Photons and sensors

(with an interlude on the history of color photography)

CS 178, Spring 2013

Marc Levoy
Computer Science Department
Stanford University

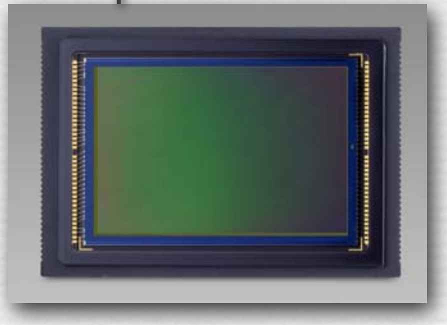
Camera pixel pipeline

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

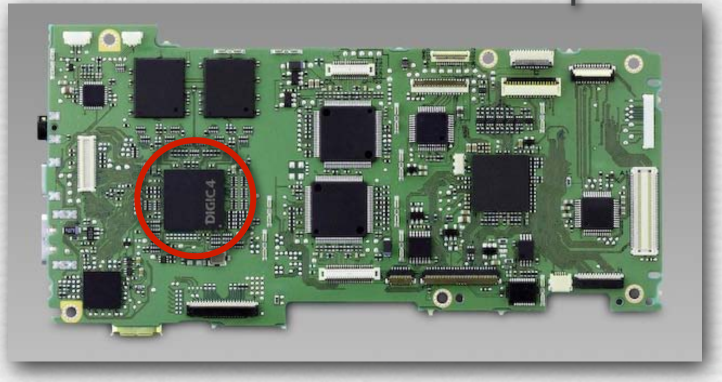
- every camera uses different algorithms
- the processing order may vary
- most of it is proprietary

Example pipeline

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



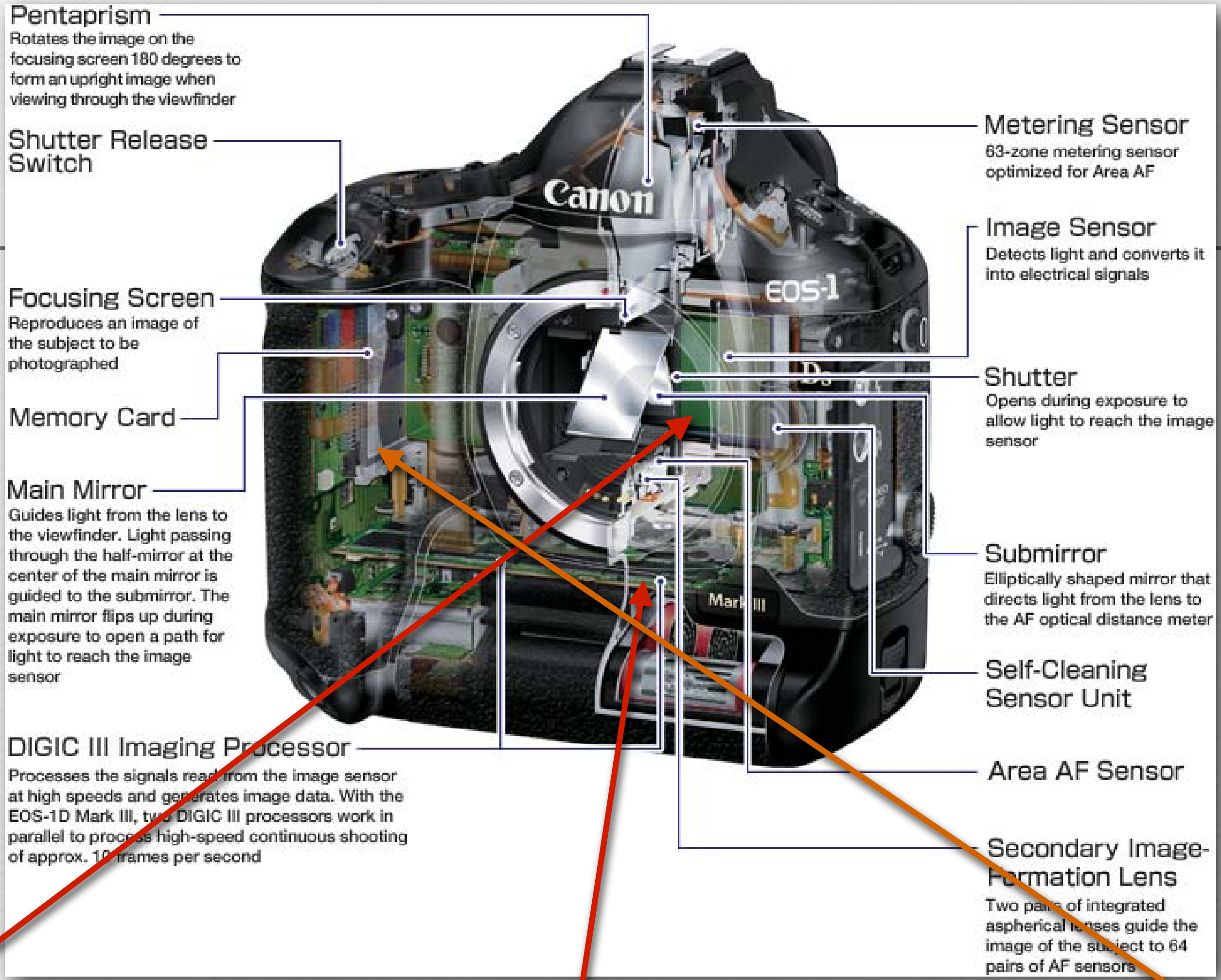
Canon 21 Mpix CMOS sensor



Canon DIGIC 4 processor

Compact Flash card

Example



Pentaprism
Rotates the image on the focusing screen 180 degrees to form an upright image when viewing through the viewfinder

Shutter Release Switch

Focusing Screen
Reproduces an image of the subject to be photographed

Memory Card

Main Mirror
Guides light from the lens to the viewfinder. Light passing through the half-mirror at the center of the main mirror is guided to the submirror. The main mirror flips up during exposure to open a path for light to reach the image sensor

DIGIC III Imaging Processor
Processes the signals read from the image sensor at high speeds and generates image data. With the EOS-1D Mark III, two DIGIC III processors work in parallel to process high-speed continuous shooting of approx. 10 frames per second

Metering Sensor
63-zone metering sensor optimized for Area AF

Image Sensor
Detects light and converts it into electrical signals

Shutter
Opens during exposure to allow light to reach the image sensor

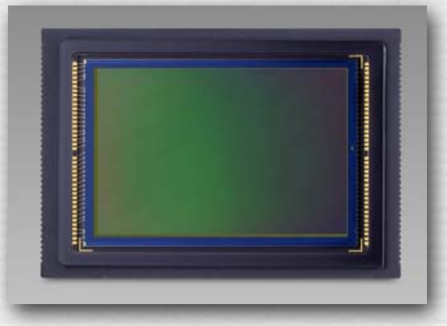
Submirror
Elliptically shaped mirror that directs light from the lens to the AF optical distance meter

Self-Cleaning Sensor Unit

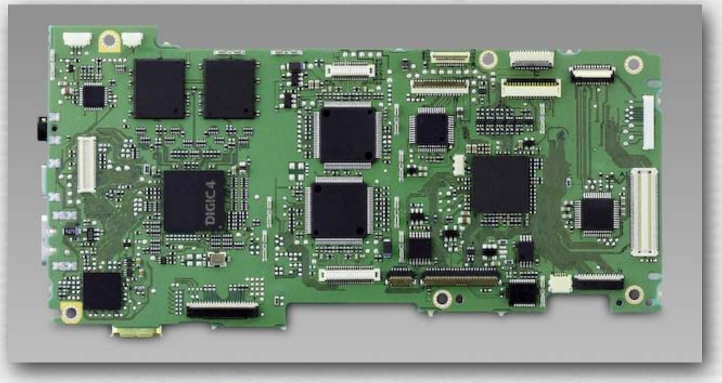
Area AF Sensor

Secondary Image-Formation Lens
Two pairs of integrated aspherical lenses guide the image of the subject to 64 pairs of AF sensors

(parts are from a Canon 5DII, but cutaway view is of 1DIII)



Canon 21 Mpix CMOS sensor



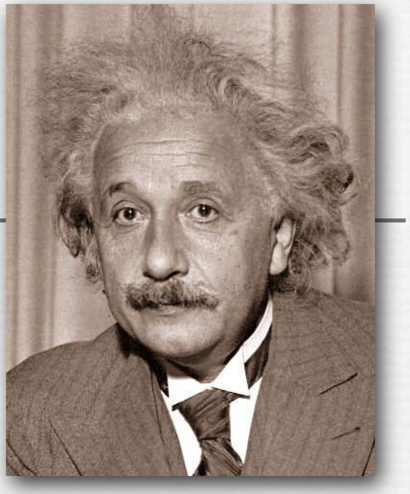
Canon DIGIC 4 processor

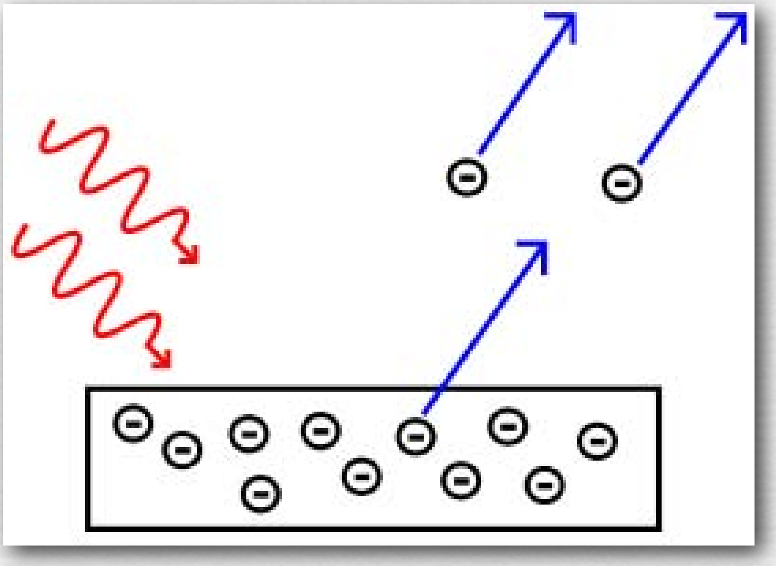
Compact Flash card

Outline

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

The photoelectric effect


Albert Einstein


(wikipedia)

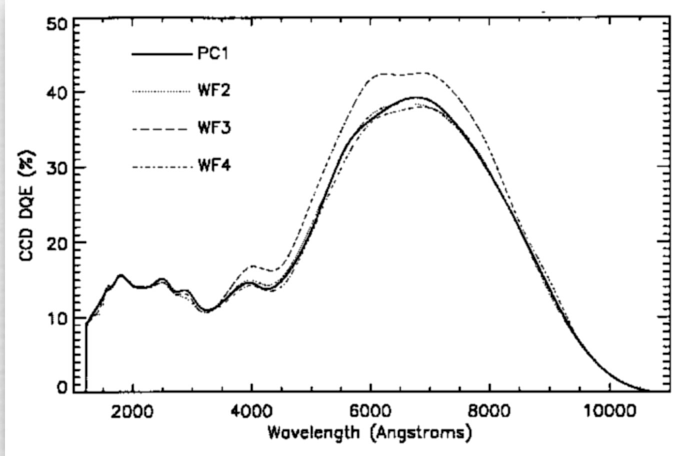
- when a photon strikes a material, an electron may be emitted
    - depends on the photon's energy, which depends on its wavelength

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

    - there is no notion of "brighter photons", only more or fewer of them

Quantum efficiency

Hubble Space
Telescope
Camera 2



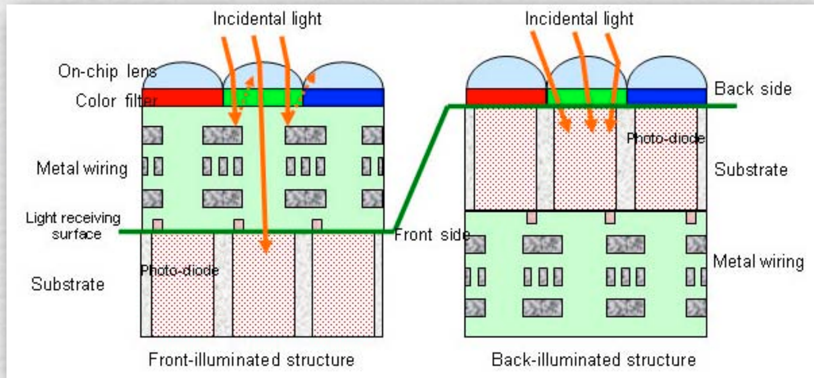
the iPhone 5 uses a back-illuminated CMOS sensor

