Computational Photography:

Camera Processing Pipeline
Scientist’s view of photography

Photo by Uwe Hermann
Scientist’s view of photography
Camera = light-measuring device

Simple models assume an image is a “quantitative measurement” of scene radiance.

Figure from Digital Image Processing, Gonzales/Woods
Image = radiant-energy measurement

Simple models assume an image is a “quantitative measurement” of scene radiance.
Assumption used in many places

- Shape from shading
- HDR Imaging
- Image Matching
- Color constancy
- Etc...
Camera = light-measuring device?
Light-measuring device?

Samsung Galaxy S6 edge

HTC One M9

LG G4

Google Camera App
All settings the same
Onboard processing (photo finishing) “Secret recipe” of a camera

Photographs taken from three different cameras with the same aperture, shutter speed, white-balance, ISO, and picture style.
Modern photography pipeline

Starting point: reality (in radiance)

Pre-Camera
- Lens Filter
- Lens
- Shutter
- Aperture

In-Camera
- CCD response (RAW)
- CCD Demosaicing (RAW)

“Photo-finishing Processing”

Post-Processing
- Touch-up
- Hist equalization
- Spatial warping
- Etc . . .

Ending point: better than reality (in RGB)

Even if we stopped here, the original CCD response potentially has had many levels of processing.
Digital cameras

- Digital cameras are far from being light-measuring devices
- They are designed to produce visually pleasing photographs
- There is a great deal of processing (photo-finishing) happening on the camera

The goal of this lecture is to discuss common processing steps that take place onboard consumer cameras
This lecture will examine

Starting point: reality (in radiance)

Pre-Camera
- Lens Filter
- Lens
- Shutter
- Aperture

In-Camera
- CCD response (RAW)
- CCD Demosaicing (RAW)
- “Photo-finishing Processing”

What is (potentially) happening on a camera.*

Camera Output: sRGB

*Camera pipelines are almost always proprietary, so knowing exactly what steps are performed for a particular make/model is not possible. This lecture examines the most common steps likely to be found on most cameras.
But, first, a “Crash Course” on Color & Color Spaces
Def *Color* (noun): The property possessed by an object of producing different sensations on the eye as a result of the way it reflects or emits light.

*Oxford Dictionary*
Color is perceptual

- **Color is not** a primary physical property
- Red, Green, Blue, Pink, Orange, Atomic Tangerine, Baby Pink, etc. . .
  - Adjectives we assign to “color sensations”

Which is “True Blue”?  

Subjective terms to describe color

**Hue**
Name of the color (yellow, red, blue, green, ...)

**Value/Lightness/Brightness**
How light or dark a color is.

**Saturation/Chroma/Color Purity**
How “strong” or “pure” a color is.

Image from Benjamin Salley
A page from a Munsell Student Color Set
Where do “color sensations” come from?

A very small range of electromagnetic radiation

Generally wavelengths from 390 to 700nm are visible to most individuals.
"White light" through a prism

Light is separated into “monochromatic” light at different wave lengths.
Sensations?

• Our eye has three receptors (cone cells) that respond to visible light and give the sensation of color.
Cones and rods

- We have additional light sensitive cells called *rods* that are not responsible for color.
- Cones are most concentrated on the fovea.
We rarely see monochromatic light in real world scenes. Instead, objects reflect a wide range of wavelengths. This can be described by a *spectral power distribution* (SPD) shown above. The SPD plot shows the relative amount of each wavelength reflected over the visible spectrum.
SPD relation to color is not unique

- Due to the accumulation effect of the cones, two different SPDs can be perceived as the same color.

Lettuce SPD: stimulating $S=0.2$, $M=0.8$, $L=0.8$

Green ink SPD: stimulating $S=0.2$, $M=0.8$, $L=0.8$

Result in the same color “sensation”.

SPD of “real lettuce”

SPD of ink in a “picture of lettuce”

SPD of ink
Tristimulus color theory

• Even before cone cells were discovered, it was empirically found that only three distinct colors (primaries) could be mixed to produce other colors

• Thomas Young (1803), Hermann von Helmholtz (1852), Hermann Grassman (1853), James Maxwell (1856) all explored the theory of trichromacy for human vision
Tristimulus color theory

Grassman’s Law states that a source color can be matched by a linear combination of three independent “primaries”.

\[
\begin{align*}
\text{Source light 1} & = \text{R1} + \text{G1} + \text{B1} \\
\text{Source light 2} & = \text{R2} + \text{G2} + \text{B2} \\
\text{Source light 3} & = (\text{R1} + \text{R2}) + (\text{G1} + \text{G2}) + (\text{B1} + \text{B2})
\end{align*}
\]

Three primaries and the weights (R1, G1, B1) of each primary needed to match the source light #1 perceived color.

Same three primaries and the weights (R2, G2, B2) of each primary needed to match the source light #2 perceived color.

If we combined source lights 1 & 2 to get a new source light 3, the amount of each primary needed to match the new source light 3 will be the sum of the weights that matched lights 1 & 2.
Radiometry vs. photometry

- **Radiometry**
  - Quantitative measurements of radiant energy
  - Often shown as spectral power distributions (SPD)
  - Measures either light coming from a source (radiance) or light falling on a surface (irradiance)

- **Photometry/ colorimetry**
  - Quantitative measurement of *perceived* radiant energy based on human’s sensitivity to light
  - Perceived in terms of “brightness” (photometry) and color (colorimetry)
Quantifying color

• We still need a way to quantify color & brightness

• SPDs go through a “black box” (human visual system) and are perceived as color

• The only way to quantify the “black box” is to perform a human study
Experiments for photometry

Chromatic source light at a particular wavelength and adjustable radiant power.

