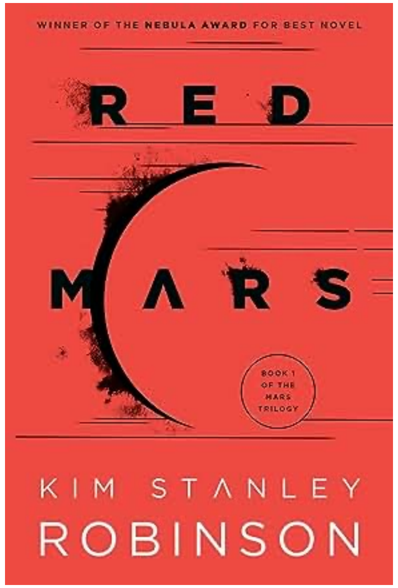
Modern Photography Pipeline

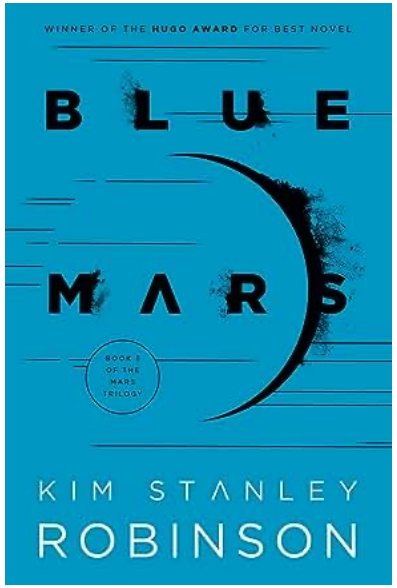
Computational Photography (CSCI 3240U)

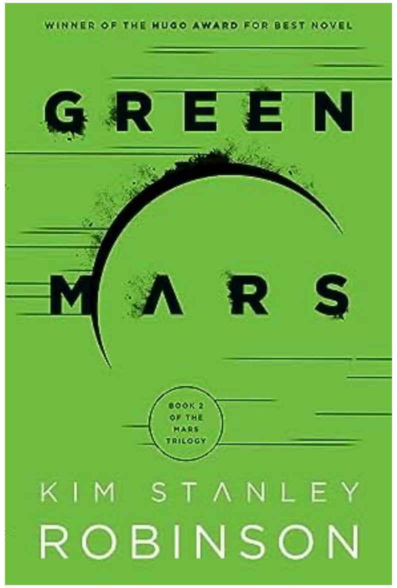
Faisal Z. Qureshi
http://vclab.science.ontariotechu.ca
Ontario Tech University

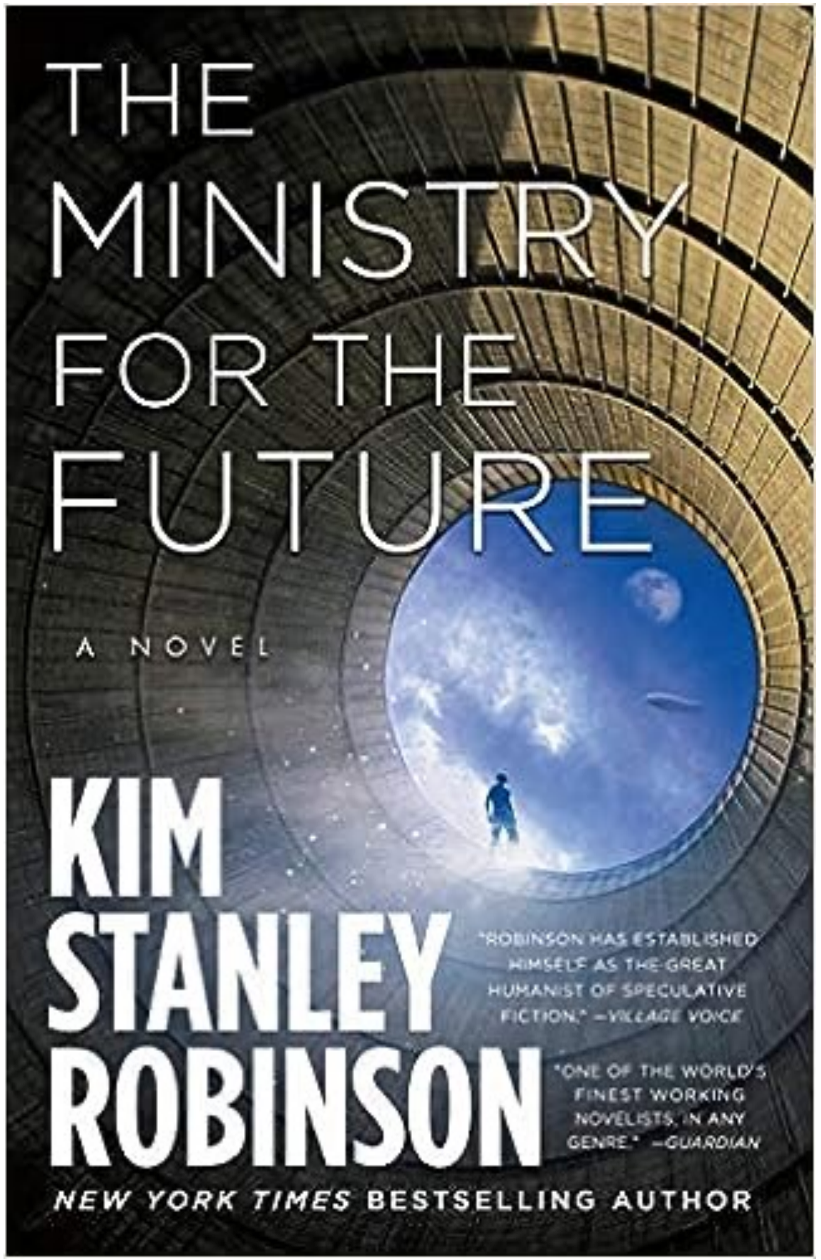
Week 1 – Lecture 2

Fun readings









Acknowledgments

- These slides borrow and adapt materials developed by others, including
  - Michael Brown
  - Kyros Kutulakos
  - David Lindell
  - Gordon Wetzstein
  - Marc Levoy
  - Fredo Durand
  - Paul Debevec
  - Ramesh Raskar

