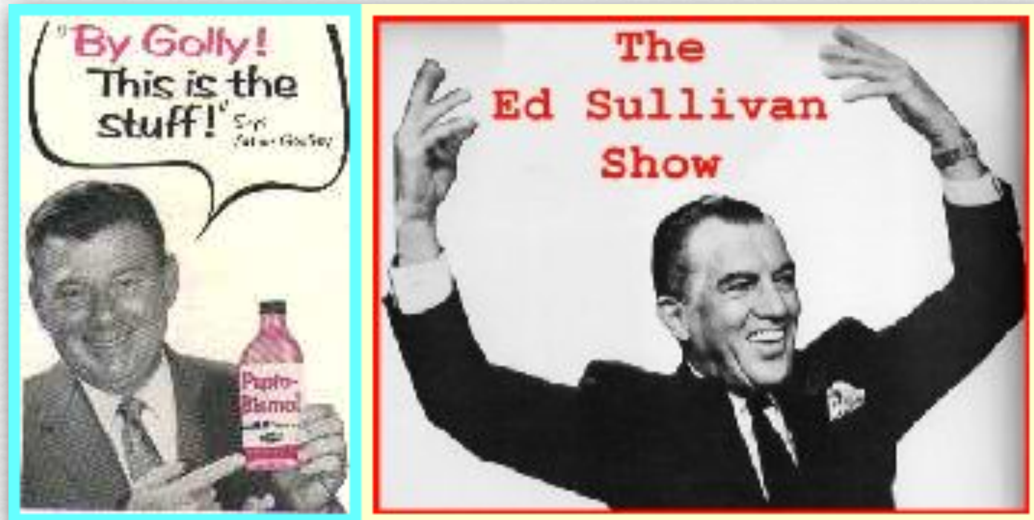
- Kodachrome, 1935
  - no longer available

First color television broadcast?

1951

- competing standards
  - U.S. NTSC 525-line, 30fps, interlaced
  - Europe PAL 625-line, 50fps, interlaced
  - France SECAM 625-line, 50fps, interlaced

First color television broadcast?



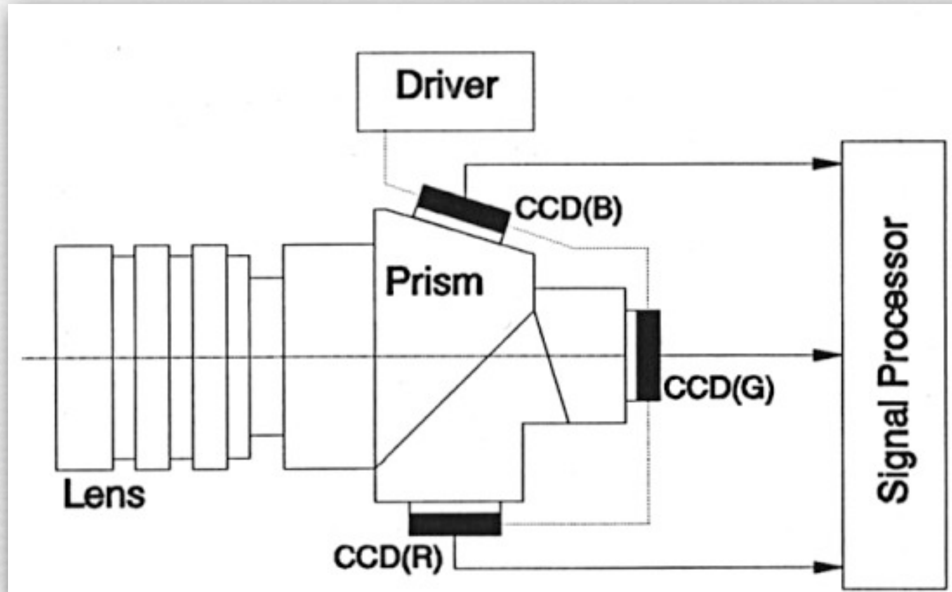
1951

- competing standards
  - U.S. NTSC Never Twice the Same Color
  - Europe PAL Pale and Lurid
  - France SECAM Système Electronique Contre les Americains

Color sensing technologies

- field-sequential - just covered
- 3-sensor
- spatial mosaic
- vertically stacked