Reference bright light with fixed radiant power.

(Alternating between source and reference @ 17Hz)

Alternate between the source light and reference light 17 times per second (17 Hz). A flicker will be noticeable unless the two lights have the same perceived “brightness”.

The viewer adjusts the radiant power of the chromatic light until the flicker disappears (i.e. the lights fuse into a constant color). The amount of radiant power needed for this fusion to happen is recorded.

Repeat this flicker fusion test for each wavelength in the source light. This allows the method to be used to determine the perceived “brightness” of each wavelength.

The “flicker photometry” experiment for photopic sensitivity.
CIE* (1924) Photopic luminosity function

The Luminosity Function (written as $\overline{y}(\lambda)$ or $V(\lambda)$) shows the eye’s sensitivity to radiant energy into luminous energy (or perceived radiant energy) based on human experiments (flicker fusion test).

*International Commission on Illumination (CIE comes from the French name Commission internationale de l'éclairage) was a body established in 1913 as an authority on light, illumination and color . . CIE is still active today -- http://www.cie.co.at
Colorimetry

• Based on tristimulus color theory, colorimetry attempts to quantify all visible colors in terms of a standard set of primaries

$$\text{Target color} = \text{R1} + \text{G1} + \text{B1}$$

Three fixed primaries
CIE RGB color matching

Human subjects matched test colors by add or subtracting three primaries.

Field of view was 2-degrees (where color cones are most concentrated)

“Standard Observer”  
(Willing participant with no eye disease)
For some test colors, no mix of the primaries could give a match! For these cases, the subjects were asked to add primaries to the test color to make the match.

This was treated as a negative value of the primary added to the test color.

“Standard Observer”
(Willing participant with no eye disease)
CIE RGB results

Plots are of the mixing coefficients of each primary needed to produce the corresponding monochromatic light at that wavelength.

Note that these functions have been scaled such that area of each curve is equal.
Negative points, the primary used did not span the full range of perceptual color.
In 1931, the CIE met and approved defining a new canonical basis, termed XYZ that would be derived from Wright-Guild’s CIE RGB data

Properties desired in this conversion:
- White point defined at X=1/3, Y=1/3, Z=1/3
- Y would be the luminosity function (V(λ))
- Quite a bit of freedom in selecting these XYZ basis
- In the end, the adopted transform was:

\[
\begin{bmatrix}
X \\ Y \\ Z
\end{bmatrix} = \begin{bmatrix}
0.4887180 & 0.3106803 & 0.2006017 \\
0.1762044 & 0.8129847 & 0.0108109 \\
0.0000000 & 0.0102048 & 0.9897952
\end{bmatrix}\begin{bmatrix}
R \\ G \\ B
\end{bmatrix}
\]

Nice article see: Fairman et al “How the CIE 1931 Color-Matching Functions Were Derived from Wright–Guild Data”, Color Research & Application, 1997
This shows the mixing coefficients $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$ for the CIE 1931 2-degree standard observer XYZ basis computed from the CIE RGB data. Coefficients are all now positive. Note that the basis XYZ are not physical SPD like in CIE RGB, but linear combinations defined by the matrix on the previous slide.
CIE XYZ 3D plot

3D plot of the CIE XYZ matching functions against the XYZ axis. Note that scaling of this plot is not uniform.
What does it mean?

• **We now have a canonical color space to describe SPDs**

• Given an SPD, \( I(\lambda) \), we can find its mapping into the CIE XYZ space

\[
X = \int_{380}^{780} I(\lambda)\bar{x}(\lambda) \, d\lambda \quad Y = \int_{380}^{780} I(\lambda)\bar{y}(\lambda) \, d\lambda \quad Z = \int_{380}^{780} I(\lambda)\bar{z}(\lambda) \, d\lambda
\]

• Given two SPDs, if their CIE XYZ values are equal, then they are considered the same perceived color, i.e.

\( I_1(\lambda), I_2(\lambda) \rightarrow (X_1, Y_1, Z_1) = (X_2, Y_2, Z_2) \) [perceived as the same color]
Two SPDs

SPD1

\[ X = 0.2841 \]
\[ Y = 0.2989 \]
\[ Z = 0.3254 \]

SPD2

\[ X = 0.2841 \]
\[ Y = 0.2989 \]
\[ Z = 0.3254 \]

From their CIE XYZ mappings, we can determine that these two SPDs will be perceived as the same color (even without needing to see the color!)

Thanks CIE XYZ!

CIE XYZ Values

**SPD1**

\[ X = 0.2841 \]
\[ Y = 0.2989 \]
\[ Z = 0.3254 \]

**SPD2**

\[ X = 0.2841 \]
\[ Y = 0.2989 \]
\[ Z = 0.3254 \]

CIE XYZ gives a way to go from radiometric to colorimetric. Imbedded is also the photometric measurement in the Y value.
What does it mean?