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

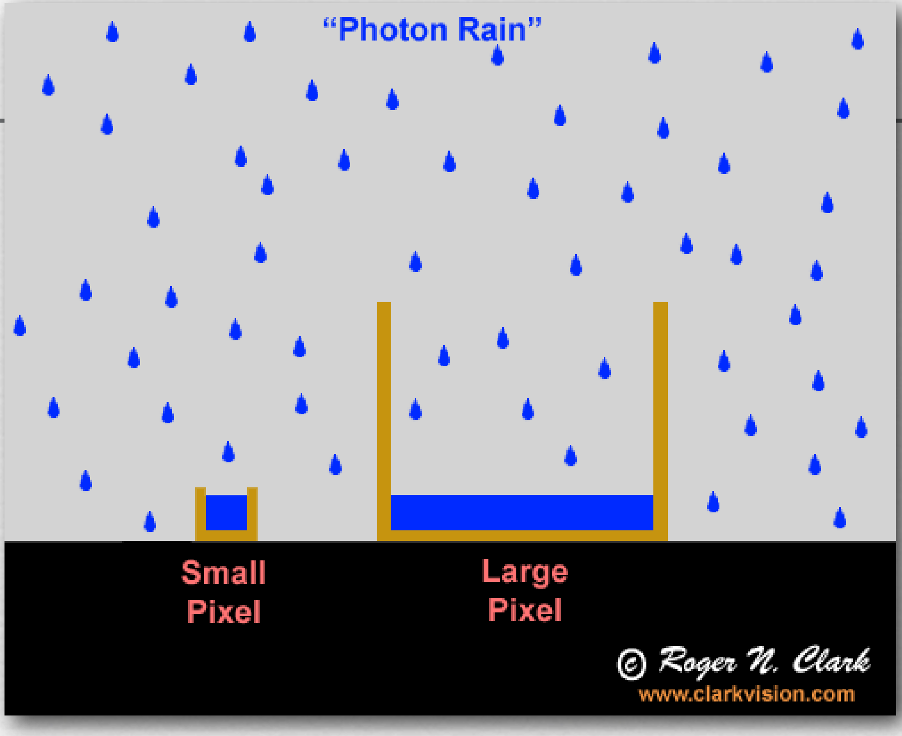
$$QE = \frac{\#\, electrons}{\#\, photons}$$

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

back-illuminated
CMOS (Sony)

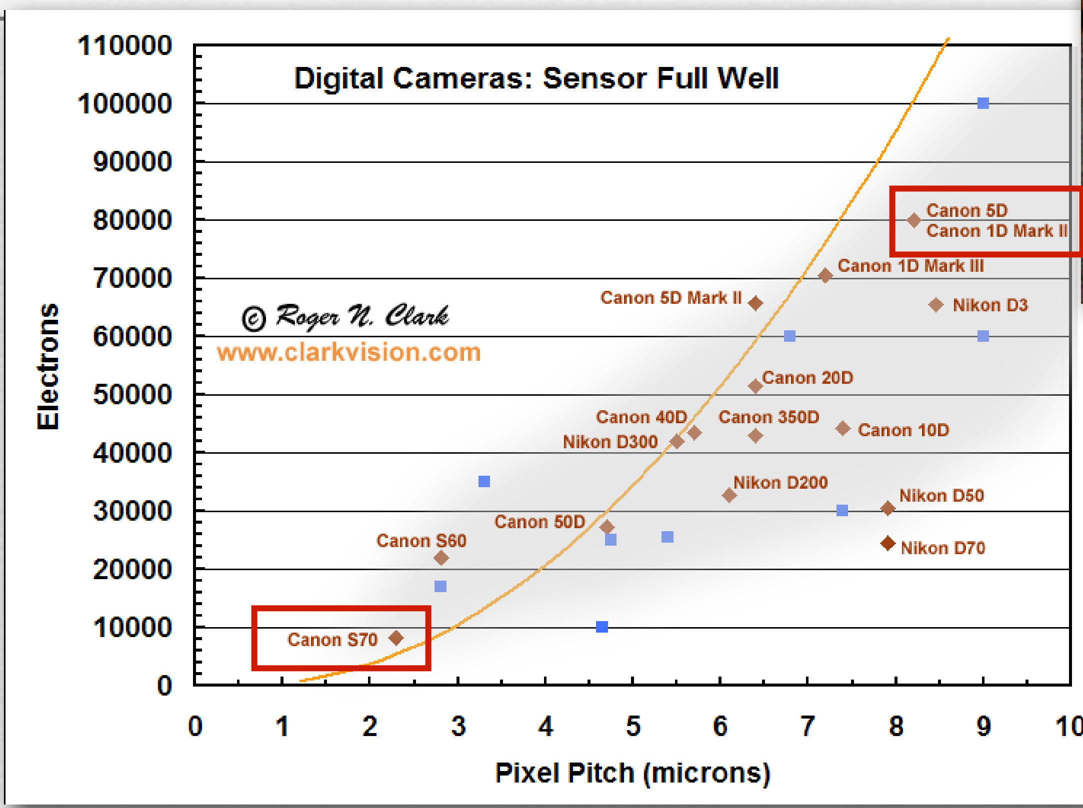


Pixel size



- ✦ the current from one electron 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)

- how many electrons can a pixel hold?
  - depends mainly on the size of the pixel (but fill factor is important)
- too many photons causes *saturation*
  - larger capacity leads to higher *dynamic range* between the brightest scene feature that won't saturate and the darkest that isn't too noisy

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

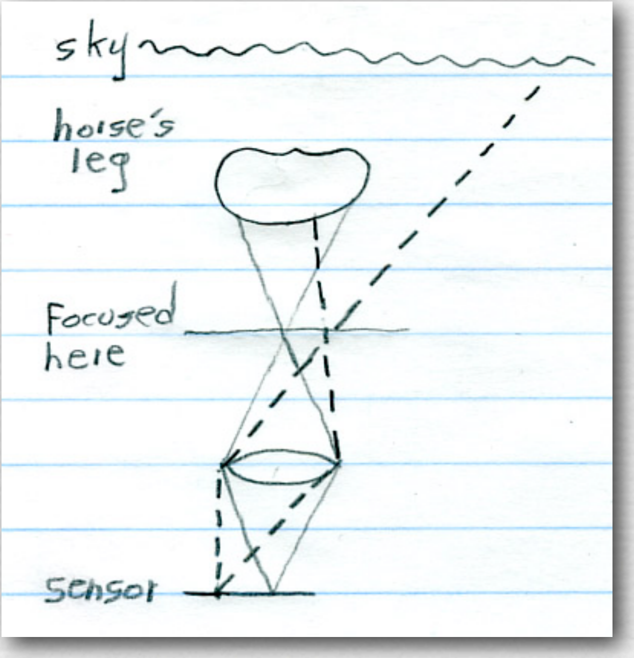
Image artifacts can be hard to diagnose

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

Q. Is this blooming?

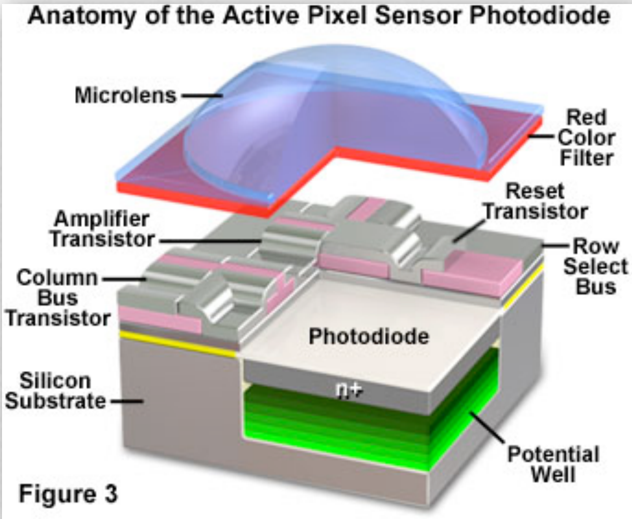
Explanation of preceding image (contents of whiteboard)

- 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 solid 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 dashed bundle of rays, corresponding to a nearby pixel, crossed at the same depth but to the side of the solid 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

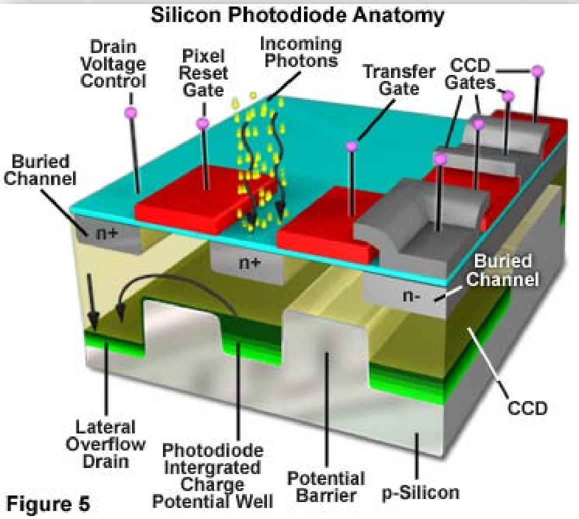


CMOS versus CCD sensors

CMOS



CCD



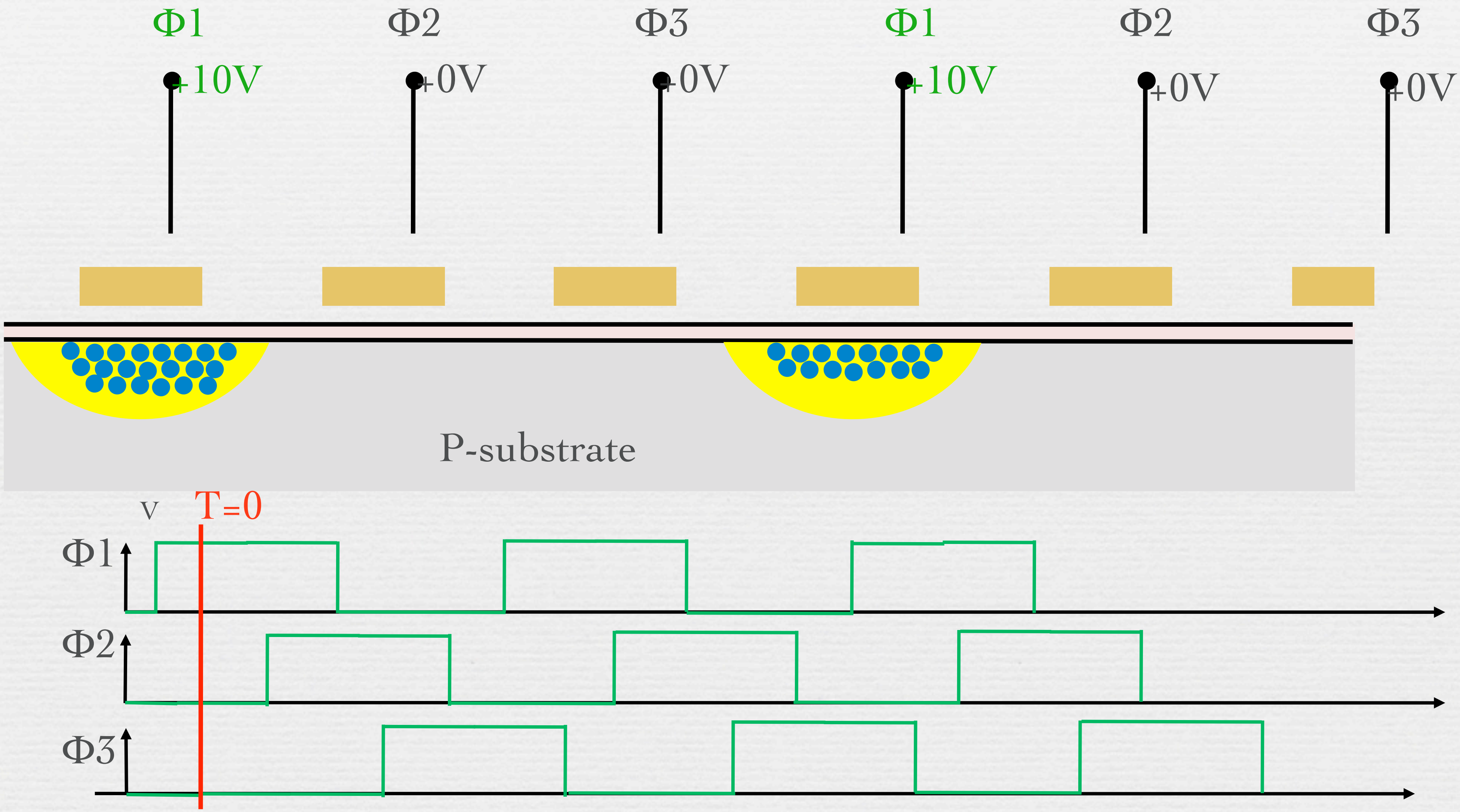
- CMOS = complementary metal-oxide semiconductor
  - an amplifier per pixel converts charge to voltage
  - low power, but noisy (but getting better)

