Photons and sensors
(with an interlude on the history of color photography)

CS 178, Spring 2014

Marc Levoy
Computer Science Department
Stanford University
Camera pixel pipeline

- 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

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

Canon 21 Mpix CMOS sensor
Canon DIGIC 4 processor
Compact Flash card

Canon DIGIC 4 processor

© Marc Levoy
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

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

\[ E_{\text{photon}} = \frac{h \times c}{\lambda} \]

- there is no notion of “brighter photons”, only more or fewer of them
Quantum efficiency

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

\[ QE = \frac{\# \text{ electrons}}{\# \text{ photons}} \]

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

- 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

- charge spilling over to nearby pixels
  - can happen on CCD and CMOS sensors
  - don’t confuse with glare or other image artifacts
Image artifacts can be hard to diagnose

Q. Is this blooming?
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 = complementary metal-oxide semiconductor
  • an amplifier per pixel converts charge to voltage
  • low power, but noisy (but getting better)

✦ 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
Gratuitous animation showing a CCD “bucket brigade” readout

\[ \Phi_1 \quad \Phi_2 \quad \Phi_3 \quad \Phi_1 \quad \Phi_2 \quad \Phi_3 \]

\[ -10V \quad +0V \quad 0V \quad +10V \quad +0V \quad -0V \]

\[ \text{P-substrate} \]

\[ v \quad T=0 \]
Gratuitous animation showing a CCD “bucket brigade” readout

Φ1
-10V

Φ2
+10V

Φ3
0V

Φ1
+10V

Φ2
+10V

Φ3
0V

P-substrate

V

T=1

Φ1

Φ2

Φ3
Gratuitous animation showing a CCD “bucket brigade” readout
Gratuitous animation showing a CCD “bucket brigade” readout
Gratuitous animation showing a CCD “bucket brigade” readout

\[ \Phi_1 \quad \Phi_2 \quad \Phi_3 \quad \Phi_1 \quad \Phi_2 \quad \Phi_3 \]

\[ -0V \quad +0V \quad +10V \quad +0V \quad +0V \quad +10V \]

P-substrate
Gratuitous animation showing a CCD “bucket brigade” readout
Smearing

- side effect of bucket-brigade readout on CCD sensors
  - along columns, so looks different than blooming
  - only happens if pixels saturate
  - doesn’t happen on CMOS sensors

© Marc Levoy
Analog to digital conversion (ADC)

- **flash ADC**
  - voltage divider
  - comparators
  - decoder
  - for n bits requires $2^n$ comparators

- practical systems use a different architecture: pipelined ADC

- recent sensors have one ADC per column of pixels
ADC must output more bits than JPEG stores (contents of whiteboard)

- 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 $output = input^\gamma$ ($0.0 \leq input \leq 1.0$)

- 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 $\geq \sim 10$ bits; otherwise, dark parts of the ramp will exhibit banding after applying gamma correction and requantizing (integerizing)
Fill factor

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

- 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

anti-aliasing filter removed

normal
Removing the antialiasing filter

✦ “hot rodding” your digital camera
  • $450 + shipping

(maxmax.com)

anti-aliasing filter removed

normal
Cameras without antialiasing filters

Nikon D800

Nikon D800E
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 two 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
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?

✦ James Clerk Maxwell, 1861
  • of Maxwell’s equations
  • 3 images, shot through filters, then simultaneously projected

(wikipedia)
Historical interlude

Q. Who made the first color print?

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

(wikipedia)
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?

- George Albert Smith’s Kinemacolor, 1906
  - alternating red and green filters, total of 32 fps
  - projected through alternating red and green filters
beam splitter leading through 2 filters to two cameras
2 strips of film, cemented together for projection
Technicolor

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

Disney’s Flowers and Trees (1932)

Wizard of Oz (1939)
First consumer color film?

- Kodachrome, 1935
  - no longer available

(wikipedia)

Picadilly Circus, 1949
First color television broadcast?

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

(Beatles in 1964 was in B&W)

started broadcasting in color in 1965
First color television broadcast?

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

(Beatles in 1964 was in B&W)

started broadcasting in color in 1965
Color sensing technologies

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

- 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

(Wikipedia)

(Theuwissen)
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

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

Bayer pattern
Sony RGB+E better color
Kodak RGB+C more dynamic range
Example of Bayer mosaic image

Small fan at Stanford women’s soccer game
(Canon 1D III)
Example of Bayer mosaic image
Before demosaicing (\texttt{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 blue)
  • color filter array (e.g. Bayer) trades off spatial resolution

Questions?
Slide credits

- Brian Curless
- Eddy Talvala
- Abbas El Gamal