• CIE XYZ space is also considered “device independent” – the XYZ values are not specific to any device

• Devices (e.g. cameras, flatbed, scanners, printers, displays) can find mappings of their device specific values to corresponding CIE XYZ values. This provides a canonical space to match between devices (at least in theory)
Luminance-chromaticity space (CIE xyY)

• CIE XYZ describes a color in terms of linear combination of three primaries (XYZ)
• Sometimes it is useful to discuss color in terms of luminance (perceived brightness) and chromaticity (we can think of as the hue-saturation combined)
• CIE xyY space is used for this purpose
Deriving CIE $xyY$

Project the CIE XYZ values onto the $X=1$, $Y=1$, $Z=1$ plane.

$$x = \frac{X}{X+Y+Z}$$

$$y = \frac{Y}{X+Y+Z}$$
A bit hard to visualize after seeing the CIE XYZ 3D plot, but we do obtain the shape on the previous slide from the projection onto X=1, Y=1, Z=1
A bit hard to visualize after seeing the CIE XYZ 3D plot, but we do obtain the shape on the previous slide from the projection onto X=1, Y=1, Z=1.
This gives us the familiar horseshoe shape of visible colors as 2D plot. Note the axis are $x$ & $y$.

Point “E” represents where $X=Y=Z$ have equal energy ($X=0.33$, $Y=0.33$, $Z=0.33$)

CIE XYZ “white point”

In the 1930s, CIE had a bad habit of over using the variables $X$, $Y$. Note that $x$, $y$ are chromaticity coordinates, $\bar{x}$, $\bar{y}$ (with the bar above) are the matching functions, and $X$, $Y$ are the imaginary SPDs of CIE XYZ.
CIE x-y chromaticity diagram

Moving way from the white point represents more saturation.

Moving along the outside of the diagram gives the hues.
CIE xyY

• Generally when we use CIE xyY, we only look at the (x,y) values on the 2D diagram of the CIE x-y chromaticity chart
• However, the Y value (the same Y from CIE XYZ) represents the perceived brightness of the color
• With values (x,y,Y) we can reconstruct back to XYZ

\[ X = \frac{Y}{y} x \quad Z = \frac{Y}{y} (1 - x - y) \]
Fast forward 80+ years

- CIE 1931 XYZ, CIE 1931 xyY (2-degree standard observer) color spaces have stood the test of time
- Many other studies have followed (most notably - CIE 1965 XYZ 10-degree standard observer), . . .
- But in the literature (and in this lecture) you’ll find CIE 1931 XYZ color space making an appearance often
What is perhaps most amazing?

• 80+ years of CIE XYZ is all down to the experiments by the “standard observers”

• How many standard observers were used? 100, 500, 1000?

A Standard Observer
CIE XYZ is based on 17 Standard Observers

10 by Wright, 7 by Guild

“The Standard Observers”
A caution on CIE x-y chromaticity

From Mark D. Fairchild book: “Color Appearance Models”

“The use of chromaticity diagrams should be avoided in most circumstances, particularly when the phenomena being investigated are highly dependent on the three-dimensional nature of color. For example, the display and comparison of the color gamuts of imaging devices in chromaticity diagrams is misleading to the point of being almost completely erroneous.”

Re-quoted from the Mark Meyer’s “Deconstructing Chromaticity” article.
We are done with color, right?
Almost . . .
An object’s SPD

- In a real scene, an object’s SPD is a combination of its reflectance properties and scene illumination.

Our earlier example ignored illumination (we could assume it was pure white light).

Instead, think of how the object reflects different wavelengths.

Tomato SPD

Wavelength ($\lambda$)
Color constancy

• Our visual system is able to compensate for the illumination

Looks the same!
Color constancy/chromatic adaptation

• Color constancy, or chromatic adaptation, is the ability of the human visual system to adapt to scene illumination
• This ability is not perfect, but it works fairly well
• Image sensors do not have this ability (it must be performed as a processing step, i.e. “white balance”)
• **Note:** Our eyes do not adjust to the lighting in the photograph -- we adjust to the viewing conditions of the scene we are viewing the photograph in!
Color constancy and illuminants

• To understand color constancy, we have to consider SPDs of different illuminants
Color temperature

- Illuminants are often described by their color temperature
- This is associated with theoretical “blackbody radiators” that produce SPDs based on a given temperature (expressed in kelvins)
- We often map light sources (both real and synthetic) to their closest color temperature (esp in Photography/Video production)

Freeware app by, **Fu-Kwun Hwang** to generate blackbody SPDs at different temperatures, showing the range in the visible spectrum.
Color temperature

**Kelvin Color Temperature Scale**

- 10,000K: Blue Sky
- 7,000K-7,500K: Cool White Seesmart LED
- 6,000K: Cloudy Sky
- 5,500K-6,000K: Day White Seesmart LED
- 4,800K: Direct Sunlight
- 4,000K-4,500K: Natural White Seesmart LED
- 4,000K: Clear Metal Halide
- 3,000K: 100W Halogen
- 2,800K: 100W Incandescent
- 2,700K-3,200K: Warm White Seesmart LED
- 2,200K: High Pressure
- 1,900K: Candle

Typical description of color temperature used in photography & lighting sources.
Plotted in CIE x-y chromaticity

Plot of color CIE xy locations of SPDs based on color temperature.
Man made illuminants SPDs

Figure from Ponce and Forsyth
CIE standard illuminants

- CIE established several “synthetic” SPDs that serve as proxies for common real illuminants
- Illuminant A
  - tungsten-filament lighting (i.e. a standard light-bulb)
- Illuminant B
  - noon sunlight
- Illuminant C
  - average daylight
- Illuminant D series
  - represent natural daylight at various color temps (5000K, 5500K, 6500K), generally denoted as D50, D55, D65
- Illuminant E
  - idea equal-energy illuminant with constant SPD
  - does not represent any real light source, but similar to D55
- Illuminant F series
  - emulates a variety of fluorescents lamps (12 in total)
CIE standard illuminants