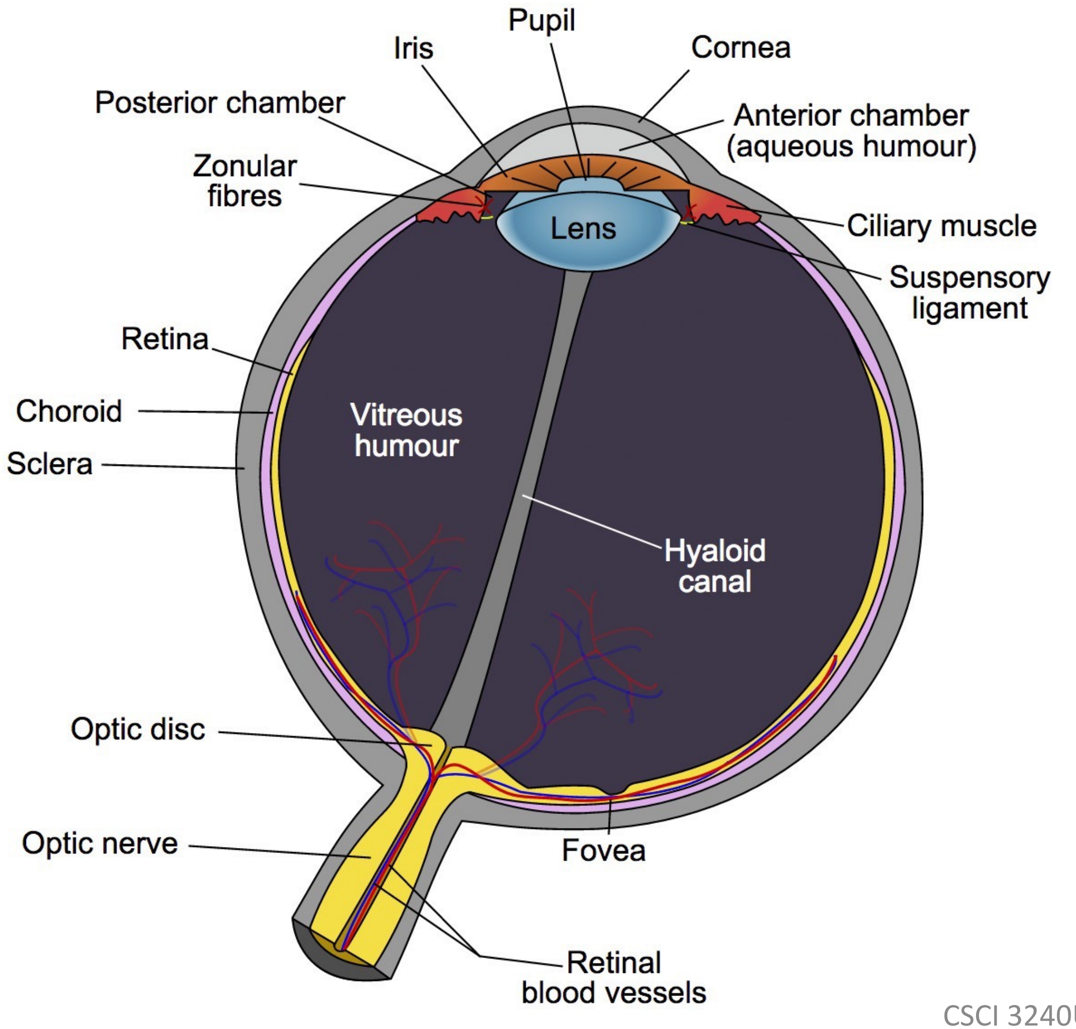
Slide credits

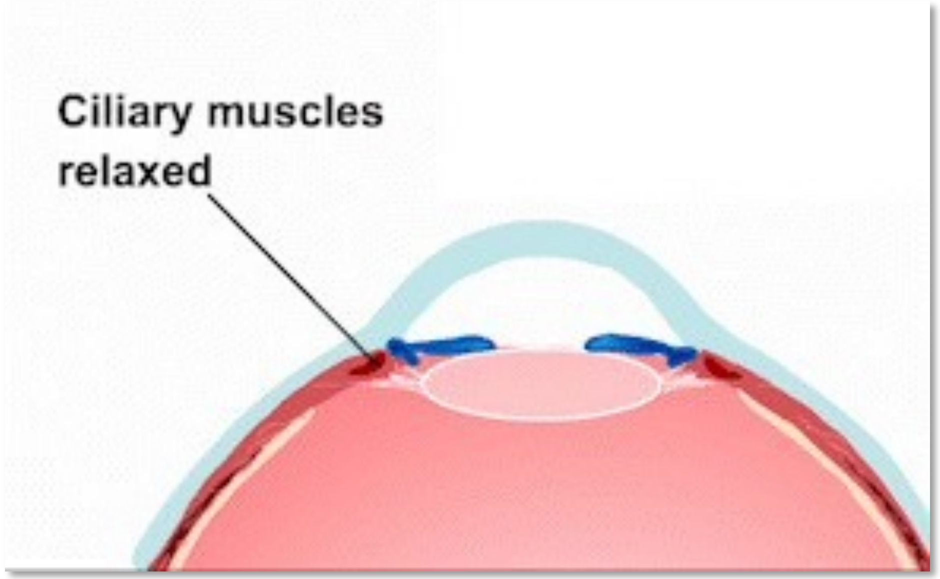
- A lot of inspiration and quite a few examples for these slides were taken directly from:
  - Kayvon Fatahalian (15-769, Fall 2016).
  - Michael Brown (CVPR 2016 Tutorial on understanding the image processing pipeline).
  - Marc Levoy (Stanford CS 178, Spring 2014).

Special thanks to Ioannis Gkioulekas

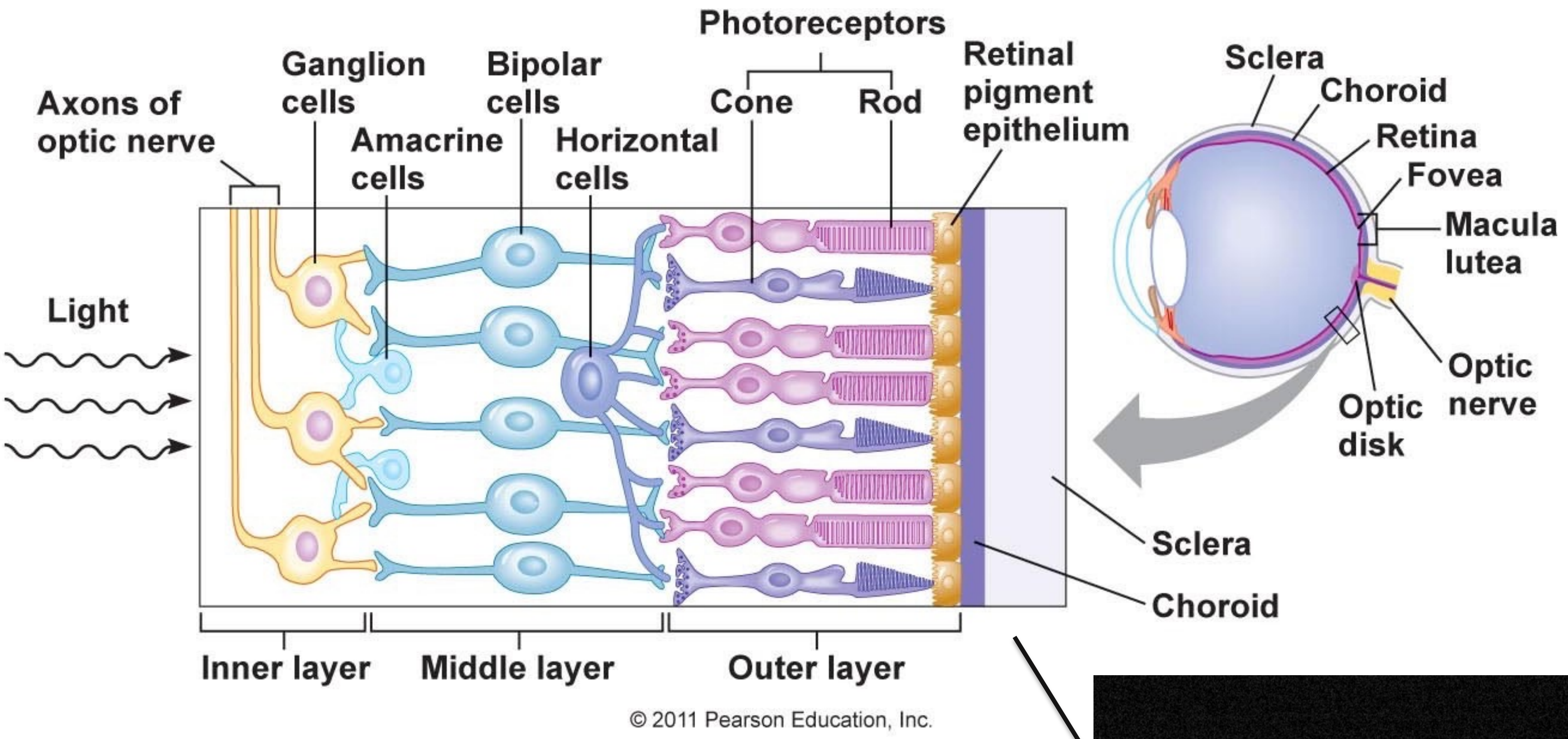
- Many of the slides are taken with his permission from the computational photography course that he has developed at CMU