Canon SLRs

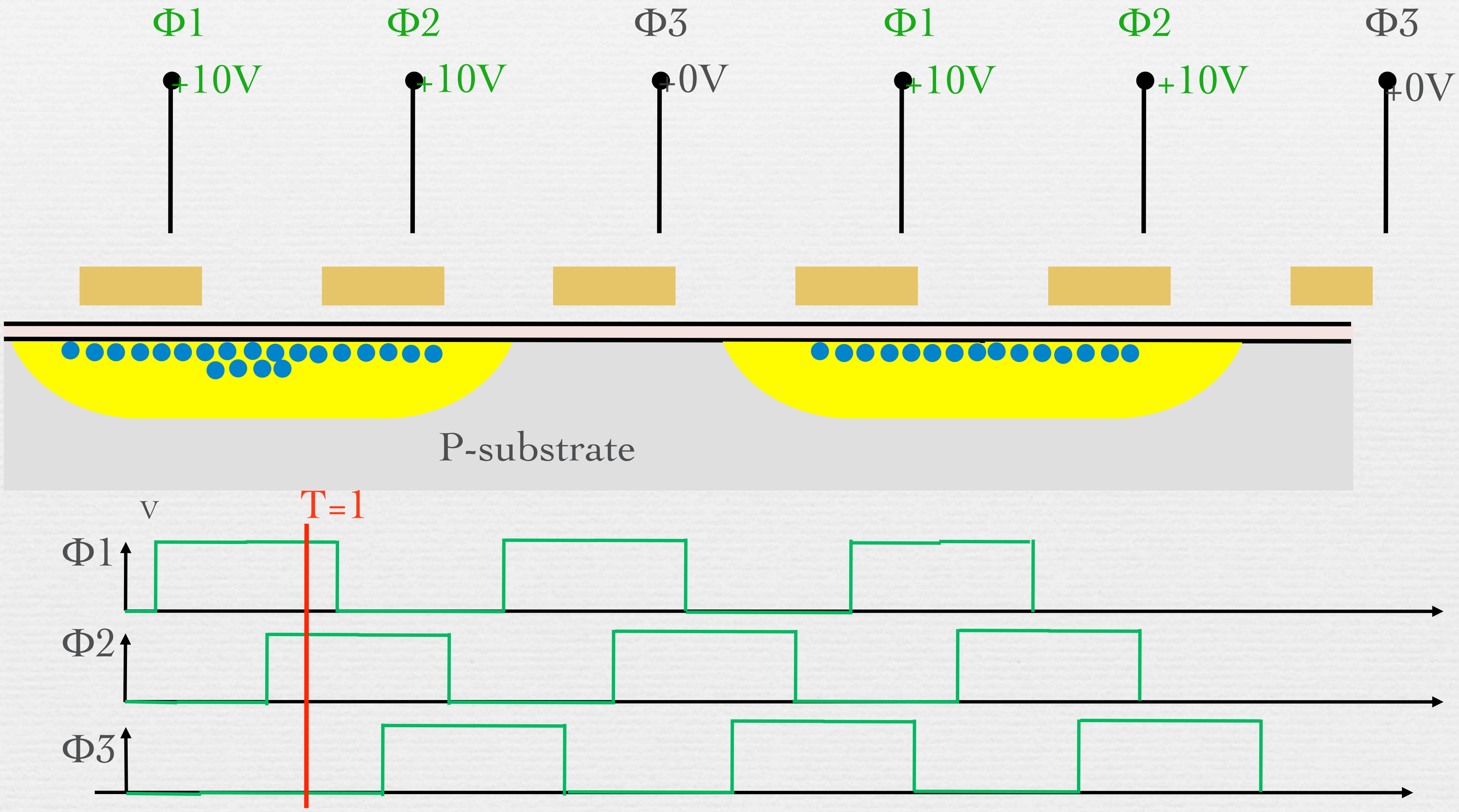
- CCD = charge-coupled device
  - charge shifted along columns to an output amplifier
  - oldest solid-state image sensor technology
  - highest image quality, but not as flexible or cheap as CMOS

Nikon D40

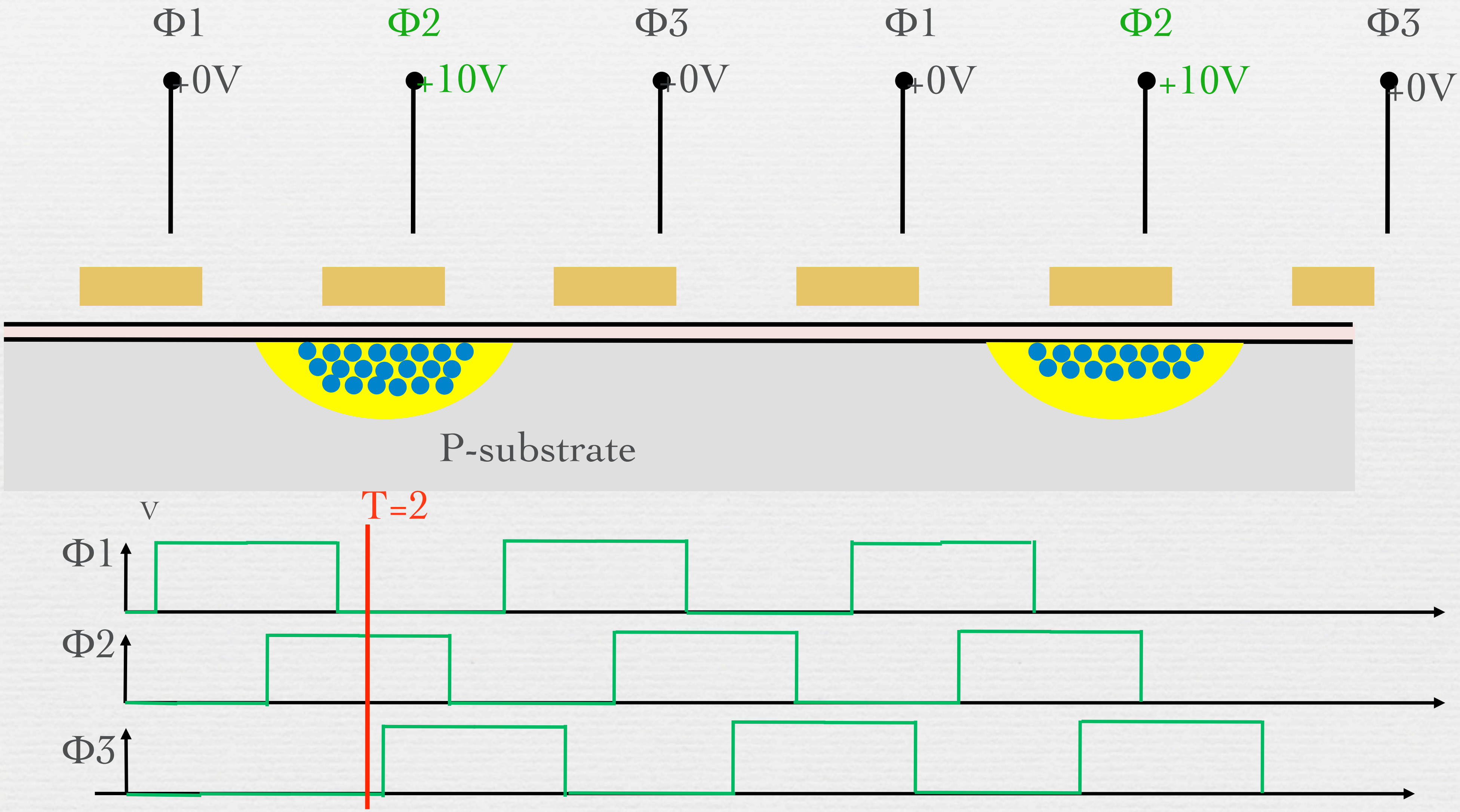
Gratuitous animation showing a CCD “bucket brigade” readout



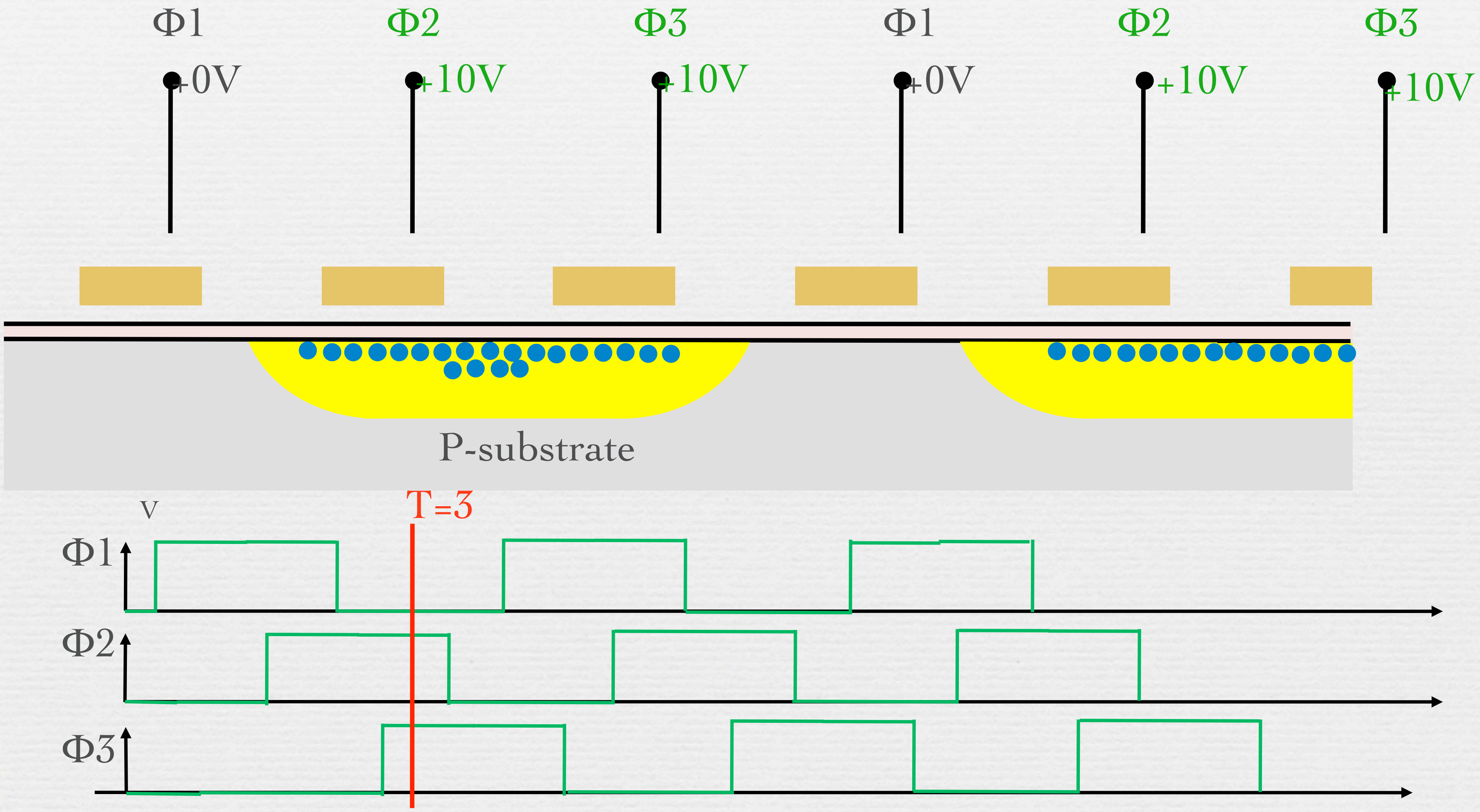
Gratuitous animation showing a CCD "bucket brigade" readout