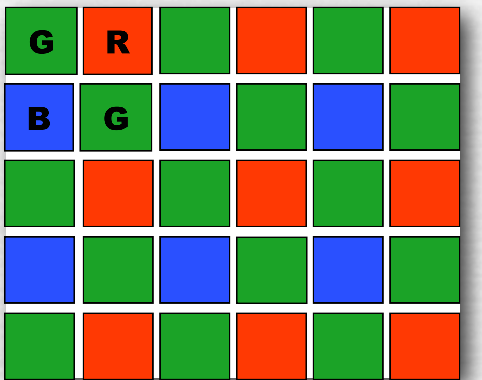
3-CCD cameras



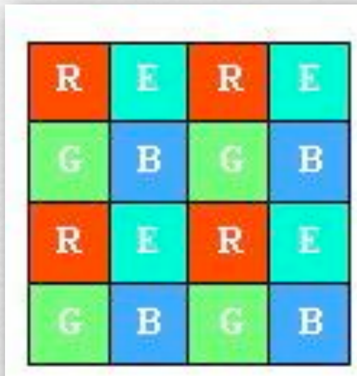
(Theuwissen)

- high-quality video cameras
- prism & dichroic mirrors split the image into 3 colors, each routed to a separate CCD sensor
- no light loss, as compared to filters
- expensive, and complicates lens design

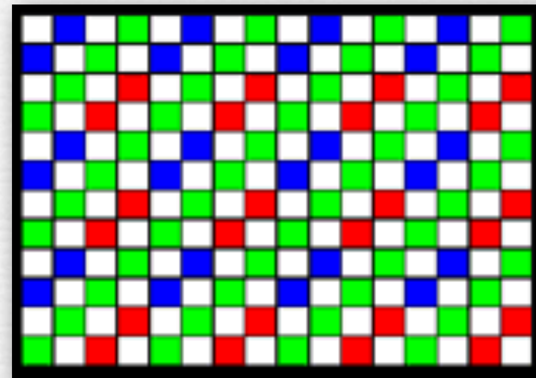
Color mosaic



Bayer pattern



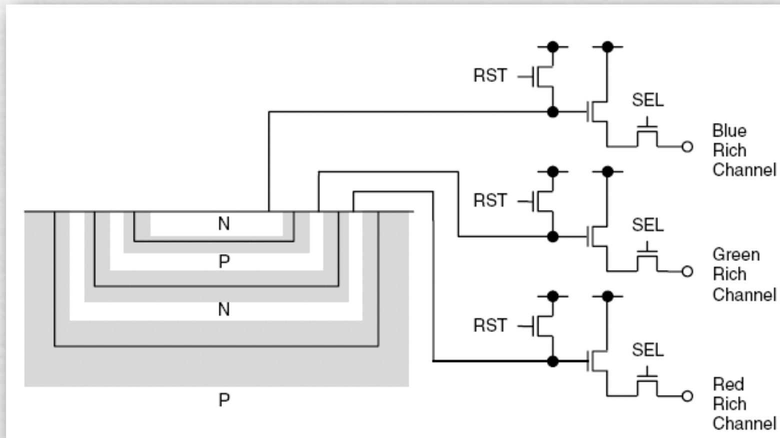
Sony RGB+E
better color


Kodak RGB+C
less noise

- ✦ uses absorption filters, so light is lost
- ✦ interpolate color information to get RGB per pixel
  - called *demosaicing*
  - hard problem, many artifacts, active research area
  - we'll return to this...

Crop from raw Bayer mosaic image

Foveon stacked sensor



- longer wavelengths penetrate deeper into silicon, so arrange a set of vertically stacked detectors
  - top gets mostly blue, middle gets green, bottom gets red
  - no control over spectral responses, so requires processing
- fewer color artifacts than color mosaics
  - but possibly worse noise performance

Not yet covered

- dynamic range

$$DR = \frac{\text{max output swing}}{\text{noise in the dark}} = \frac{\text{saturation level - dark current}}{\text{dark shot noise + readout noise}}$$

- noise and ISO
- spectral characteristics of color filters
- demosaicing

Slide credits

- Brian Curless
- Eddy Talvala
- Abbas El Gamal

- Theuwissen A., *Solid-State Imaging with Charge-Coupled Devices,* Kluwer Academic Publishers, 1995.