Anatomy of the Human Eye

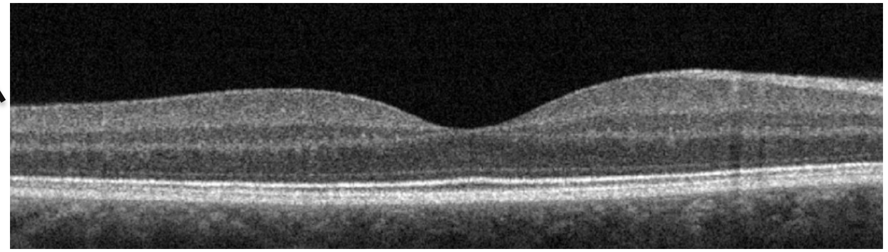




Retina

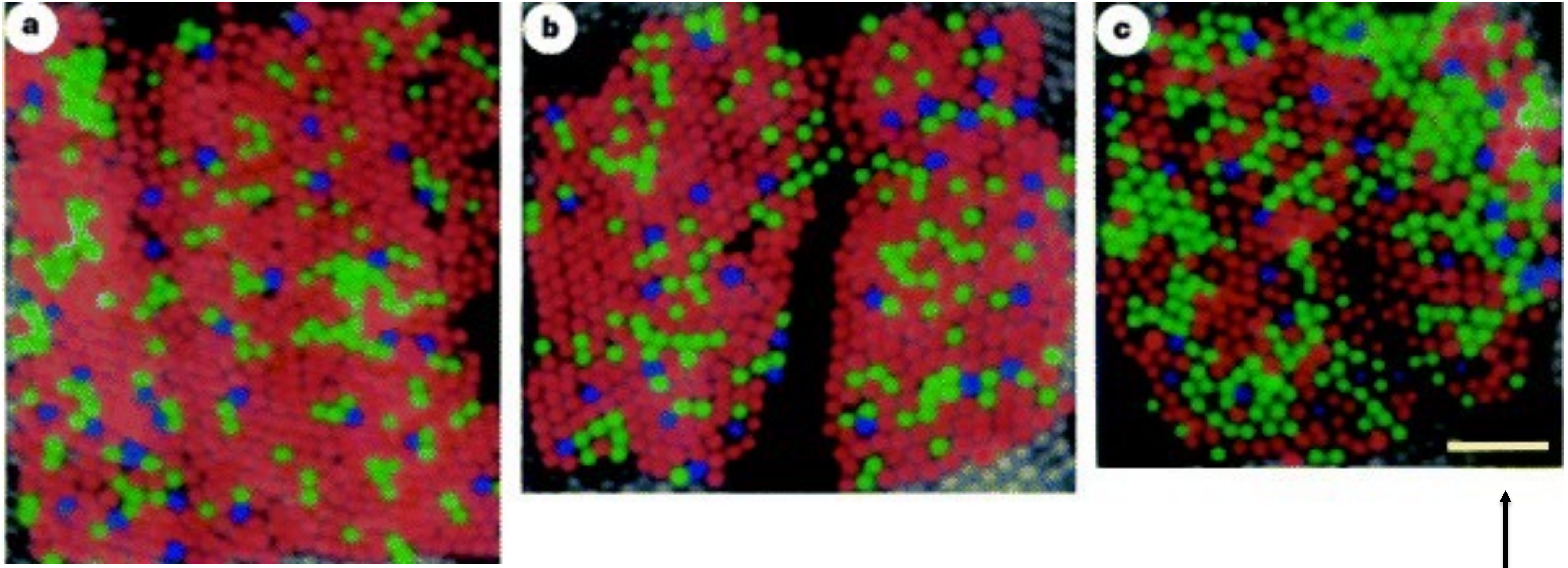


Photoreceptors
Axons of optic nerve
Ganglion cells
Amacrine cells
Bipolar cells
Horizontal cells
Cone
Rod
Retinal pigment epithelium
Light
Inner layer
Middle layer
Outer layer
Sclera
Choroid
Retina
Fovea
Macula lutea
Optic nerve
Optic disk
Sclera
Choroid
© 2011 Pearson Education, Inc.