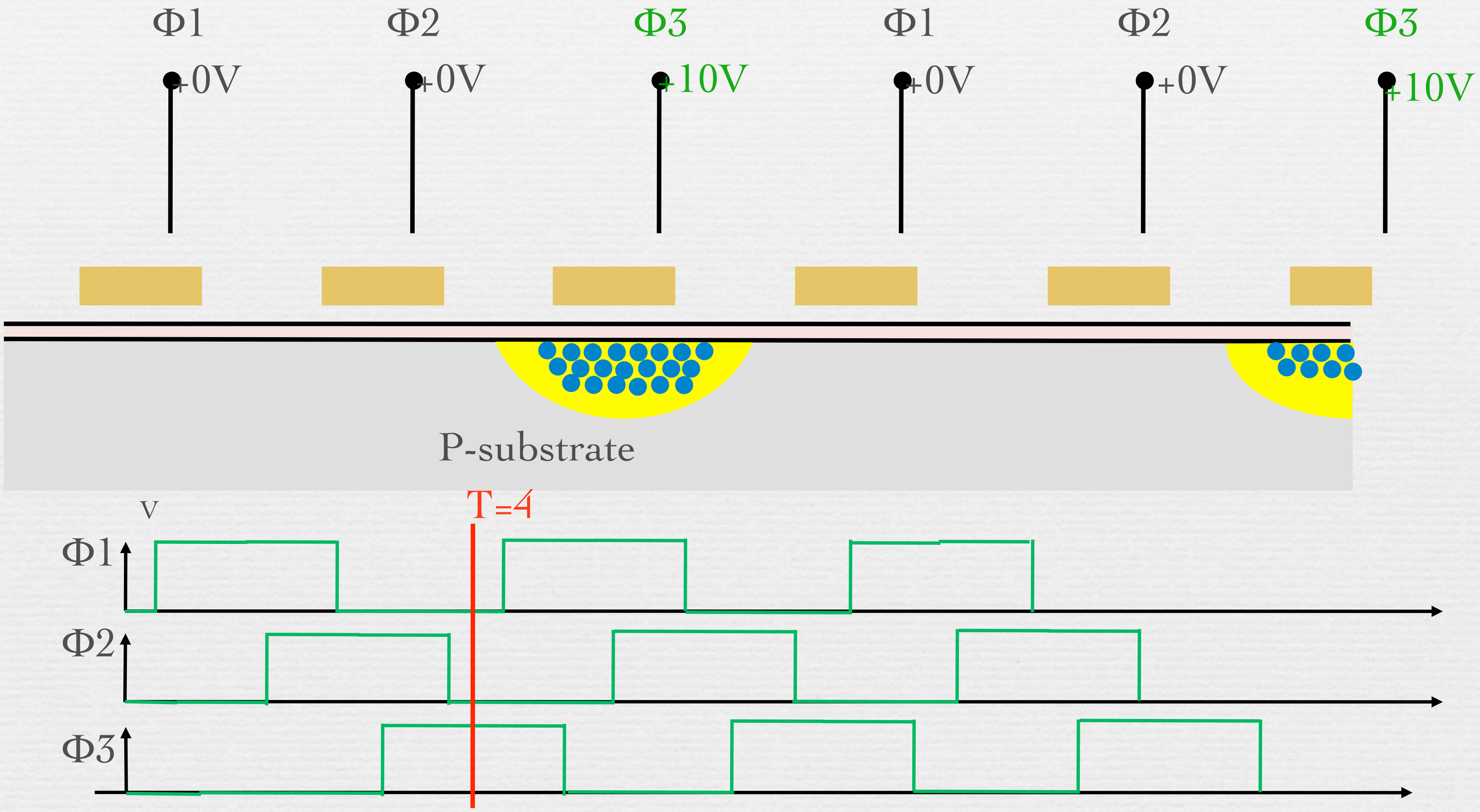
Gratuitous animation showing a CCD "bucket brigade" readout



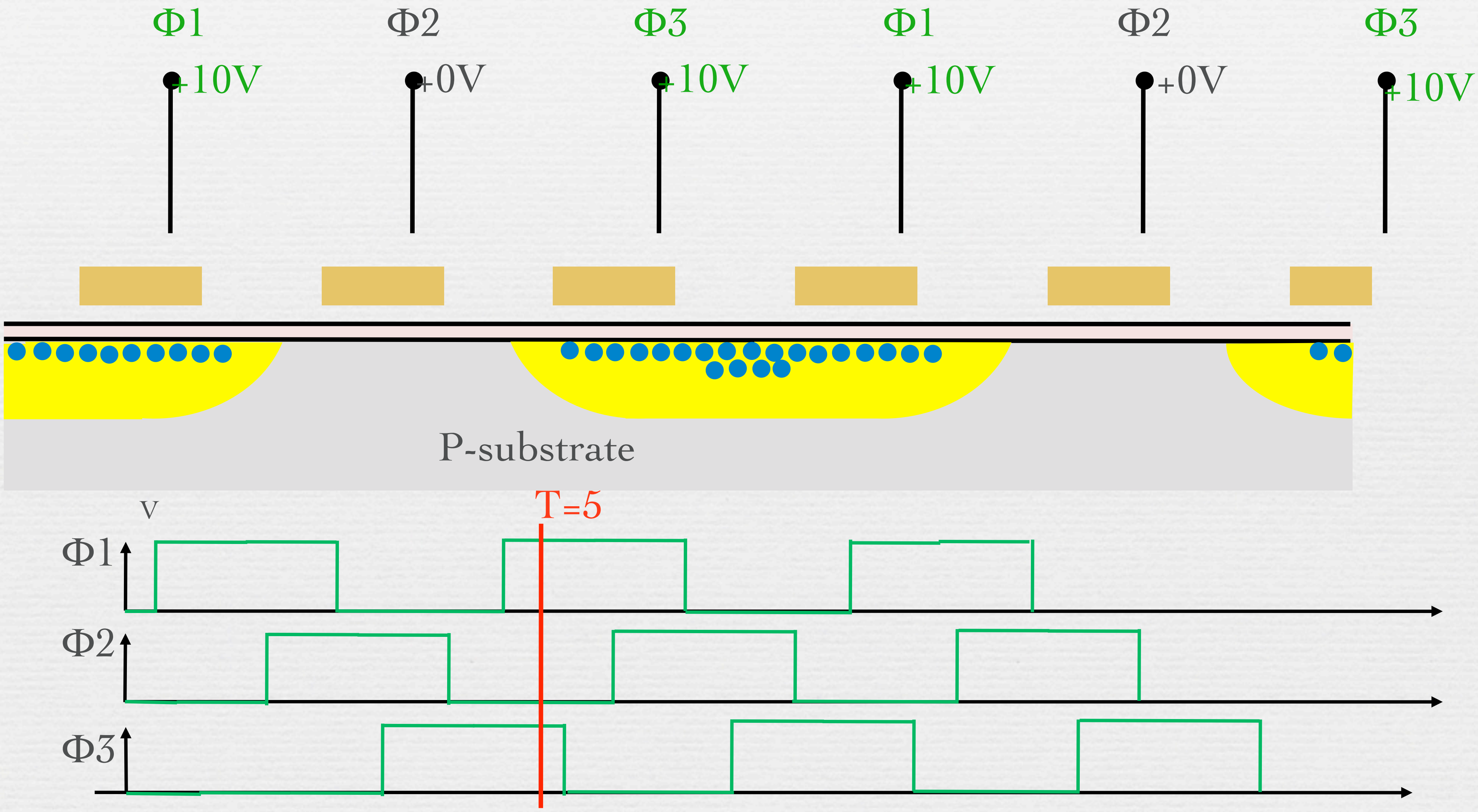
Gratuitous animation showing a CCD "bucket brigade" readout