SPDs for CIE standard illuminant A, B, C

SPDs for CIE standard illuminant D50, D55, D65

SPDs for CIE standard illuminant E

SPDs for CIE standard illuminants F2, F8, F11

D, E, and F series images from http://www.image-engineering.de
White point

• A white point is a CIE XYZ or CIE xyY value of an ideal “white target” or “white reference”

• This is essentially an illuminants SPD in terms of CIE XYZ/CIE xyY
  – We can assume the white reference is reflecting the illuminant

• The idea of chromatic adaptation is to make white points the same between scenes
White points in CIE x-y chromaticity

CIE Illuminants
A, B, C, D65, E in terms of CIE x-y

<table>
<thead>
<tr>
<th>CIE</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.44757</td>
<td>0.40745</td>
</tr>
<tr>
<td>B</td>
<td>0.34842</td>
<td>0.35161</td>
</tr>
<tr>
<td>C</td>
<td>0.31006</td>
<td>0.31616</td>
</tr>
<tr>
<td>D65</td>
<td>0.31271</td>
<td>0.32902</td>
</tr>
<tr>
<td>E</td>
<td>0.33333</td>
<td>0.33333</td>
</tr>
</tbody>
</table>
Color constancy (at its simplest)

• (Johannes) *Von Kries* transform
• Compensate for each channel corresponding to the L, M, S cone response

\[
\begin{bmatrix}
L_2 \\
M_2 \\
S_2
\end{bmatrix} = \begin{bmatrix}
1/L_{1w} & 0 & 0 \\
0 & 1/M_{1w} & 0 \\
0 & 0 & 1/S_{1w}
\end{bmatrix}\begin{bmatrix}
L_1 \\
M_1 \\
S_1
\end{bmatrix}
\]

$L_2, M_2, S_2$ is the new LMS response with the illuminant divided “out”. In this case white is equal to [1,1,1]  
$L_{1w}, M_{1w}, S_{1w}$ is the LMS response to “white” under this illuminant  
$L_1, M_1, S_1$ are the input LMS space under an illuminant.
Illuminant to illuminant mapping

- More appropriate would be to map to another illuminant’s LMS response (e.g. in the desired viewing condition)
- \((\text{LMS})_1\) under an illuminant with white-response \((L_{1w}, M_{1w}, S_{1w})\)
- \((\text{LMS})_2\) under an illuminant with white-response \((L_{2w}, M_{2w}, S_{2w})\)

\[
\begin{bmatrix}
L_2 \\
M_2 \\
S_2
\end{bmatrix} = \begin{bmatrix}
L_{2w}/L_{1w} & 0 & 0 \\
0 & M_{2w}/M_{1w} & 0 \\
0 & 0 & S_{2w}/S_{1w}
\end{bmatrix} \begin{bmatrix}
L_1 \\
M_1 \\
S_1
\end{bmatrix}
\]

- \(L_2, M_2, S_2\) is the new LMS response with the illuminant divided “out” and scaled to LMS\(_2\) illuminant
- \(L_{1w}, M_{1w}, S_{1w}\) is the LMS response to “white” the input illuminant, \(L_{2w}, M_{2w}, S_{2w}\) response to “white” of output illuminant
- \(L_1, M_1, S_1\) are the input LMS space under an illuminant.
Example

*Simulation* of different “white points” by photographing a “white” object under different illumination.

Images courtesy of Sharon Albert (Weizmann Institute)
Here, we have mapped the two input images to one below to mimic chromatic adaptation. The “white” part of the cup is shown before and after to help show that the illumination falling on white appears similar after the “chromatic adaptation”.
Now we are finally done with color?
Almost (really) . . .
CIE XYZ and RGB

• While CIE XYZ is a canonical color space, images/devices rarely work directly with XYZ
• XYZ are not real primaries
• RGB primaries dominate the industry
• We are all familiar with the RGB color cube

But by now, you should realize that Red, Green, Blue have no quantitative meaning. We need to know their corresponding SPDs or CIE XYZ values
Device specific RGB values

The RGB values span a subspace of CIE-XYZ to define the devices gamut.

If you have RGB values, they are specific to a particular device.
Trouble with RGB

RGB values have no meaning if the primaries between devices are not the same! This is a huge problem for color reproduction from one device to the next.
In 1996, Microsoft and HP defined a set of “standard” RGB primaries.

R = CIE xyY (0.64, 0.33, 0.2126)
G = CIE xyY (0.30, 0.60, 0.7153)
B = CIE xyY (0.15, 0.06, 0.0721)

This was considered an RGB space achievable by most devices at the time.

White point was set to the D65 illuminant. **This is an important thing to note.** It means sRGB has built in the assumed viewing condition (6500K daylight).
CIE XYZ to sRGB conversion

Matrix conversion:

\[
\begin{bmatrix}
R \\
G \\
B \\
\end{bmatrix} = \begin{bmatrix}
3.2404542 & -1.5371385 & 0.4985314 \\
-0.9692660 & 1.8760108 & 0.0415560 \\
0.0556434 & -0.2040259 & 1.0572252 \\
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
\end{bmatrix}
\]

- R=G=B=1 is defined as illuminant D65 is in CIE XYZ
- This is the linear-sRGB space
- sRGB also specifies a gamma correction of the values
Actual formula is a bit complicated, but effectively this is gamma \( (I' = 255*I^{(1/2.2)}) \), where \( I' \) is the output intensity and \( I \) is the linear sRGB ranged 0-1, with a small linear transfer for linearized sRGB values close to 0 (not shown in this plot).
Gamma justification

- Gamma encoding is used in hardware to compensate for the non-linear characteristics of cathode ray tubes (CRT).
- The application of the gamma=1/(2.2) will be undone by the CRT’s non-linear responsible to voltage to beam intensity to produce the desired result.
Before (linear sRGB) & after (sRGB)
Standardization isn’t new - NTSC/PAL
CIE XYZ ↔ NTSC/sRGB
(know your color space!)