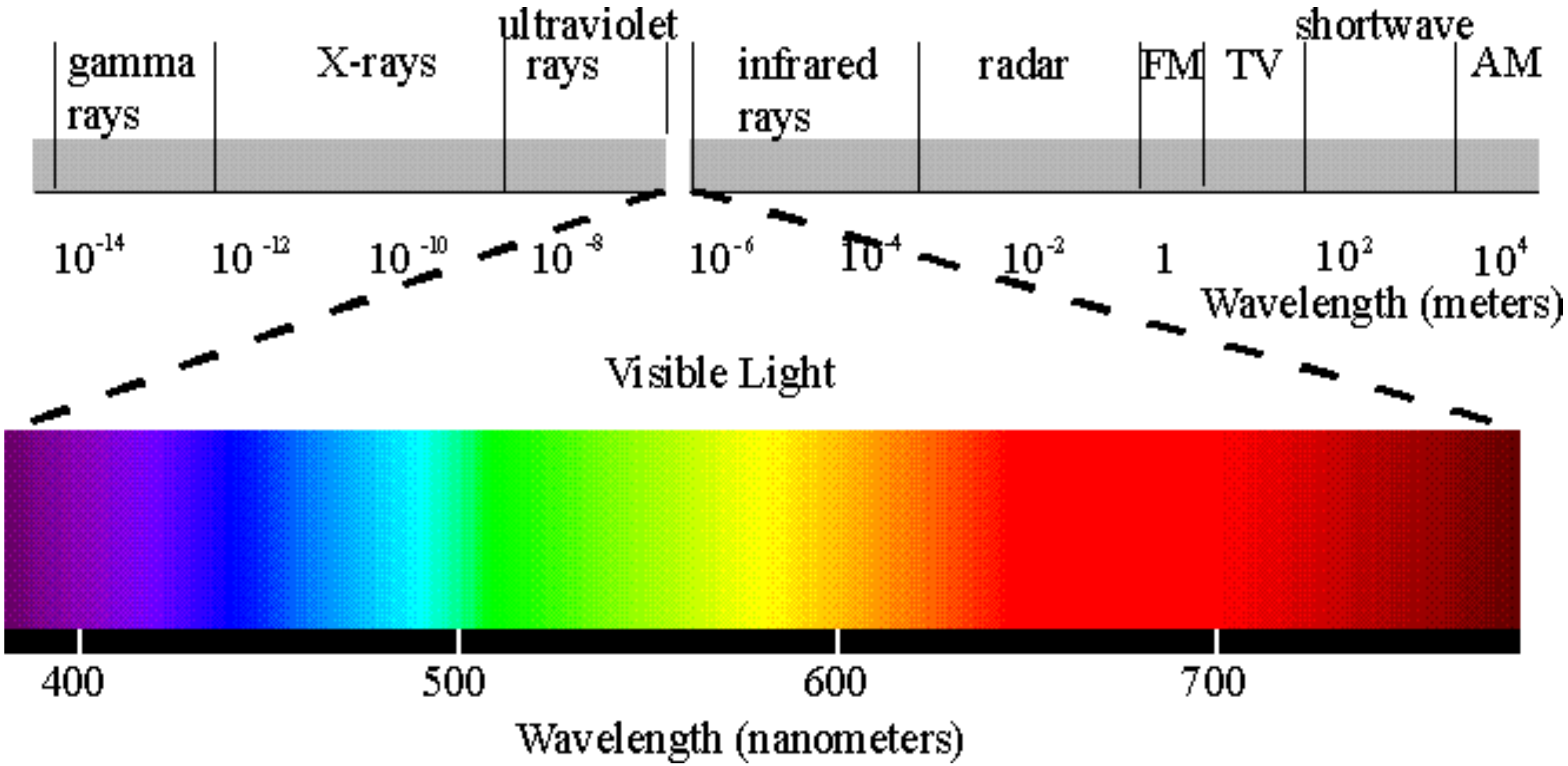
Retina

Roorda & Williams, 1999, Nature



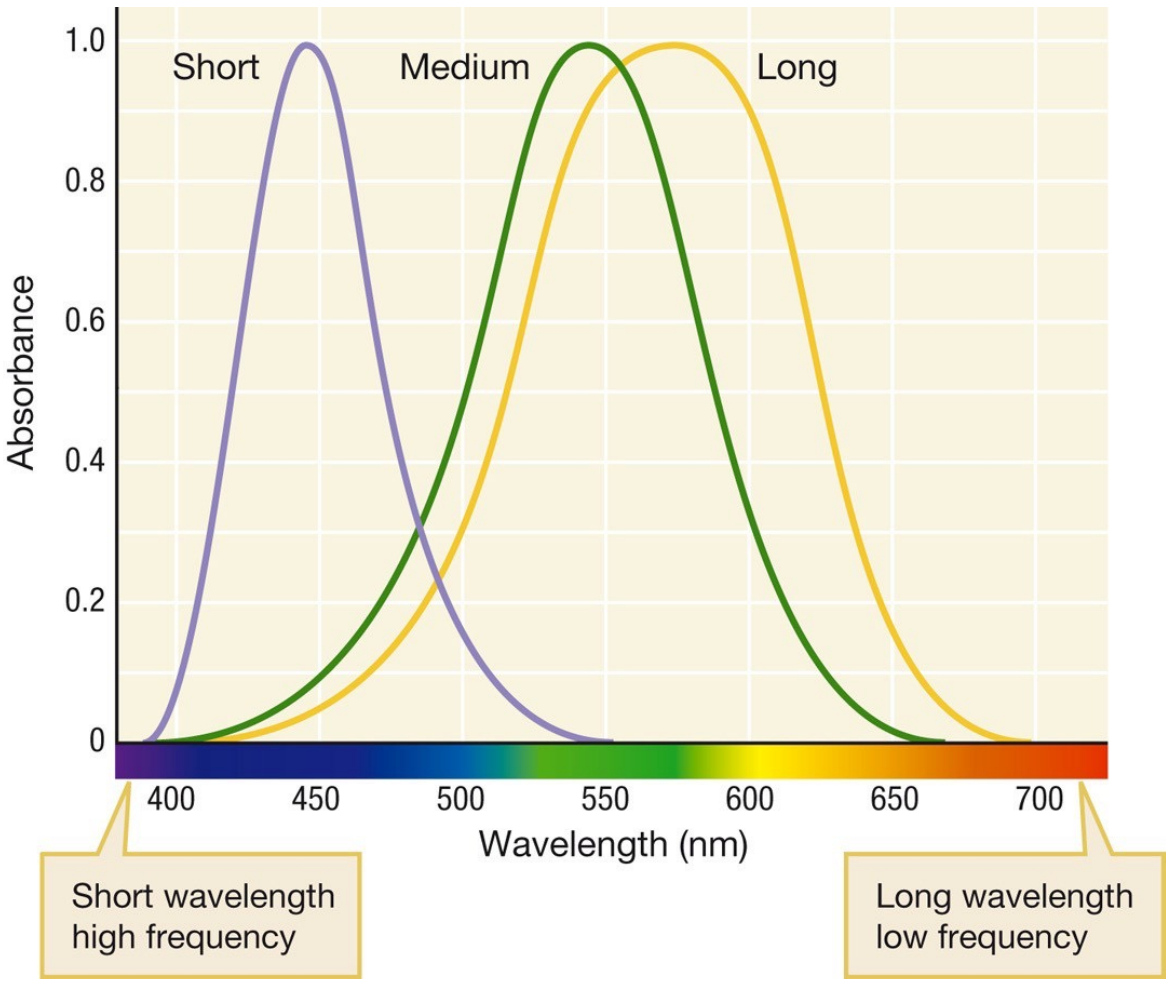
5 arcmin visual angle

Color perception

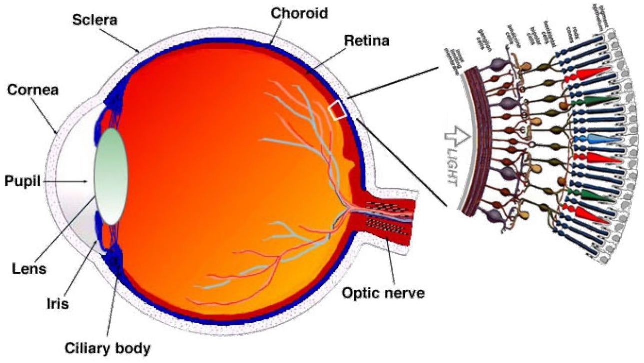


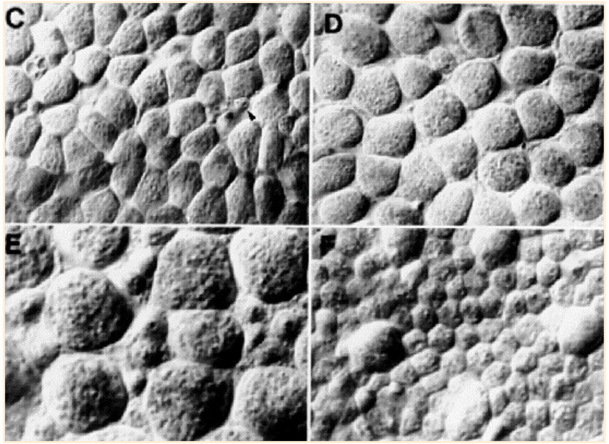
Sensitivity of Cones

reddit.com

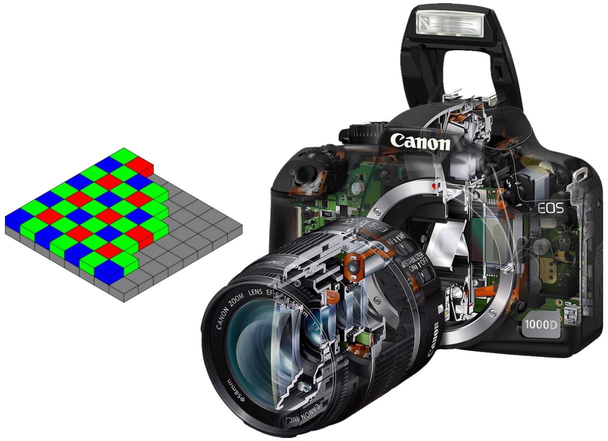


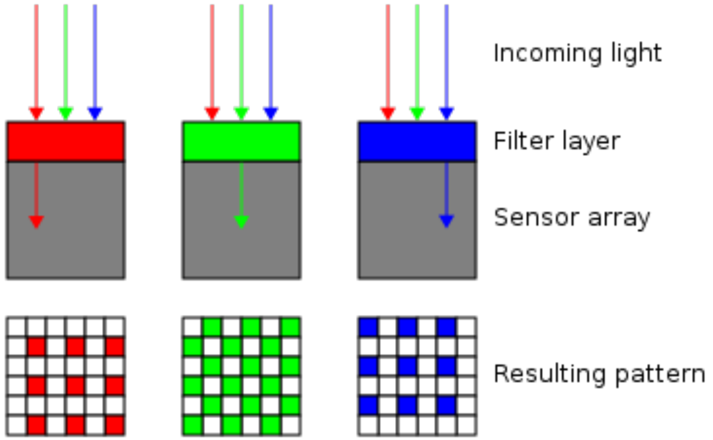
Eye vs. camera





[Williams 91]

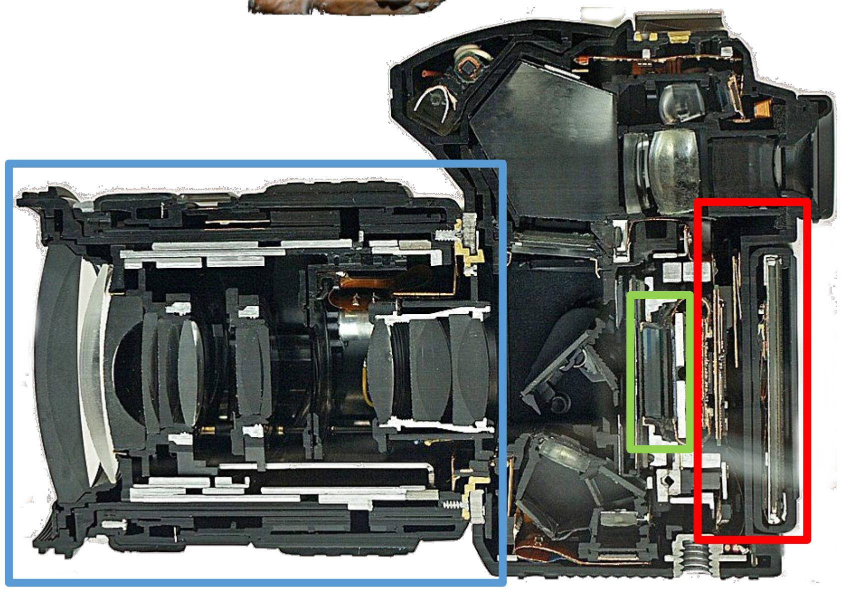




wikipedia

The modern photography pipeline

The modern photography pipeline



Optics and optical controls → Sensor, analog circuitry, and color filter array → In-camera image processing pipeline

On a Heuristic Viewpoint Concerning the Production and Transformation of Light

132

6. *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt;*
von A. Einstein.

Zwischen den theoretischen Vorstellungen, welche sich die Physiker über die Gase und andere ponderable Körper gebildet haben, und der Maxwellschen Theorie der elektromagnetischen Prozesse im sogenannten leeren Raume besteht ein tiefgreifender formaler Unterschied. Während wir uns nämlich den Zustand eines Körpers durch die Lagen und Geschwindigkeiten einer zwar sehr großen, jedoch endlichen Anzahl von Atomen und Elektronen für vollkommen bestimmt ansehen, bedienen wir uns zur Bestimmung des elektromagnetischen Zustandes eines Raumes kontinuierlicher räumlicher Funktionen, so daß also eine endliche Anzahl von Größen nicht als genügend anzusehen ist zur vollständigen Festlegung des elektromagnetischen Zustandes eines Raumes. Nach der Maxwellschen Theorie ist bei allen rein elektromagnetischen Erscheinungen, also auch beim Licht, die Energie als kontinuierliche Raumfunktion aufzufassen, während die Energie eines ponderabeln Körpers nach der gegenwärtigen Auffassung der Physiker als eine über die Atome und Elektronen erstreckte Summe darzustellen ist. Die Energie eines ponderabeln Körpers kann nicht in beliebig viele, beliebig kleine Teile zerfallen, während sich die Energie eines von einer punktförmigen Lichtquelle ausgesandten Lichtstrahles nach der Maxwellschen Theorie (oder allgemeiner nach jeder Undulationstheorie) des Lichtes auf ein stets wachsendes Volumen sich kontinuierlich verteilt.