Gratuitous animation showing a CCD "bucket brigade" readout

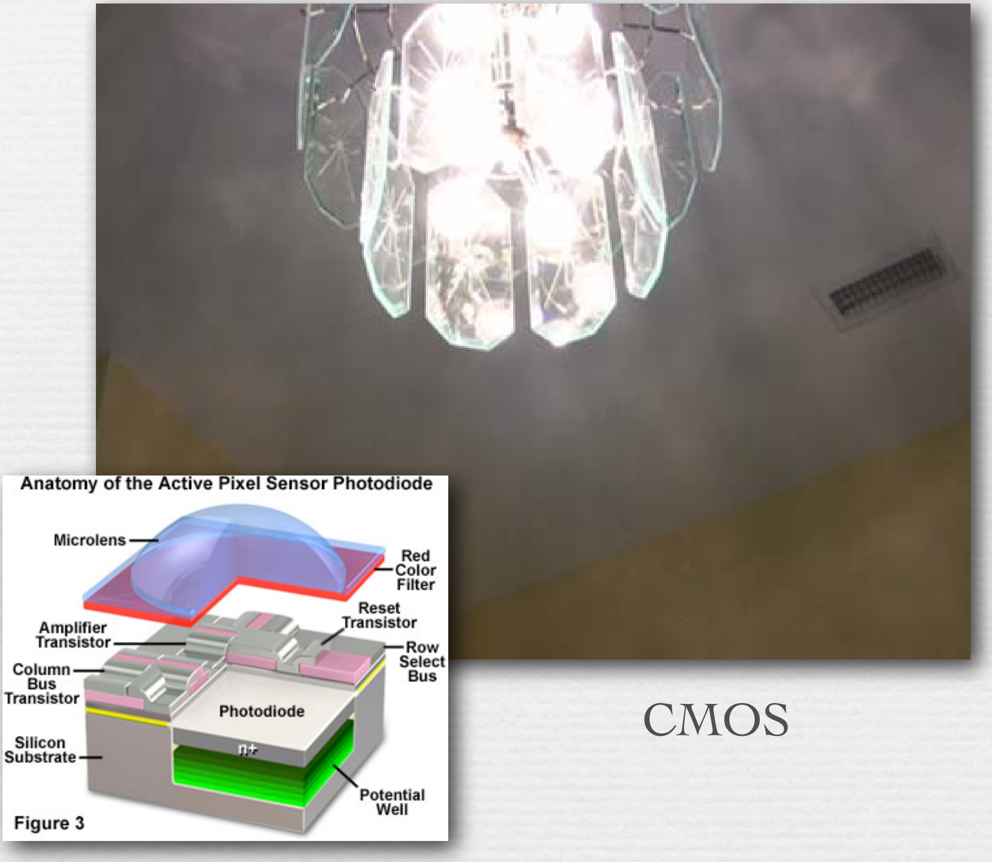


Gratuitous animation showing a CCD “bucket brigade” readout

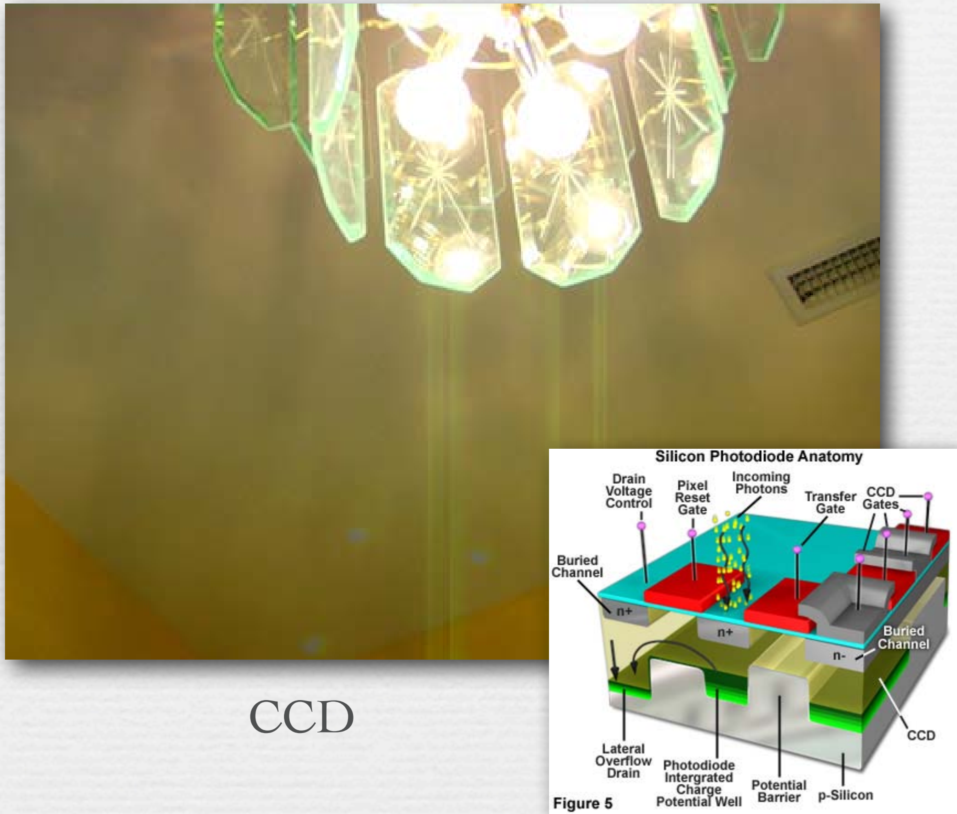


Smearing

(dvxuser.com)



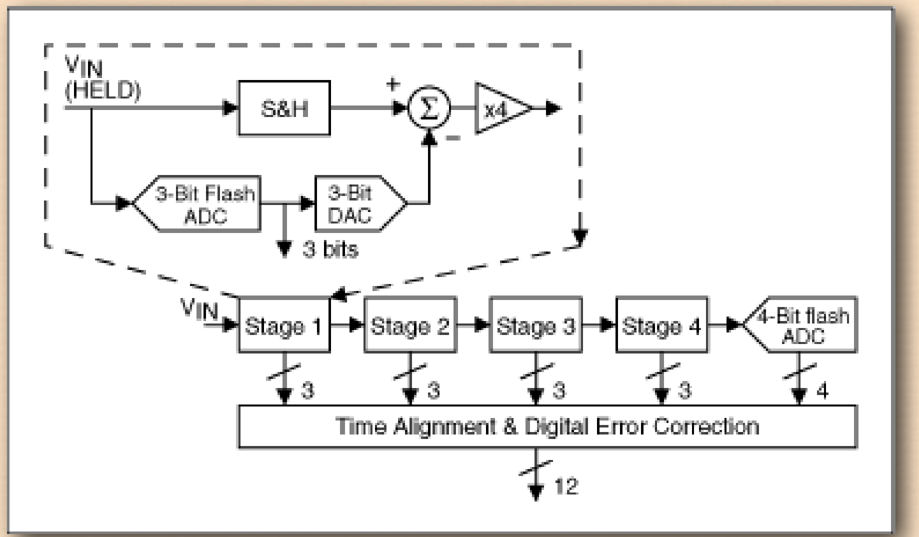
CMOS

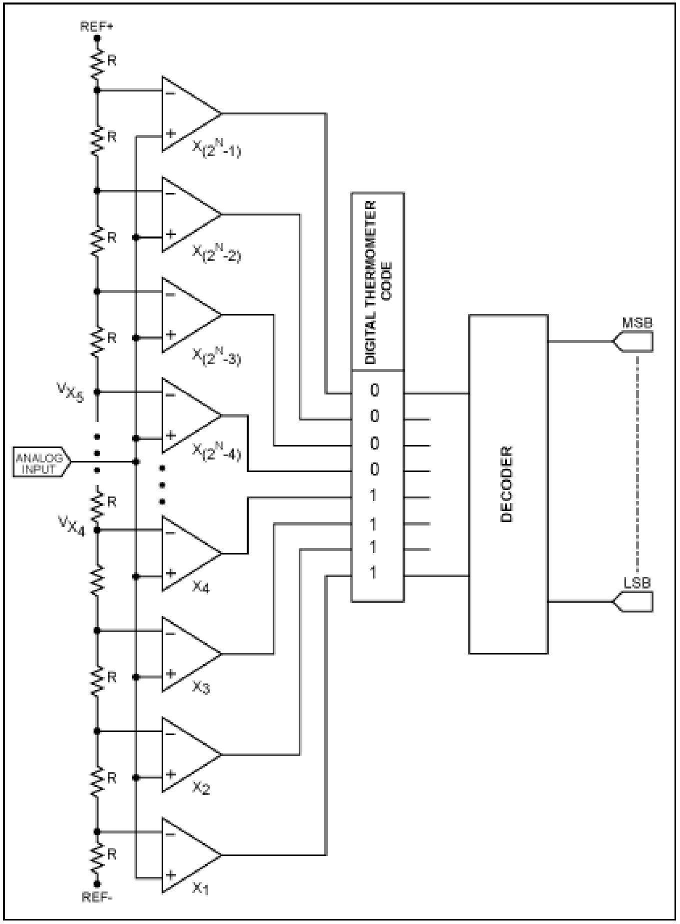


CCD

- ✦ 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)

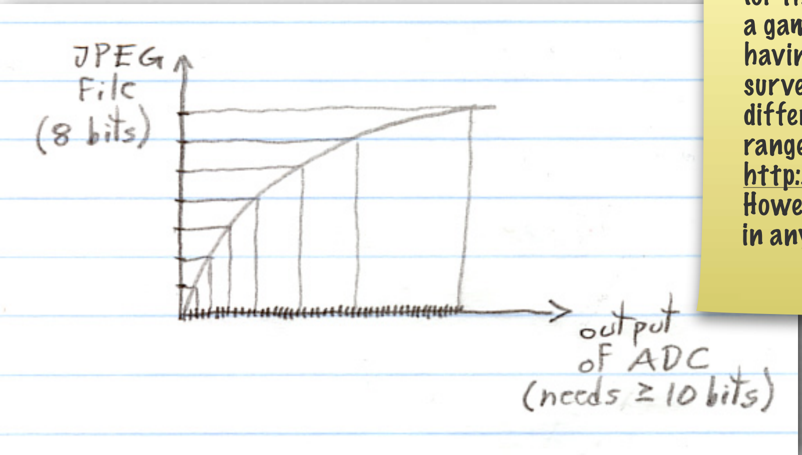




(maxim-ic.com)

- 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

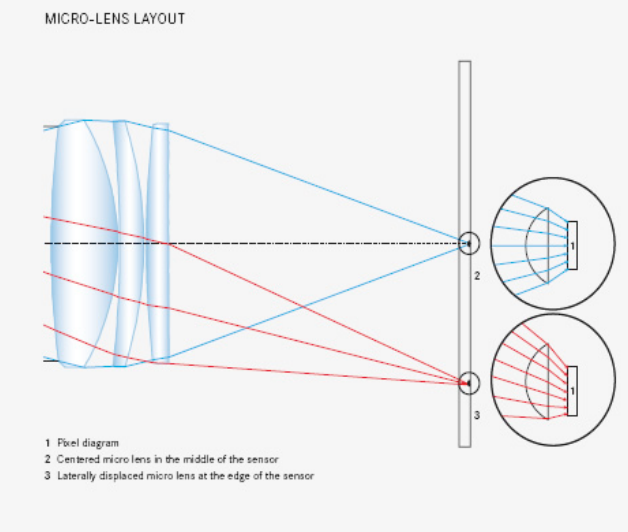
ADC must output more bits than JPEG stores (contents of whiteboard)



A student asked why not make the output of the ADC (or the RAW file, which is equivalent) logarithmic (or a gamma function, which is similar), instead of having more bits as suggested here. Indeed, some surveillance cameras do exactly this, but for a different reason - as a way of achieving HDR dynamic range within a given bitdepth in a single shot. See http://en.wikipedia.org/wiki/Wide_dynamic_range. However, this strategy is not (to my knowledge) used in any consumer photographic camera.

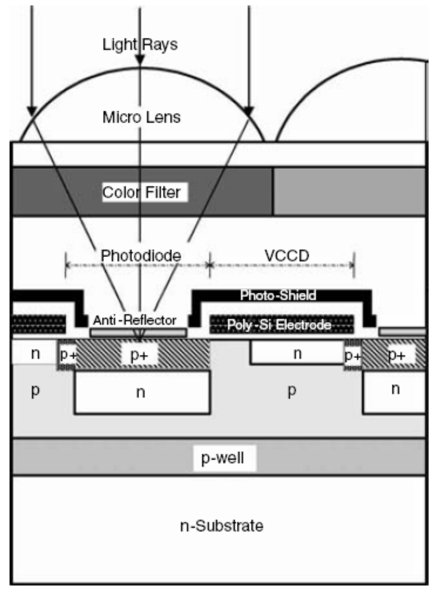
- converting from analog-to-digital converter (ADC) values (as stored in a RAW file) to the values stored in a JPEG file includes a *tone mapping;* as introduced in the exposure metering lecture, this mapping is typically non-linear and includes a step called *gamma correction*, which has the form output = input^{γ} $(0.0 \leq \text{input} \leq 1.0)$
- since JPEG files only store 8 bits/pixel for each color component, in order for a scene consisting of a smooth gray ramp to fill each of these 256 buckets, the camera's ADC needs to output ≥ ~10 bits; otherwise, dark parts of the ramp will exhibit banding after applying gamma correction and requantizing (integerizing)

Fill factor



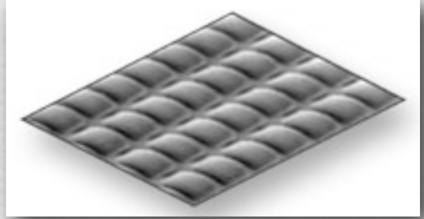
shifted microlenses on M9

Leica M9
(digital full-frame)



on a CCD sensor

on a (front-illuminated) CMOS sensor



oblique view
of microlenses

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

Spatio-temporal prefiltering in photography

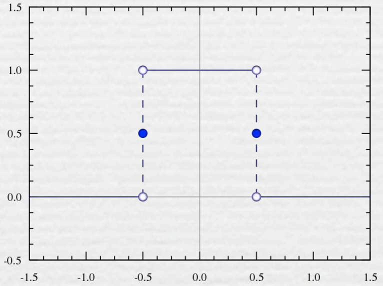
- integrating light over an area at each pixel site instead of point sampling serves two functions:
  - captures more photons, to improve *dynamic range*
  - convolves the image with a prefilter, to avoid *aliasing*

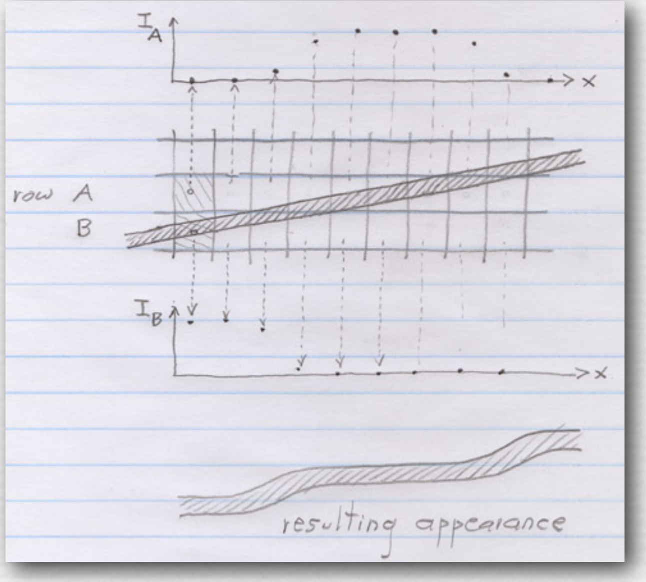
- microlenses gather more light and improve the prefilter
  - microlenses ensure that the *spatial prefilter* is a 2D rect of width roughly equal to the pixel spacing