It is important to know which color space your image is in.

Linear-sRGB back to XYZ

\[
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} = \begin{bmatrix}
0.4124 & 0.3576 & 0.1805 \\
0.2126 & 0.7152 & 0.0722 \\
0.0193 & 0.1192 & 0.9505
\end{bmatrix} \begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

Linear-NTSC back to XYZ

\[
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} = \begin{bmatrix}
0.6071 & 0.1736 & 0.1995 \\
0.2990 & 0.5870 & 0.1140 \\
0.0000 & 0.0661 & 1.1115
\end{bmatrix} \begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]
CIE XYZ: The mother of color spaces

ProPhoto
Adobe RGB
L*ab
sRGB
YUV
HSV
HSI
NTSC
YIQ
CIE XYZ: The mother of color spaces

But, be cautious. All Ys are not equal. The YUV and YIQ are defined from the gamma encoded RGB values. Technically they are not CIE Y luminance, but “luma”. They should be written with a Y’ to distinguish them.
Standard color spaces are great
Benefits of sRGB

• Like CIE XYZ, sRGB is a device independent color space (often called an output color space)
• If you have two pixels with the same sRGB values, they will have the same CIE XYZ value, which means “in theory” they will appear as the same perceived value
• Does this happen in practice?
See for yourself

Canon  Nikon  Sony

sRGB values that are the same will be perceived as the same on the screen. However, from “real scene SPD” to final sRGB, cameras are clearly doing different things.
Congratulations!

Certificate of Completion

This Certificate is presented to

You

on this 15 day of Feb 2017

for completion of

“Crash Course on Color”
Crash course on color is over!

- A lot of information to absorb
- Understanding colorimetry is required to understand imaging devices
- CIE XYZ and CIE illuminants will make many appearances in color imaging/processing discussions
Overview of the Camera Imaging Pipeline
Pipeline for sRGB (JPEG)

Sensor with color filter array (CCD/CMOS) → Gain Control A/D Converter Possible LUT → White Balance

AFE – Analog Front End Sensor related processing

Tone Reproduction ← Color Space Transform + Color Preferences ← Noise Reduction/Sharpening ← CFA Demoasicing

JPEG Compression ← Exif File Info ← Save to storage

* Note that steps can be optional (e.g. noise reduction) or applied in slightly different order.
Pipeline for RAW

Sensor with color filter array (CCD/CMOS)

Gain Control
A/D Converter
Possible LUT

AFE – Analog Front End
Sensor related processing

Lossless Compression

TIFF encoding
(meta data)

Save to storage

Additional steps could be “estimated” but not applied for RAW*. For example, white point/illuminant estimation, which is saved in the meta-data but not applied to the RAW. Some cameras even save the full “Exif” data with RAW so the RAW can be rendered as if it was fully processed as an sRGB/JPEG.

*Technically speaking, RAW shouldn’t be all caps, it is not an acronym.
Imaging sensor

• We will discuss the basics of the imaging sensor at a high level
  – Could have a whole lecture just on sensors
• The assumption is after this step, the digital values coming off the sensor are linear with respect to the amount of falling on the pixel over a given amount of time
• Will discuss some common steps to make this happen
Imaging sensor

Basic imaging sensor design

Microlens to help increase light collection on the sensor

Color filters place over the sensor. This forms a Color Filter Array (CFA) also called a “Bayer Pattern” after inventor Bryce Bayer.

Silicon/Circuitry

Photodiode (Photon comes in, electron pops out)
CCD & CMOS

CCD (charge coupled device) has a different readout technology to convert charge to voltage and buffer it for output. The plus side is there is more space on the pixel for the photo sensor.

CMOS (complementary metal oxide semiconductor) converts charge to voltage at the pixel site. This allows faster readouts, but less space for the photo sensor per pixel.

Figures from Dave Litwiller, “CCD vs. CMOS”, Photonics Spectra, Jan 2001
Camera RGB sensitivity

- The color filter array (CFA) on the camera filters the light into three primaries.
Black light subtraction

• Sensor values for pixels with “no light” should be zero
• But, often, this is not the case for various reasons
  – Cross talk on the sensor, etc.
  – This can also change with sensor temperature
• This can be corrected by capturing a set of pixels that do not see light
• Place a dark-shield around sensor
• Subtract the level from the “black” pixels
Optical black (OB)

Sensor area(s) capturing "optical black"

Black light capturing areas (likely exaggerated) from Sony US Patent US8227734B2 (Filed 2008)
Signal amplification (gain)

• Imaging sensor signal is amplified
• Amplification to assist A/D
  – Need to get the voltage to the range required to the desired digital output
• This gain could also be used to accommodate camera ISO settings
  – Unclear if all cameras do this here or use a simple post-processing gain to the RAW for ISO settings
  – DSLR cameras RAW is modified by the ISO setting
Defective pixel mask