Die mit kontinuierlichen Raumfunktionen operierende Undulationstheorie des Lichtes hat sich zur Darstellung der rein optischen Phänomene vortrefflich bewährt und wird wohl nie durch eine andere Theorie ersetzt werden. Es ist jedoch im Auge zu behalten, daß sich die optischen Beobachtungen auf zeitliche Mittelwerte, nicht aber auf Momentanwerte beziehen, und es ist trotz der vollständigen Bestätigung der Theorie der Beugung, Reflexion, Brechung, Dispersion etc. durch das

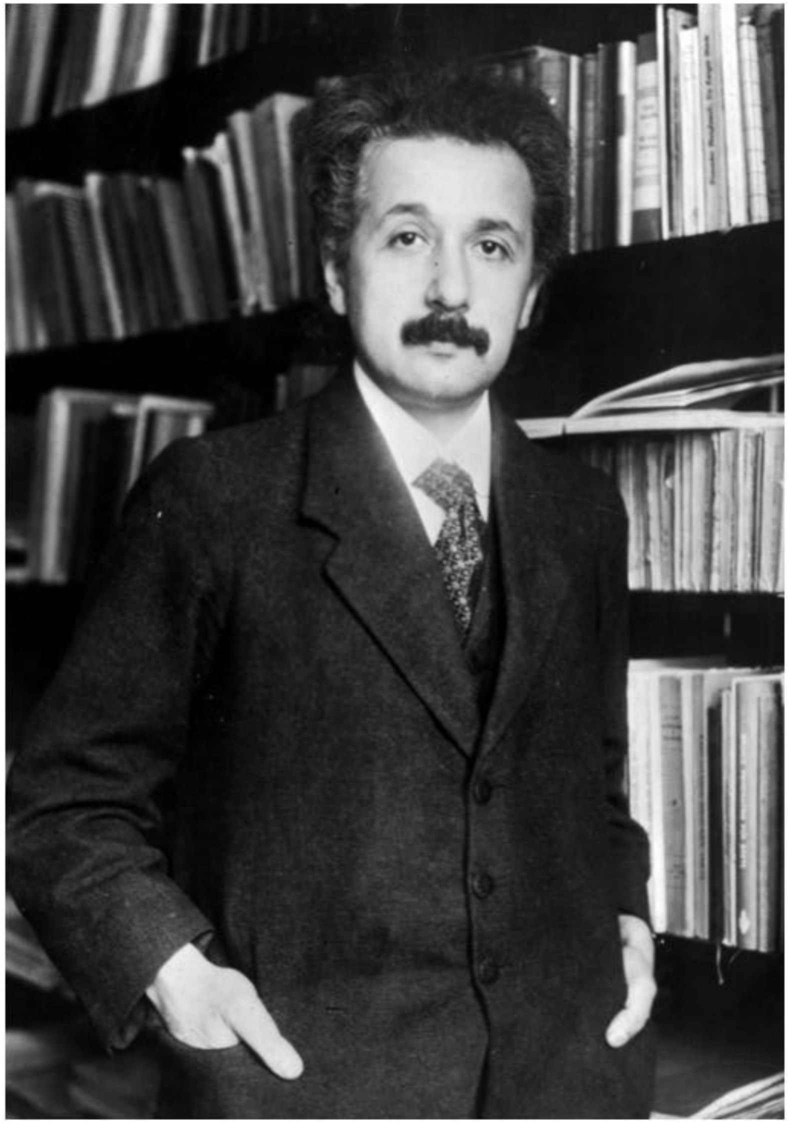
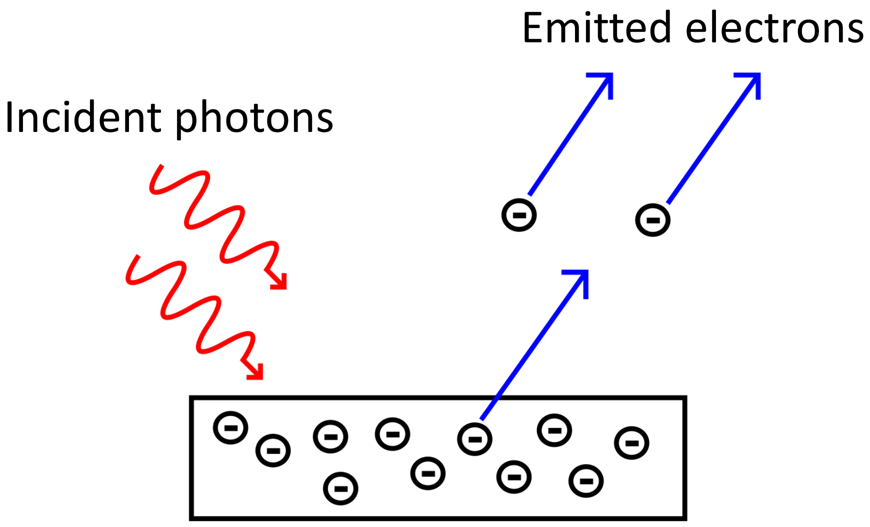


Photo-Electric Effect

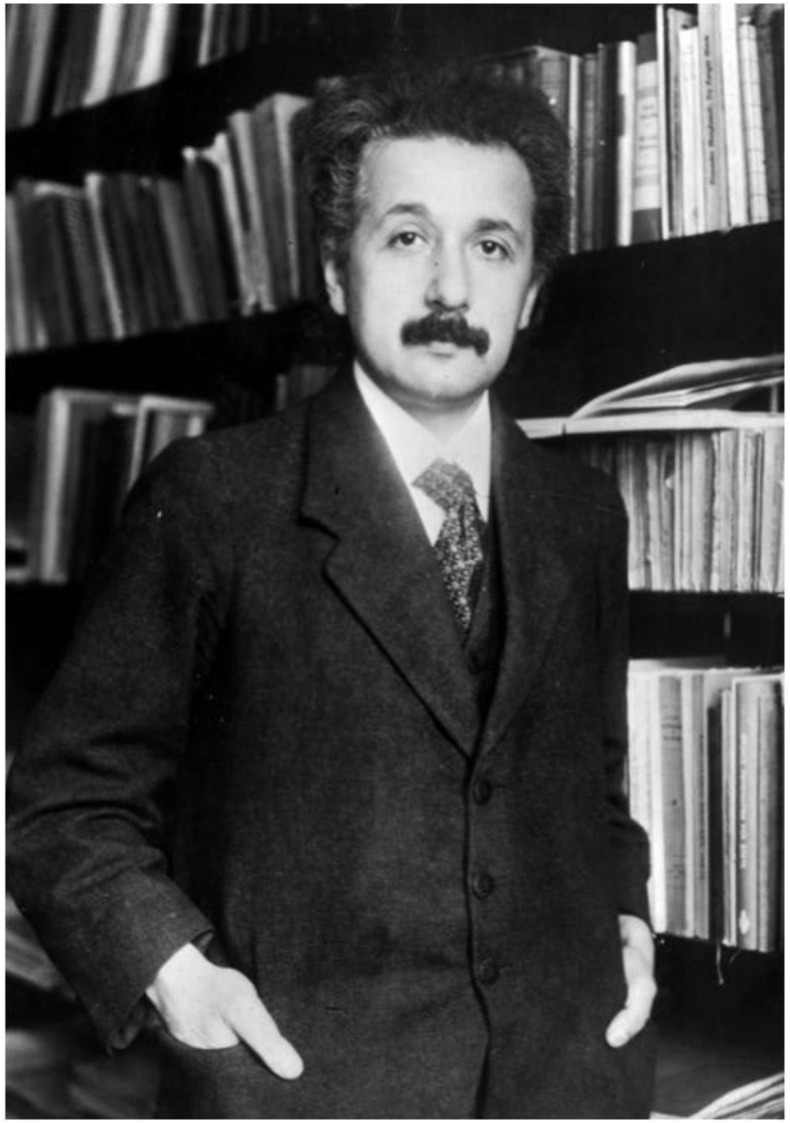


132

6. *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt;*
von A. Einstein.

Zwischen den theoretischen Vorstellungen, welche sich die Physiker über die Gase und andere ponderable Körper gebildet haben, und der Maxwellschen Theorie der elektromagnetischen Prozesse im sogenannten leeren Raume besteht ein tiefgreifender formaler Unterschied. Während wir uns nämlich den Zustand eines Körpers durch die Lagen und Geschwindigkeiten einer zwar sehr großen, jedoch endlichen Anzahl von Atomen und Elektronen für vollkommen bestimmt ansehen, bedienen wir uns zur Bestimmung des elektromagnetischen Zustandes eines Raumes kontinuierlicher räumlicher Funktionen, so daß also eine endliche Anzahl von Größen nicht als genügend anzusehen ist zur vollständigen Festlegung des elektromagnetischen Zustandes eines Raumes. Nach der Maxwellschen Theorie ist bei allen rein elektromagnetischen Erscheinungen, also auch beim Licht, die Energie als kontinuierliche Raumfunktion aufzufassen, während die Energie eines ponderabeln Körpers nach der gegenwärtigen Auffassung der Physiker als eine über die Atome und Elektronen erstreckte Summe darzustellen ist. Die Energie eines ponderabeln Körpers kann nicht in beliebig viele, beliebig kleine Teile zerfallen, während sich die Energie eines von einer punktförmigen Lichtquelle ausgesandten Lichtstrahles nach der Maxwellschen Theorie (oder allgemeiner nach jeder Undulationstheorie) des Lichtes auf ein stets wachsendes Volumen sich kontinuierlich verteilt.