- integrating light over the exposure time does the same:
  - captures more photons
  - convolves the scene with a *temporal prefilter*, roughly a 1D rect, creating motion blur if the camera or scene moves

However, a rect is not an ideal pre-filter (contents of whiteboard)

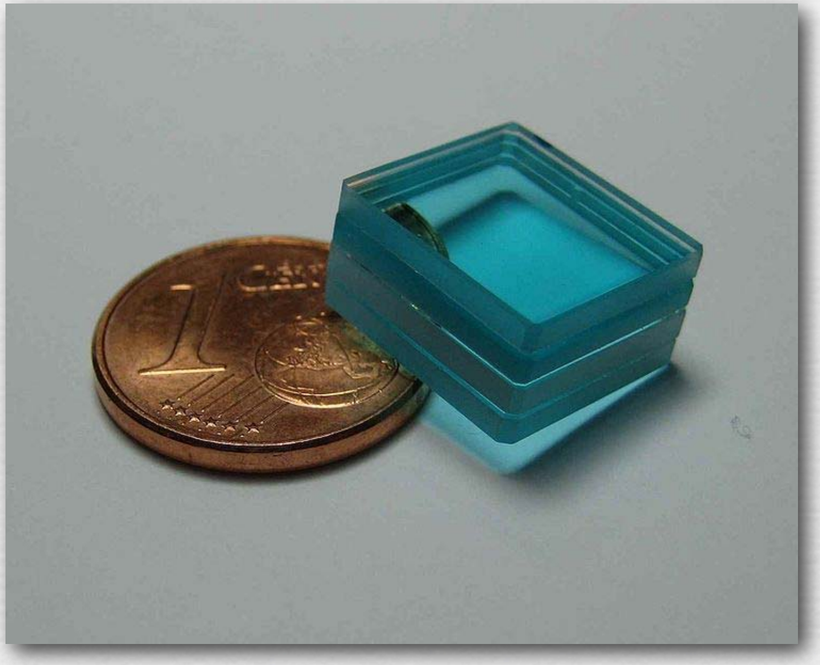
$$rect(x) = \Pi(x) = \begin{cases} 0 & \text{if } |x| > \frac{1}{2} \\ \frac{1}{2} & \text{if } |x| = \frac{1}{2} \\ 1 & \text{if } |x| < \frac{1}{2} \end{cases}$$



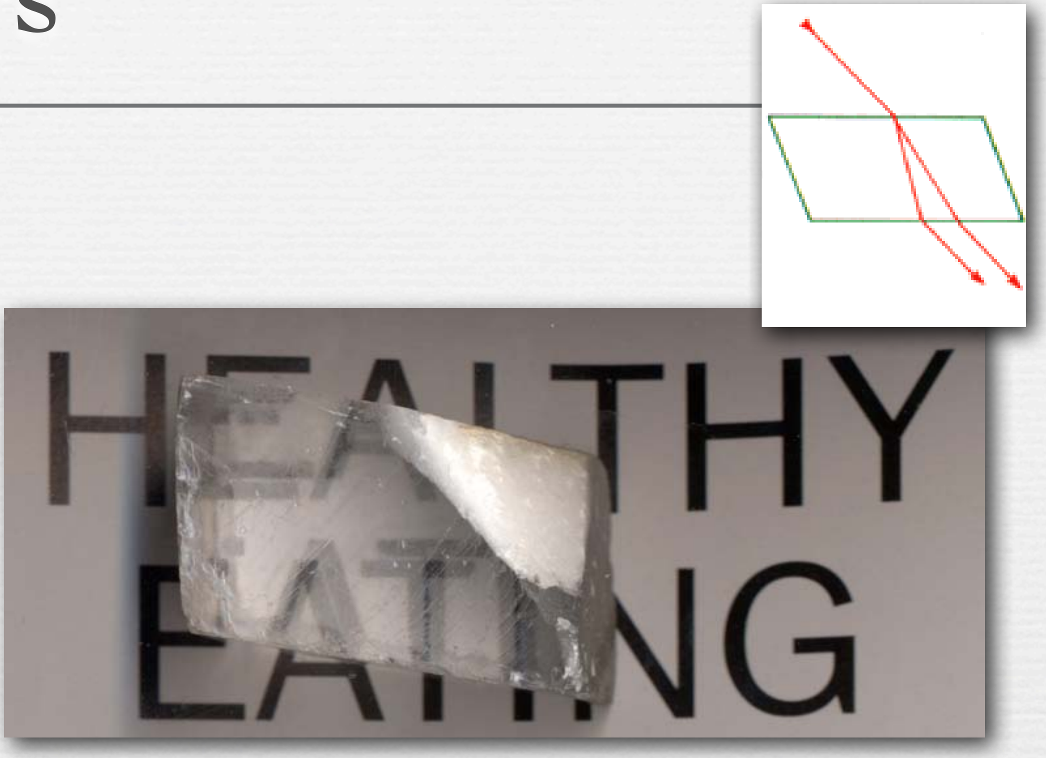


- as you know, convolving a focused image by a 2D rect (a 1D rect is defined at left above) of width equal to the pixel spacing is equivalent to computing the average intensities in the squares forming each pixel
- assuming such a 2D rect, a narrow angled stripe object will produce for row A the intensities shown in plot I_A , rising quickly, staying constant for a while, then dropping; the resulting ropey appearance is aliasing
- if this were a film and each frame were a 1D rect over time, a small object would appear to move quickly, then pause, then move again

Antialiasing filters



infrared & antialiasing filter

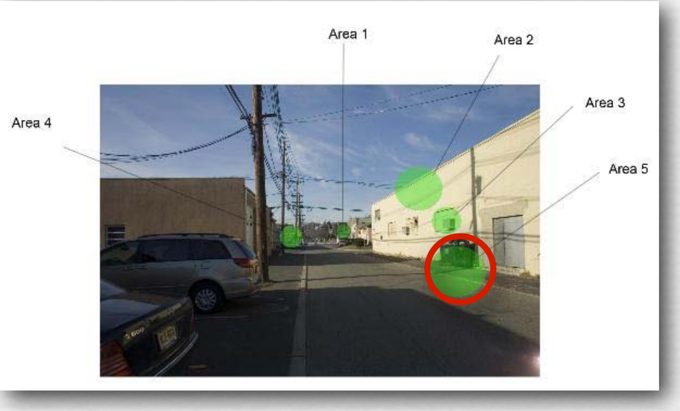


birefringence in a calcite crystal

- 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



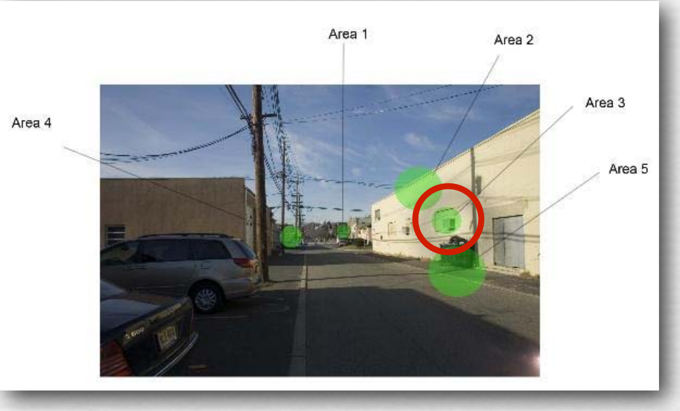
(maxmax.com)

anti-aliasing filter removed

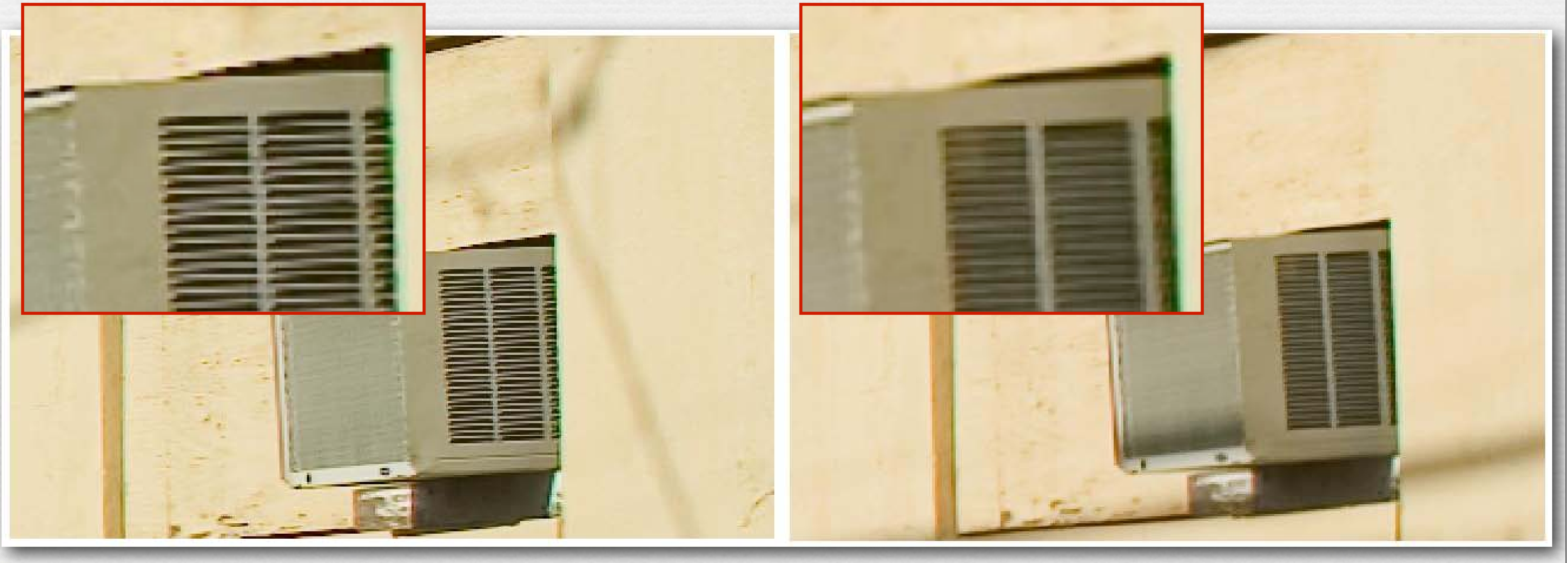
normal

Removing the antialiasing filter

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



(maxmax.com)



anti-aliasing filter removed

normal

Recap

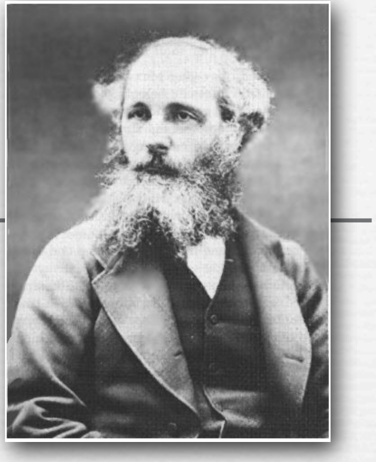
- photons strike a sensor and are converted to electrons
  - performance factors include *quantum efficiency* and *pixel size*
- sensors are typically CCD or CMOS
  - both can suffer *blooming*; only CCDs can suffer *smearing*
- integrating light over an area serves <u>two</u> functions
  - capturing more photons, to improve *dynamic range*
  - convolving the image with a prefilter, to avoid *aliasing*
  - to ensure that the area spans pixel spacing, use *microlenses*
  - to improve further on the prefilter, use an *antialiasing filter*
- integrating light over time serves the same two functions
  - captures more photons, but may produce motion blur

Questions?