- CCD/CMOS have pixels that are defective
- Dead pixel masks are pre-calibrated at the factory
  - Using “dark current” calibration
  - Take an image with no light
  - Record locations reporting values to make “mask”
- Bad pixels in the mask are interpolated
- This process seems to happen before RAW is saved
  - If you see dead pixels in RAW, these are generally new dead pixels that have appeared after leaving the factory
Example

Identifying “dead pixels”

After interpolation

Image courtesy of Lu Zheng (CityU Hong Kong) and Moshe Ben-Ezra (MIT)
Nonlinear response correction

• Some image sensor (generally CMOS) often have a non-linear response to different amounts of irradiance

• A non-linear adjustment or look up table (LUT) interpolation can be used to correct this
Other possible distortion correction

• Sensor readout could have spatial distortion for various reasons, e.g. sensor cross-talk
• For point-and-shoot/mobile cameras with fixed lens, vignetting correction for lens distortion could be applied
• Such corrections can be applied using a LUT or polynomial function (in the case of vignetting)
Ex: Flat field correction
Ex: Flat field correction
(non-uniform gain)
At this stage

- We now have a reading from the sensor that is linear with respect to light coming in
- Defective pixels have been interpolated
- Potential distortion has been reduced
RAW – stop here

• If saving in RAW, we can stop here.
• Convert to TIFF + metadata, save to media
• RAW generally represents gained + linearized sensor response before CFA demosaicing and white-balance correction
• We like to think of RAW as “unprocessed” sensor response

Important: RAW image color space will be camera-specific
White balance

• White balance is intended to mimic chromatic adaptation of the eye

• Users can manually set the white balance
  – Camera specific white-balance matrices are used selected illuminant (see next slides)
  – This is often stored in the Exif metadata

• Otherwise auto white balance (AWB) is performed
## WB manual settings

<table>
<thead>
<tr>
<th>WB SETTINGS</th>
<th>COLOR TEMPERATURE</th>
<th>LIGHT SOURCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>☁️</td>
<td>10000 - 15000 K</td>
<td>Clear Blue Sky</td>
</tr>
<tr>
<td>☀️</td>
<td>6500 - 8000 K</td>
<td>Cloudy Sky / Shade</td>
</tr>
<tr>
<td>☀️</td>
<td>6000 - 7000 K</td>
<td>Noon Sunlight</td>
</tr>
<tr>
<td>🌅</td>
<td>5500 - 6500 K</td>
<td>Average Daylight</td>
</tr>
<tr>
<td>⚡️</td>
<td>5000 - 5500 K</td>
<td>Electronic Flash</td>
</tr>
<tr>
<td>🌅</td>
<td>4000 - 5000 K</td>
<td>Fluorescent Light</td>
</tr>
<tr>
<td>🕯️</td>
<td>3000 - 4000 K</td>
<td>Early AM / Late PM</td>
</tr>
<tr>
<td>🕯️</td>
<td>2500 - 3000 K</td>
<td>Domestic Lightning</td>
</tr>
<tr>
<td>🕯️</td>
<td>1000 - 2000 K</td>
<td>Candle Flame</td>
</tr>
</tbody>
</table>

Typical mapping of WB icons to related color temperature. White-balance matrix is often stored in the exif file.
Examples of manual WB matrices

**Sunny**

\[
\begin{bmatrix}
2.0273 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 1.3906 \\
\end{bmatrix}
\]

**Nikon D7000**

**Incandescent**

\[
\begin{bmatrix}
1.3047 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 2.2148 \\
\end{bmatrix}
\]

**Shade**

\[
\begin{bmatrix}
2.4922 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 1.1367 \\
\end{bmatrix}
\]

**Daylight**

\[
\begin{bmatrix}
2.0938 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 1.5020 \\
\end{bmatrix}
\]

**Canon 1D**

**Tungsten**

\[
\begin{bmatrix}
1.4511 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 2.3487 \\
\end{bmatrix}
\]

**Shade**

\[
\begin{bmatrix}
2.4628 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 1.2275 \\
\end{bmatrix}
\]

**Daylight**

\[
\begin{bmatrix}
2.6836 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 1.5586 \\
\end{bmatrix}
\]

**Sony A57K**

**Tungsten**

\[
\begin{bmatrix}
1.6523 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 2.7422 \\
\end{bmatrix}
\]

**Shade**

\[
\begin{bmatrix}
3.1953 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 1.2891 \\
\end{bmatrix}
\]
Auto white balance (AWB)

• If manual white balance is not used, then an AWB algorithm is performed
• This is not entirely the same as chromatic adaptation, because it doesn’t have a target illuminant, instead AWB (as the name implies) attempts to make what is assumed to be white map to “pure white”
• Next slides introduce two well known methods: “Gray World” and “White Patch”
AWB: Gray world algorithm

• This method assumes that average reflectance of a scene is achromatic (i.e. gray)
  – Gray is just the white point not at its brightest, so it serves as an estimate of the illuminant
  – This means that image average should have equal energy, i.e. R=G=B

• Based on this assumption, the algorithm adjusts the input average to be gray as follows:

\[
R_{avg} = \frac{1}{N_r} \sum R_{sensor}(r) \quad G_{avg} = \frac{1}{N_g} \sum G_{sensor}(g) \quad B_{avg} = \frac{1}{N_b} \sum B_{sensor}(b)
\]

- r = red pixels values, g = green pixels values, b = blue pixels values
- \(N_r\) = # of red pixels, \(N_g\) = # of green pixels, \(N_b\) = # blue pixels

First, estimate the average response:

Note: # of pixel per channel may be different if white balance is applied to the RAW image before demosaicing. Some pipelines may also transform into another colorspace, e.g. LMS, to perform the white-balance procedure.
AWB: Gray world algorithm

• Based on averages, white balance can be expressed as a matrix as:

\[
\begin{bmatrix}
R' \\
G' \\
B'
\end{bmatrix} = \begin{bmatrix}
G_{avg}/R_{avg} & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & G_{avg}/B_{avg}
\end{bmatrix}\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

Matrix scales each channel by its average and then normalizes to the green channel average.