Die mit kontinuierlichen Raumfunktionen operierende Undulationstheorie des Lichtes hat sich zur Darstellung der rein optischen Phänomene vortrefflich bewährt und wird wohl nie durch eine andere Theorie ersetzt werden. Es ist jedoch im Auge zu behalten, daß sich die optischen Beobachtungen auf zeitliche Mittelwerte, nicht aber auf Momentanwerte beziehen, und es ist trotz der vollständigen Bestätigung der Theorie der Beugung, Reflexion, Brechung, Dispersion etc. durch das



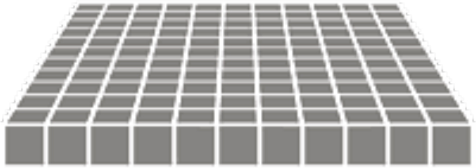
Einstein's Nobel Prize in Physics in 1921, "for his services to Theoretical Physics, and especially for his discovery of the law of photoelectric effect."

Role of an imaging sensor

Camera shutter opens

↓

Exposure begins



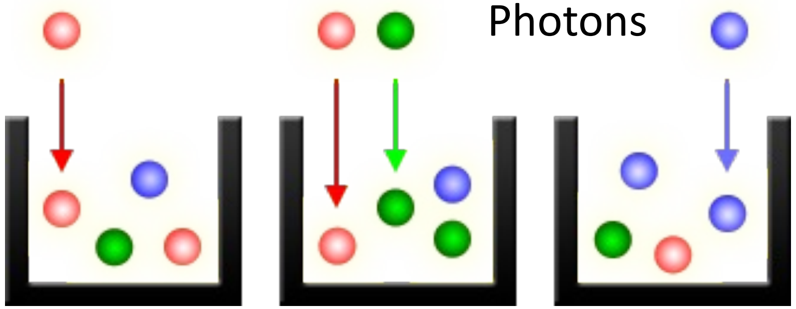
Array of photon buckets

↓

Photon buckets begin to store photons

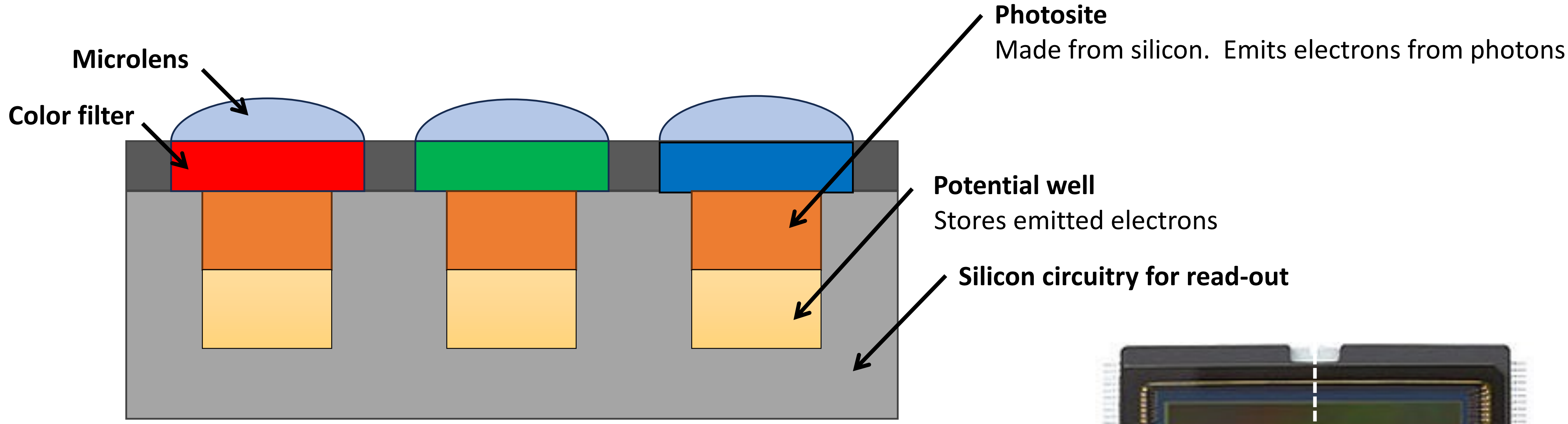
↓

Camera shutter closes and “stored photons” are converted to intensity values

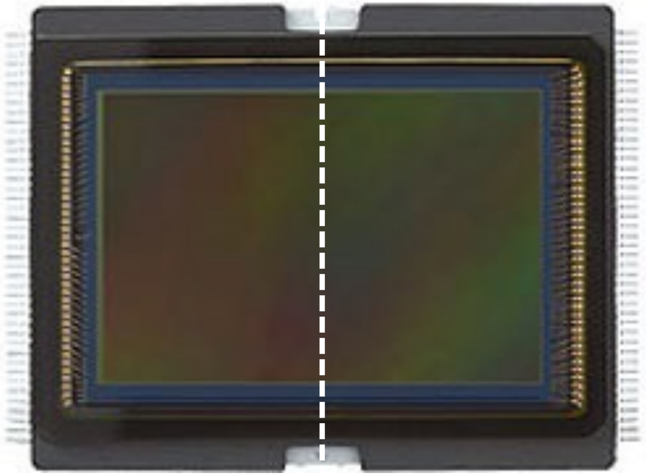


Close-up view of photon buckets

Basic Imaging Sensor Design



Canon 6D sensor
(20.2 MP, Full-Frame)



Quantum efficiency of Photosite (QE)

- How many incident photons will the photosite convert to electrons

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

- Fundamental optical performance metric of imaging sensors
    - But not the only one