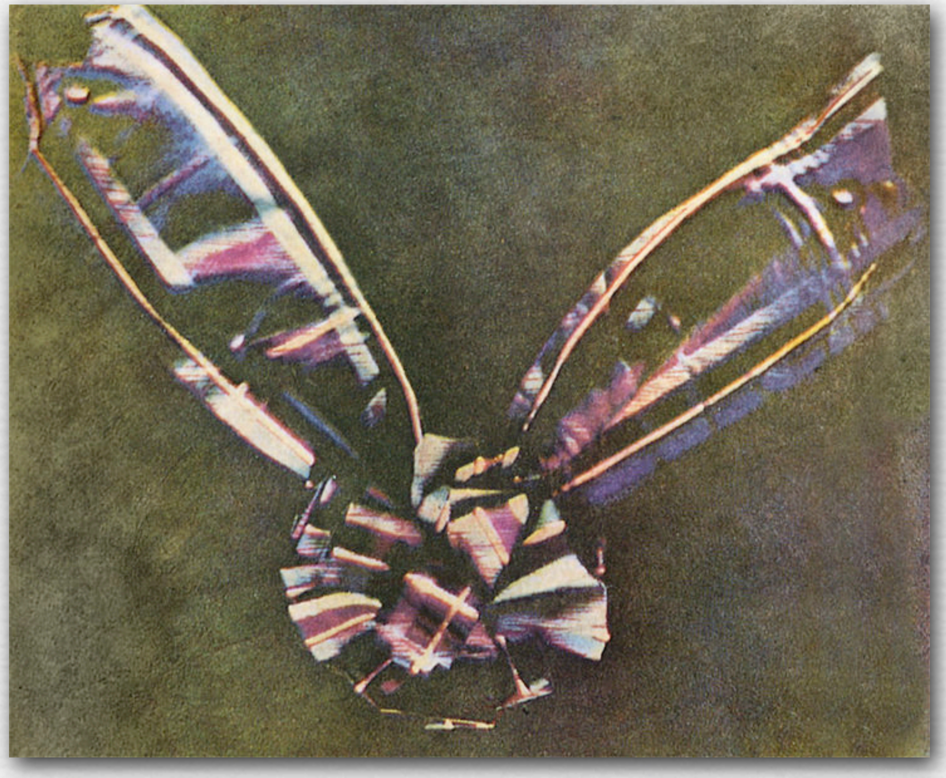
Color

- silicon detects all visible frequencies well
- can't differentiate wavelengths after photon knocks an electron loose
  - all electrons look alike
- must select desired frequencies before light reaches photodetector
  - block using a filter, or separate using a prism or dichroic
- 3 spectral responses is enough
  - a few consumer cameras record 4
- silicon is also sensitive to near infrared (NIR)
  - most sensors have an IR filter to block it
  - to make a NIR camera, remove this filter

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

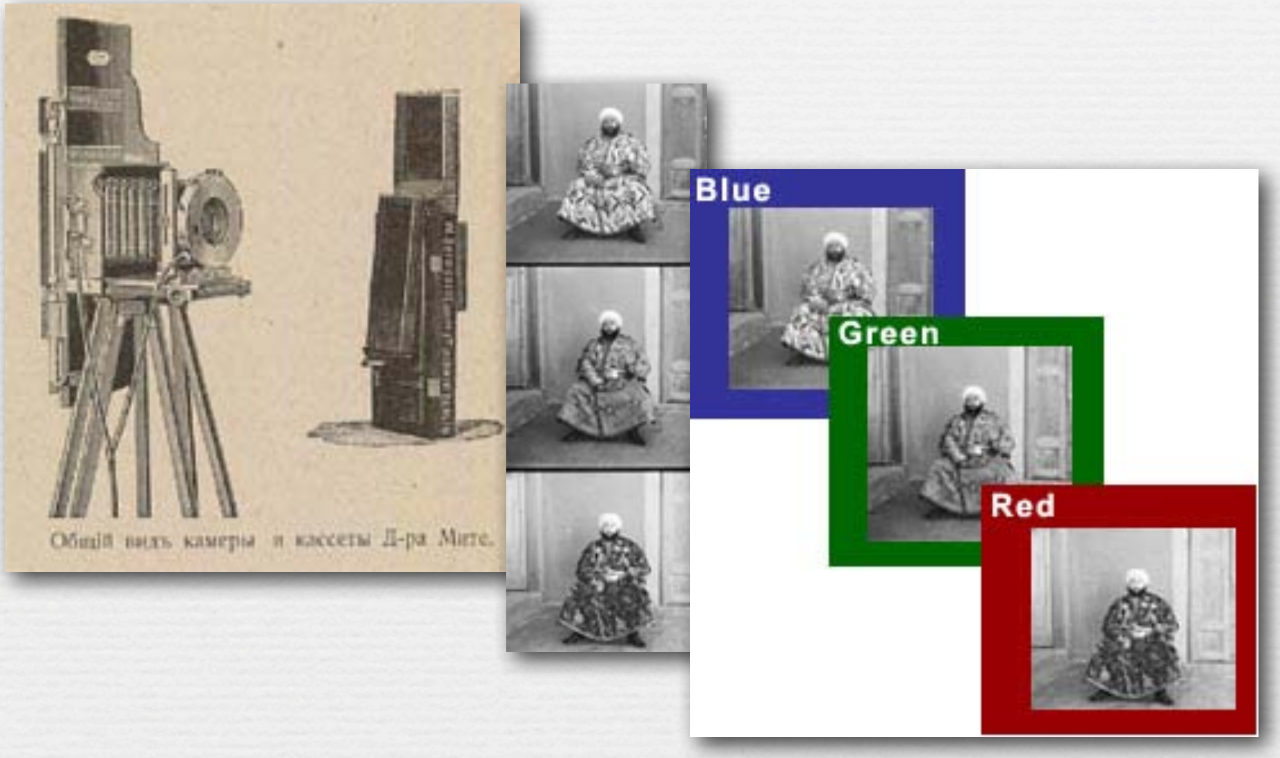
Fig. 16. — Louis Ducos du Hauron (1).

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



- shot sequentially through R, G, B filters
- 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)

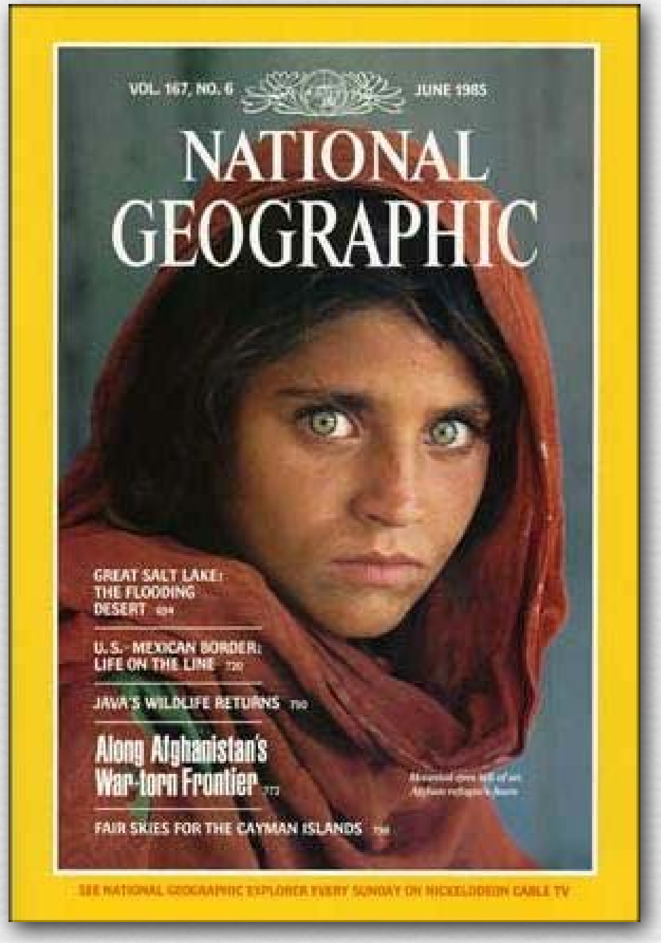
Wizard of Oz (1939)

- 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?

(Beatles in 1964 was in B&W)

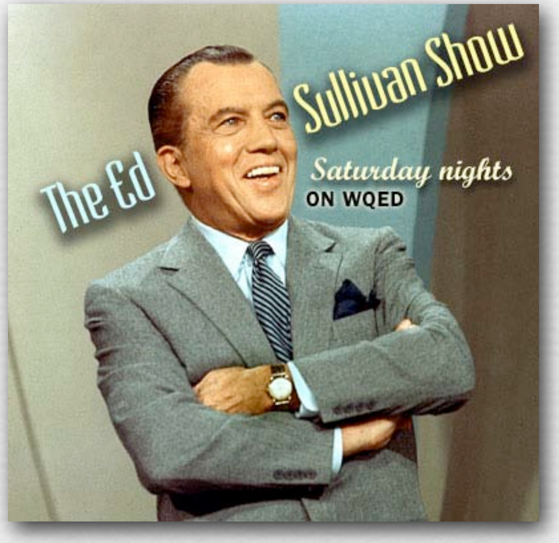
started broadcasting in color in 1965

- competing standards

U.S.	NTSC	525-line, 30fps, interlaced
Europe	PAL	625-line, 25fps, interlaced
France	SECAM	625-line, 25fps, interlaced

First color television broadcast?

(Beatles in 1964 was in B&W)



started broadcasting in color in 1965

- competing standards

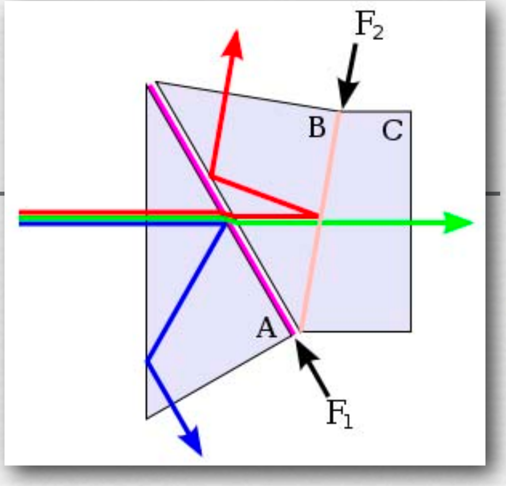
• U.S.	NTSC	Never Twice the Same Color
• Europe	PAL	Pale and Lurid
• France	SECAM	Système Electronique Contre les Americains

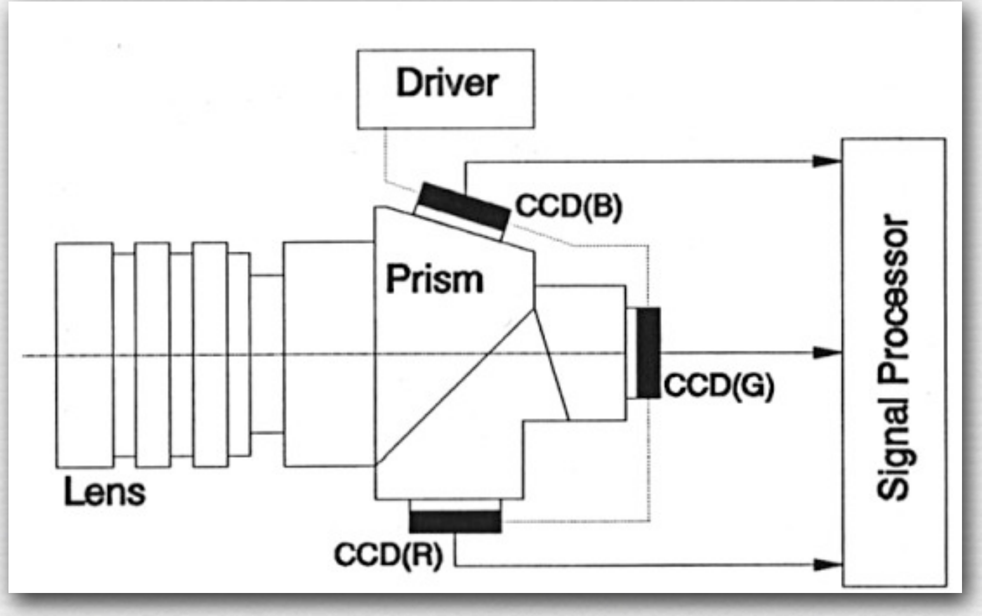
Color sensing technologies

- field-sequential - just covered
- 3-chip
- vertically stacked
- color filter arrays