Note: some (perhaps most) pipelines may also transform into another colorspace, e.g. LMS, to perform the white-balance procedure.
AWB: White patch algorithm

• This method assumes that highlights represent specular reflections of the illuminant
  – This means that maximum R, G, B values are a good estimate of the white point

• Based on this assumption, the algorithm works as follows:

\[ R_{max} = \max(R_{sensor}(r)) \quad G_{max} = \max(G_{sensor}(g)) \quad B_{max} = \max(B_{sensor}(b)) \]

\[ r = \text{red pixels values, } g = \text{green pixels values, } b = \text{blue pixels values} \]
AWB: White patch algorithm

• Based on RGB max, white balance can be expressed as a matrix as:

\[
\begin{bmatrix}
R' \\
G' \\
B'
\end{bmatrix} = \begin{bmatrix}
G_{\text{max}}/R_{\text{max}} & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & G_{\text{max}}/B_{\text{max}}
\end{bmatrix}\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

Matrix scales each channel by its maximum value and then normalizes to the green channel’s maximum.
AWB example

Input

Gray World

White Patch
Better AWB methods

• Gray world and white patch are *very basic* algorithms
  – These both tend to fail when the image is dominated by large regions of a single color (e.g. a sky image)

• **There are many improved versions**

• Most improvements focus on how to perform these white point estimation more robustly

• **Note**: AWB matrix values are often stored in an images Exif file
CFA demosaicing

• Color filter array/Bayer pattern placed over pixel sensors
• We want an RGB value at each pixel, so we need to perform interpolation

Sensor RGB layout

Desired output with RGB per pixel.
Simple interpolation

Simply interpolate based on neighbor values.
Simple “edge aware” interpolation

If \(|G2 - G8| \&\& |(G4 - G8)| \text{ both } < \text{ Thres}\)

\[ G5 = \frac{G2 + G4 + G6 + G8}{4} \]

elseif \(|G2 - G8| > \text{ Thres}\)

\[ G5 = \frac{G4 + G6}{2} \]

else

\[ G5 = \frac{G2 + G8}{2} \]

Case 1

All about the same.

Case 2

G2 and G8 differ

G4 and G6 differ
Demosaicing

• These examples are simple algorithms
• Cameras almost certainly use more complex and proprietary algorithms
• Demosaicing can be combined with additional processing
  – Highlight clipping
  – Sharpening
  – Noise reduction
Noise Reduction (NR)

- All sensors inherently have noise
- Some of this is mitigated in the sensor circuitry (discussed previously)
- Some cameras apply NR after A/D conversion
- A couple simple methods are presented here
- For high-end cameras, it is likely that cameras apply slightly different strategies depending on ISO settings, e.g. high ISO will result in more noise, so a more aggressive NR could be used
- Examples given are more likely on lower-end point-and-shoot cameras
NR – Rank order statistics

• Sliding window **median** filter
• Sort all pixels in a 3x3 (or larger) widow about center pixel by value
• Select median (i.e. pixel #5 in rank)

![Input and Output](image)

Median filter is nice because it preserves edges.
NR – Rank order statistics

• Sliding window **despeckle** filter
• Sort all pixels in a 3x3 (or larger) widow about center pixel by value
• If center pixel maps to extreme (rank 1 or 9) and is significantly different than closest neighbor, take neighboring ranked pixel value

Center pixel mapped to an extreme. 
|Rank 9 (19) – Rank 8 (11)| > threshold
Replace with Rank 8
NR + Sharpening

- Another strategy is to apply a blur and add the detail back in “content regions”, or even boost content to perform sharpening.

Sketch of the procedure here:

For values with high-response assume this is “content” and add it back to the blur version (can even give a slight boost = sharpening). If low response, don’t do anything, then this region will have NR applied.
Color space manipulation

• We are now at one of the most crucial steps in the pipeline
• Up until this point, the “RGB” values have been related to the camera’s RGB sensitivity, i.e. sensor RGB color space
• We now need to map these values to the output color space (sRGB)
Color space transform

• Color Correction Matrix (CCM)
  – Transforms sensor native RGB values into some canonical colorspace (e.g. CIE XYZ) that will eventual be transform to the final sRGB colorspace

• It is important that the white-balance has been performed correctly
Color transform conceptually

These LUT used to correct sensor RAW

RAW sensor RGB values come in, and sRGB values come out

Tone-scale and sRGB gamma are applied

Matrix decompositions

Focusing on the matrix part only. Conceptually we could breakdown the color transform into a series of transforms. It may desirable to perform AWB in a different colorspace, e.g. LMS, then to sRGB (using CIE XYZ as an intermediate colorspace to try things together).