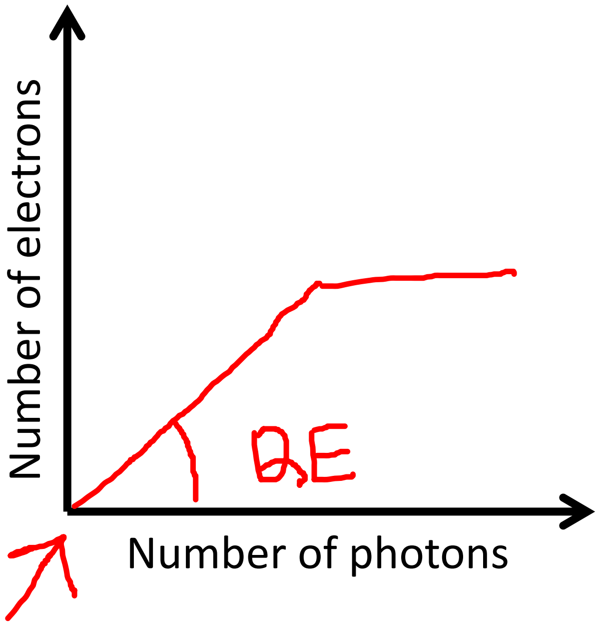
Emitted electrons

Incident photons

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

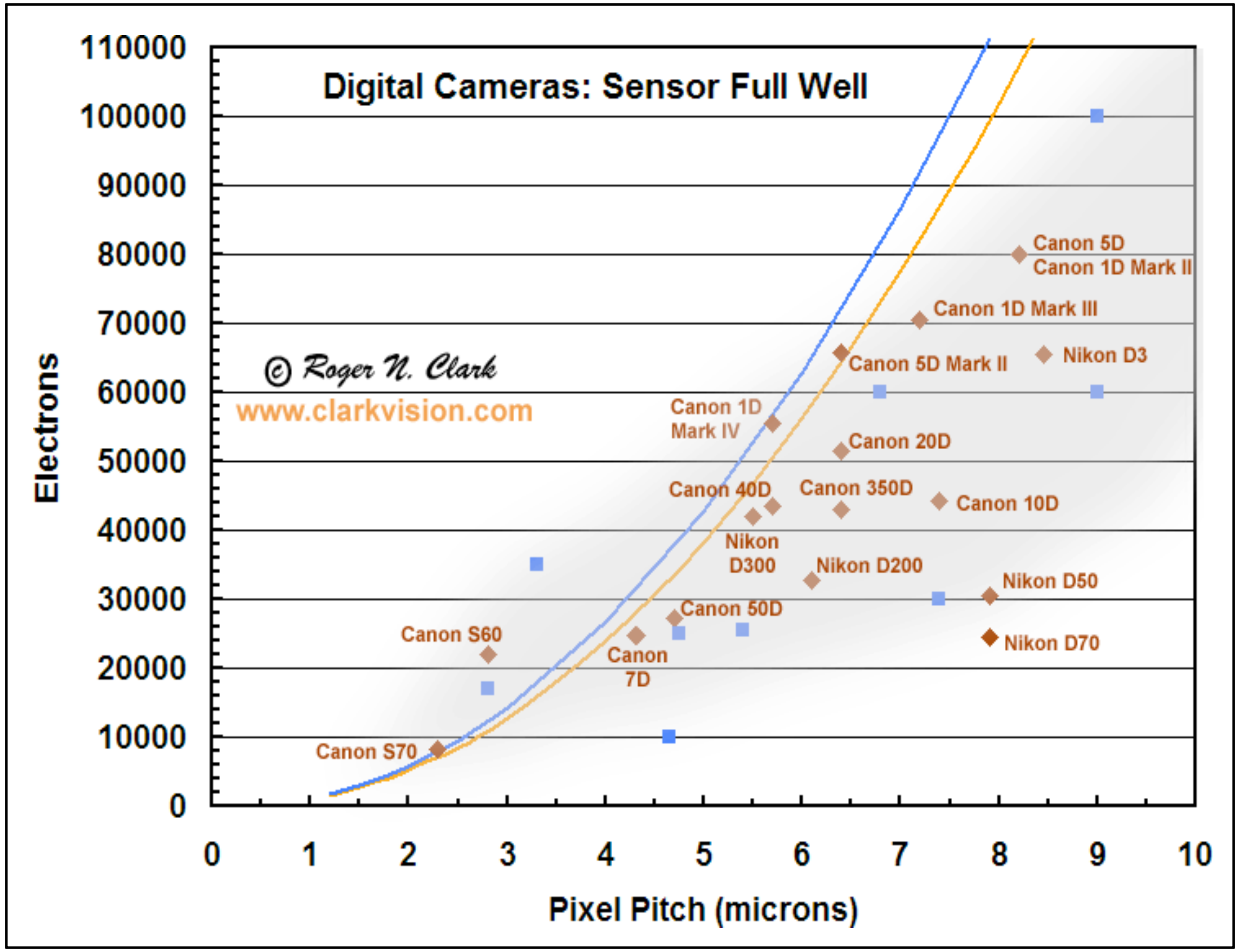
Photosite response

- The photosite response is mostly linear; however, it is non-linear under
  - Saturation (too many photons, potential well is full before exposure ends); and
  - Under-exposure (too few photons, sensor noise or thermal noise)

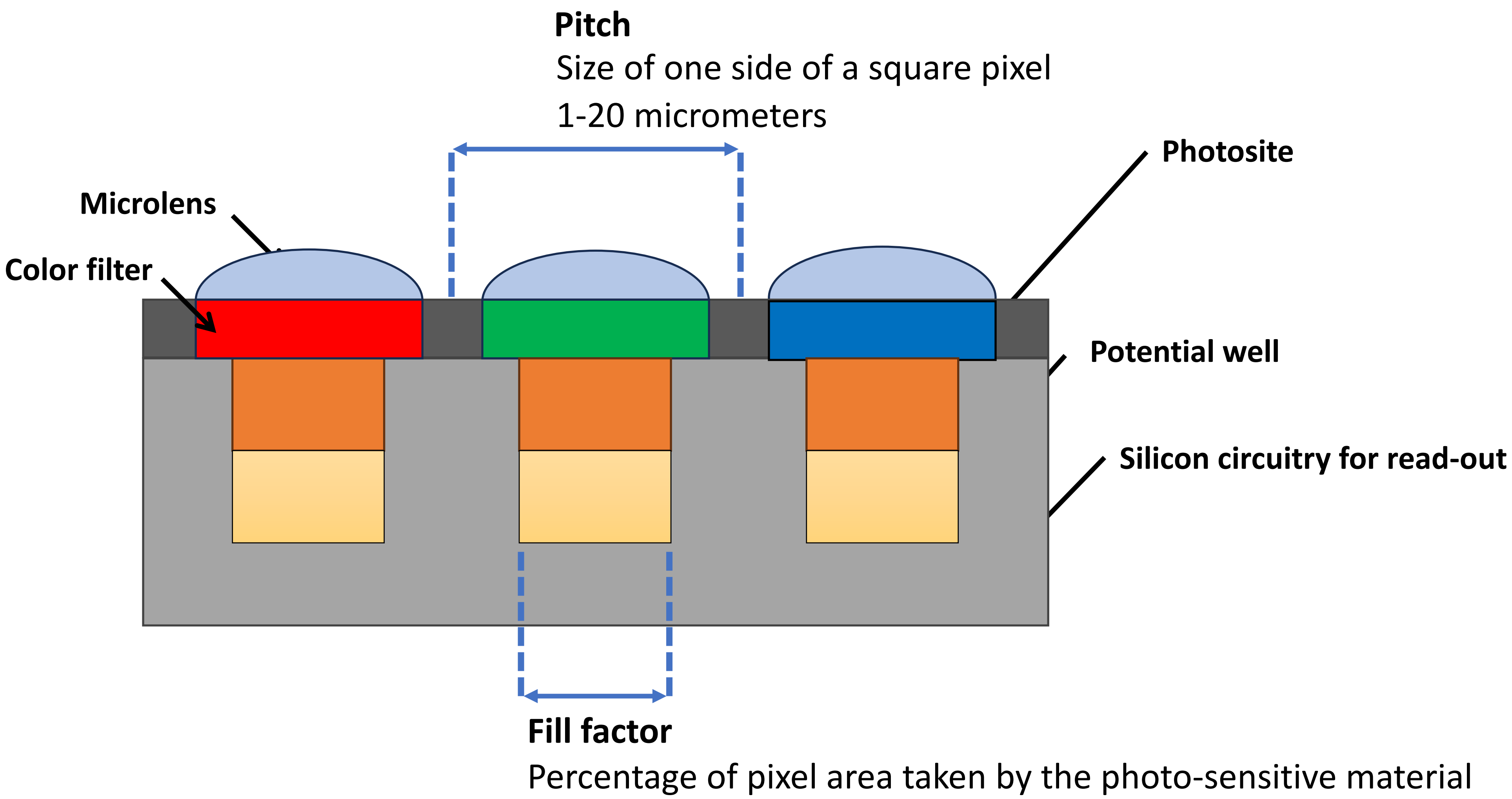


Photosite full-well capacity

- How many electrons can a photosite stores before saturation?
- Important optical performance metric of imaging sensor