3-chip cameras

(Wikipedia)

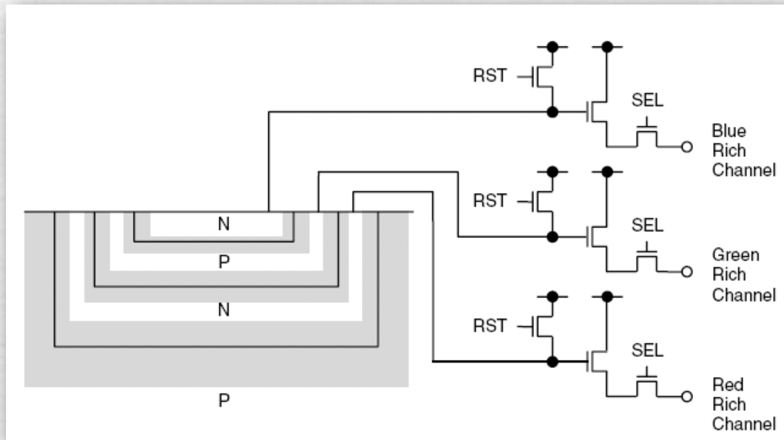




(Theuwissen)

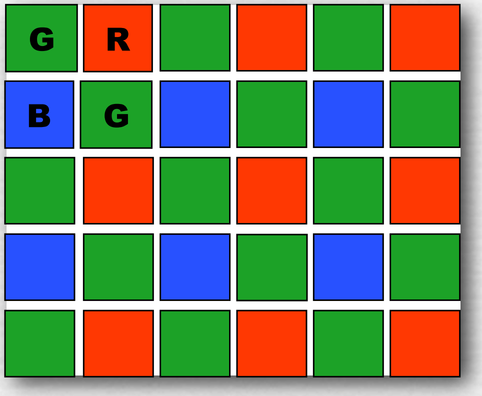
- high-quality video cameras
- prism & dichroic mirrors split the image into 3 colors, each routed to a separate sensor (typically CCD)
- no light loss, as compared to filters (which absorb light)
- expensive, and complicates lens design

Foveon stacked sensor

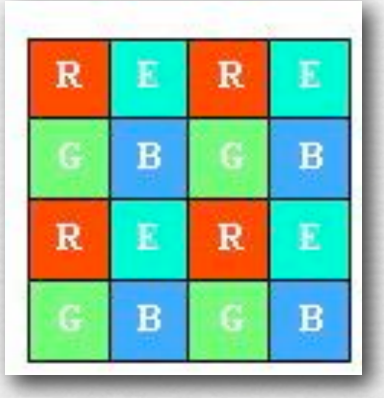


- longer wavelengths penetrate deeper into silicon, so arrange a set of vertically stacked detectors
  - top gets mostly blue, middle gets green, bottom gets red
  - no control over spectral responses, so requires processing
- fewer moiré artifacts than color filter arrays + demosaicing
  - but possibly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

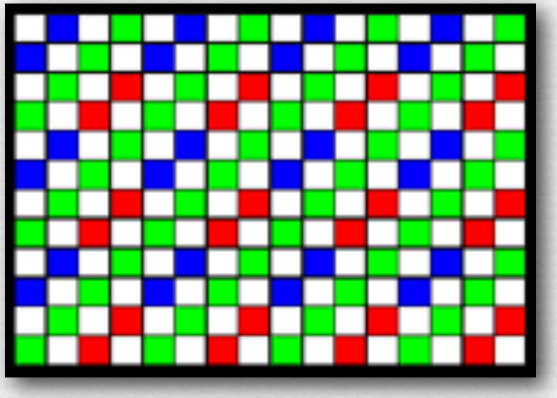
Color filter arrays



Bayer pattern

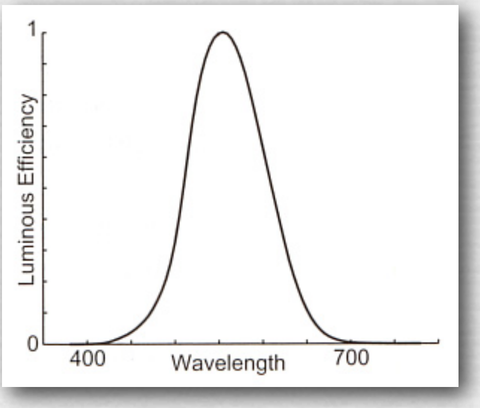


Sony RGB+E
better color



Kodak RGB+C
more dynamic range

- Why more green pixels than red or blue?
  - because humans are most sensitive in the middle of the visible spectrum
  - sensitivity given by the human luminous efficiency curve



(Stone)

Example of Bayer mosaic image

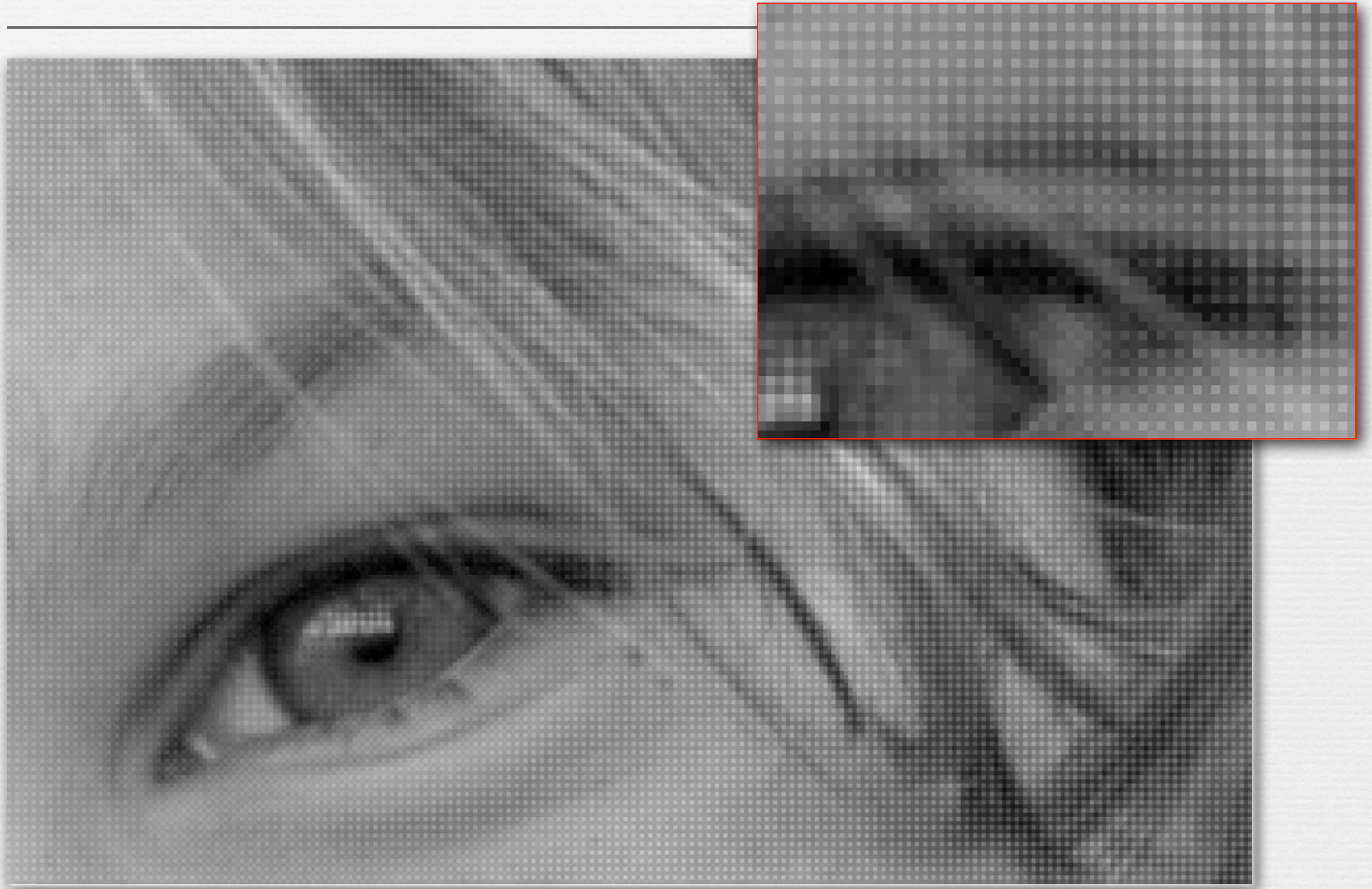
Small fan at Stanford women's soccer game

(Canon 1D III)

Example of Bayer mosaic image

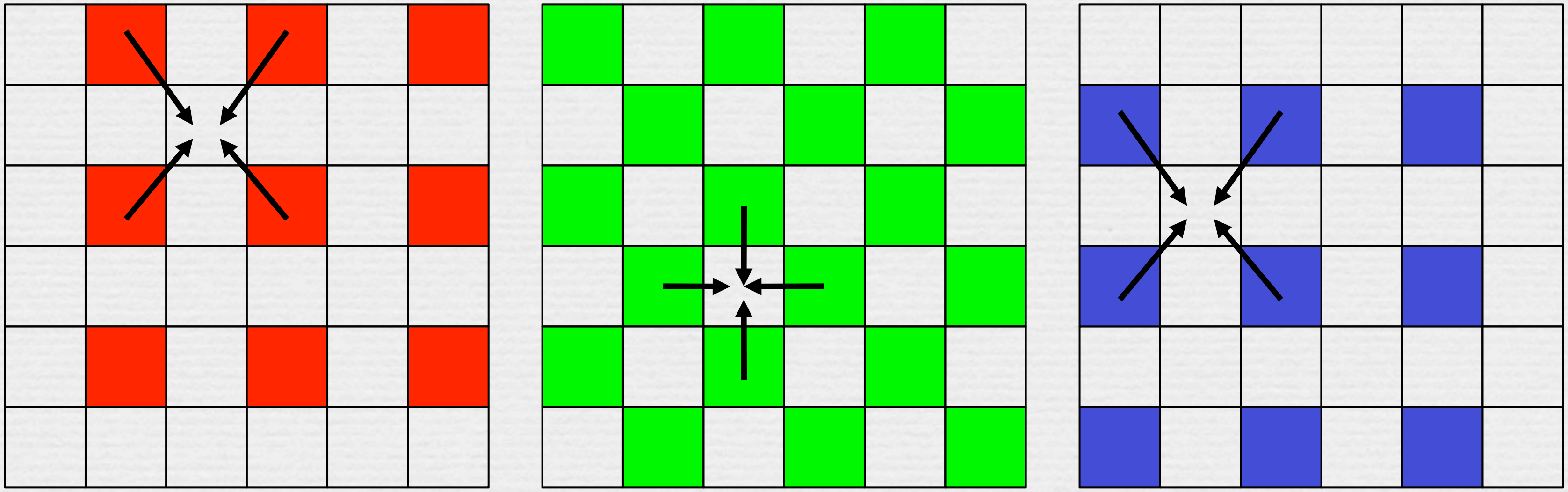
© Marc Levoy

Before demosaicing (`dcraw -d`)



Demosaicing

- linear interpolation
  - average of the 4 nearest neighbors of the same color
- cameras typically use more complicated scheme
  - try to avoid interpolating across contrasty edges
  - demosaicing is often combined with denoising, sharpening...
- due to demosaicing, 2/3 of your data is "made up"!



Recap

- color can only be measured by selecting certain light frequencies to reach certain sensor sites or layers
  - selection can employ *filters* or *dichroics* or *penetration depth*

- measuring color requires making a tradeoff
  - *field sequential* cameras trade off capture duration
  - *3-chip* cameras trade off weight and expense
  - *vertically stacked* sensors (Foveon) trade off noise (in red)
  - *color filter array* (e.g. Bayer) trades off spatial resolution

Questions?

Slide credits

- Brian Curless
- Eddy Talvala
- Abbas El Gamal

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