First, convert sensor RGB to CIE (1931) XYZ

This is from Parulski & Spaulding, however, it is more common to perform white-balance first – more on this later.
Tone-mapping

• Non-linear mapping of RGB tones
• Applied to achieve some preferred tone-reproduction
  – This is not sRGB gamma
  – This is to make the images look nice
• To some degree this mimics the nonlinearity in film (known as the Hurter-Driffield Curves)
• Each camera has its own unique tone-mapping (possibly multiple ones)
Examples

(RAW or linearize sRGB)

Here are several examples of tone-mappings for a white range of cameras. The x-axis is the input “RAW” (or linearized-sRGB), the y-axis is the output After tone mapping.

From Grossberg & Nayar (CVPR'03)
Tone-mapped examples
Note about tone-mapping

• It is worth noting, that up until this stage, our color values (either RAW or linear sRGB) are related to incoming light in a linear fashion
• After this step, that relationship is broken
• Unlike the sRGB gamma (which is known), the tone-mapping is propriety and can only be found by a calibration procedure
JPEG compression

• Joint Photographic Experts Group (JPEG)
• Lossy compression strategy based on the 2D Discrete Cosine Transformation (DCT)
• The by far the most widely adopted standard for image storage
JPEG applies almost every compression trick known.
1) Transform coding, 2) psychovisual (loss), 3) Run-length-encoding (RLE), 4) Difference coding, and Huffman.
JPEG quality

• The amount of quantization applied on the DCT coefficients amounts to a “quality” factor
  – More quantization = better compression (smaller file size)
  – More quantization = lower quality

• Cameras generally allow a range that you can select
Exif metadata

• Exchangeable image file format (Exif)
• Created by the Japan Electronics and Information Technology Industries Association (JEITA)
• Associates meta data with images
  – Date/time
  – Camera settings (basic)
    • Image size, aperture, shutter speed, focal length, ISO speed, metering mode (how exposure was estimated)
  – Additional info (from in some Exif files)
    • White-balance settings, even matrix coefficients of white-balance
    • Picture style (e.g. landscape, vivid, standard, portrait)
    • Output color space (e.g. sRGB, Adobe RGB, RAW)
    • GPS info
    • More . . .
Saved to storage. We are done!

Sensor with color filter array (CCD/CMOS)

Gain Control
A/D Converter
Possible LUT

AFE – Analog Front End
Sensor related processing

White Balance

Tone Reproduction

Color Space Transform + Color Preferences

Noise Reduction/Sharpening

CFA Demoasicing

JPEG Compression

Exif File Info

Save to storage

* Note that steps can be optional (e.g. noise reduction) or applied in slightly different order.
Pipeline comments

• Again, important to stress that the exact steps mentioned in these notes only serve as a guide of common steps that take place

• For different camera makes/models, these could be performed in different order (e.g. white-balance after demosaicing) and in different ways (e.g. combining sharpening with demosaicing)
These two steps combined is often referred to as “color rendering” or “color processing/interpolation”. What actually happens here is very much specific to individual cameras, and even settings on the camera.
ICC and color profiles

• International Color Consortium (ICC)
  – In charge of developing several ISO standards for color management
• Promote the use of ICC profiles
• ICC profiles are intended for device manufacturers to describe how their respective color spaces (e.g. sensor RGB) map to canonical color spaces called Profile Connection Spaces (PCS)
• PCS are similar to linking all devices to CIE XYZ, but are more flexible allowing for additional spaces to be defined (beyond CIE XYZ)
From the ICC – ISO 22028

Photography and graphic technology — Extended colour encodings for digital image storage, manipulation and interchange —

In this case, “sensor characterization” is related to the Color Adaptation Matrix described earlier.

We can see tone-mapping is not explicitly denoted, instead it is grouped with “color render”.

In-camera processing
- Exposure adjustment
- White balancing
- Sensor characterization
- Colour rendering

Output-referred CRT-ready colour encoding

Printer colour transform

Softcopy image

Hardcopy image
From the ICC ISO 22028

This describes a basic digital camera pipeline in more detail.

RGB values linked to the device are considered “scene referred”.

After the color transform to sRGB they are denoted as “output referred” color encodings.

We can see a more detailed breakdown of color rendering.
Tone mapping/color rendering

• From the ICC ISO standards, we see this part is more complex than just matrix + tone-map
• In fact, it can often involve 3D LUTs
Camera image processors

• Steps applied after the sensor output are generally performed by an “image processor” on the camera

• Different cameras use different processing boards/software

• High-end cameras and associated processors
  – Nikon – Expeed
  – Fuji – Real photo engine
  – Canon – DIGIC
  – Sony – BIONZ
  – . . .

• Mobile-devices
  – Qualcomm - Snapdragon
  – Nvidia - Tegra
Example: Expeed block diagram

Example: Sony Exmor + BIONZ

Not a lot of detail . . .

From Sony’s website
Concluding remarks

• Your camera performs a great deal of processing

• RAW
  – Represents relatively unprocessed sensor response
  – Does not represent CIE XYZ
  – Transform to CIE XYZ is not readily available
  – There is no standard RAW!

• sRGB
  – Is defined with respect to CIE XYZ
  – However, incoming radiant energy (SPDs) to sRGB is not standardize (otherwise all camera images would look the same)
Understanding Color Perception

Hermann Grassmann

Johannes von Kries

W. David Wright

James Clerk Maxwell

Hermann von Helmholtz

Thomas Young
Digital Cameras

Bryce E. Bayer

Willard Boyle and George Smith
(Nobel Prize for inventing the CCD)
Photo: Reuters

Eric R. Fossum
(Invented CMOS)
And of course

“The Standard Observers”