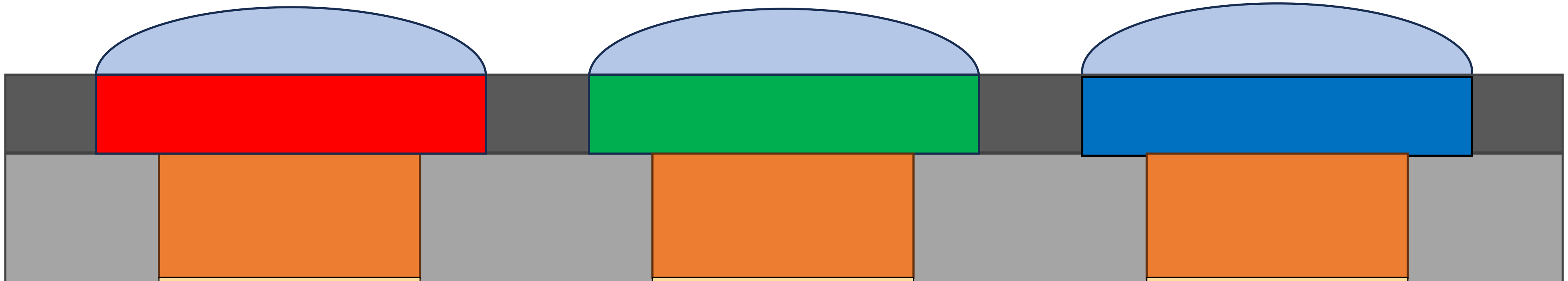


Pixel pitch and fill factor

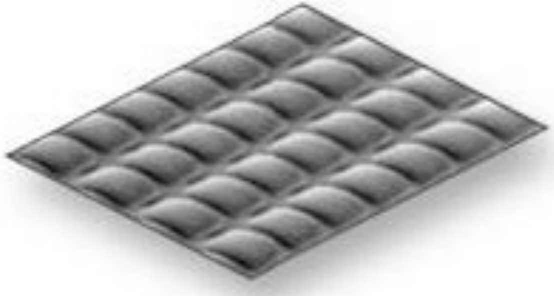


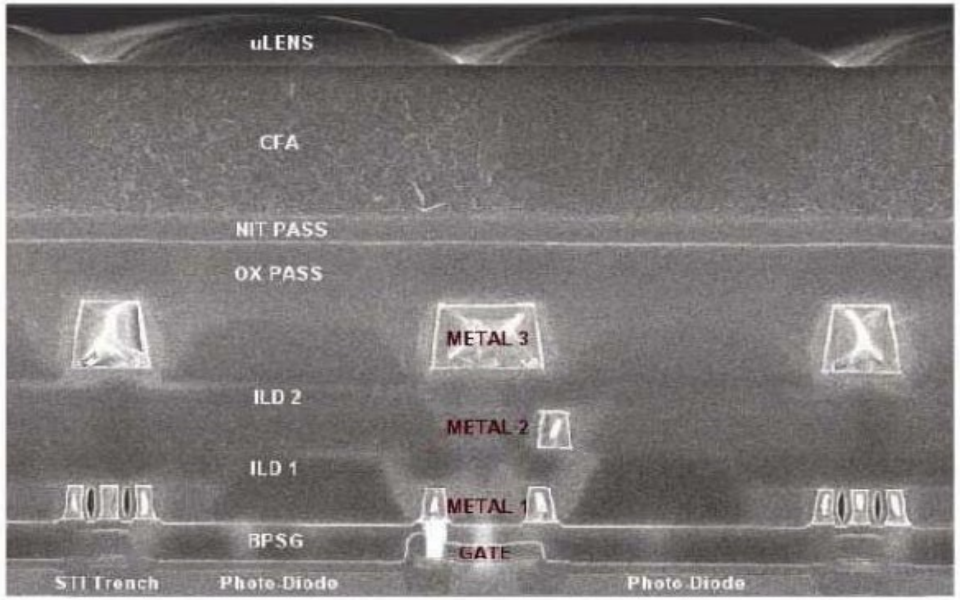
Microlenses

- **Microlenses** increase the effective fill area by bending light towards the photosite
- These are sometimes called lenslets
- Microlenses implement a pixel-sized 2d rect filter and act as spatial low-pass filter, thus preventing anti-aliasing

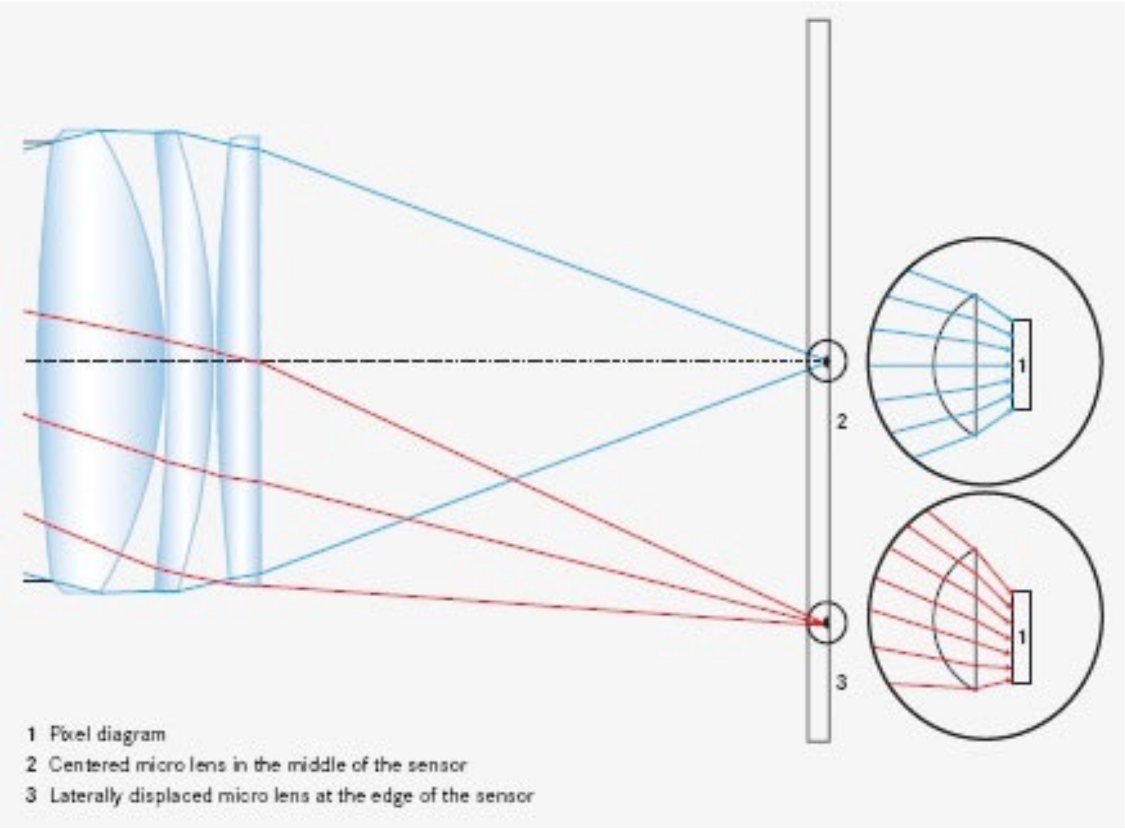


Microlenses





Sensor cross-section



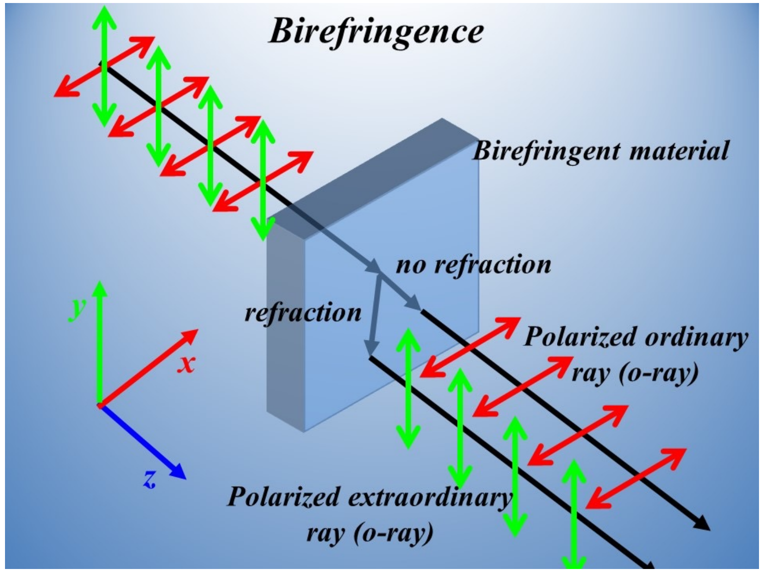
Shifted microlenses for increased fill factor

Optical Low-Pass Filter (OLPF)

- Oftentimes an optional low-pass filter is placed in front of the sensor
  - Sensors have a separate glass sheet in front of the sensor, which acts as OPLF
- Improves pre-filter
- Prevents anti-aliasing

Optical Low-Pass Filter (OLPF)

- Two birefringent layers
- Split one ray to four rays, implementing a four-tap discrete convolutional filter

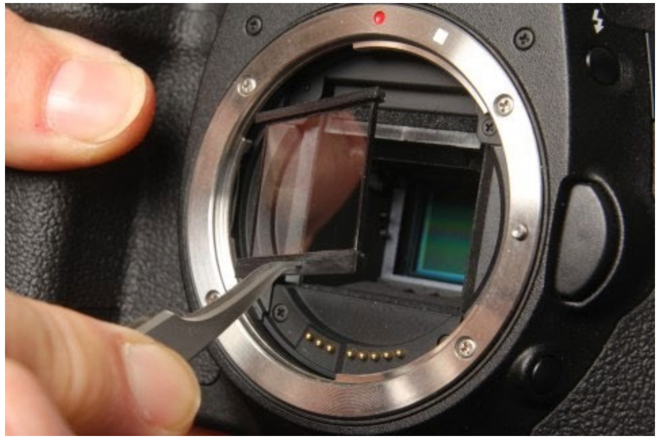


ScienceDirect

Birefringence in a calcite crystal

Optical Low-Pass Filter (OLPF)

- The OLPF means you also lose resolution.
- Nowadays, due the large number of pixels, OLPF are becoming unnecessary.
  - Photographers often hack their cameras to remove the OLPF, to avoid the loss of resolution (“hot rodding”).
  - Camera manufacturers offer camera versions with and without an OLPF.
  - The OLPF can be problematic also when working with coherent light
    - Coherent light is electromagnetic radiation that has a certain wavelength (think, lasers)



Optical Low-Pass Filter (OLPF)

OLPF

Optical Low-Pass Filter (OLPF)

OLPF

Optical Low-Pass Filter (OLPF)

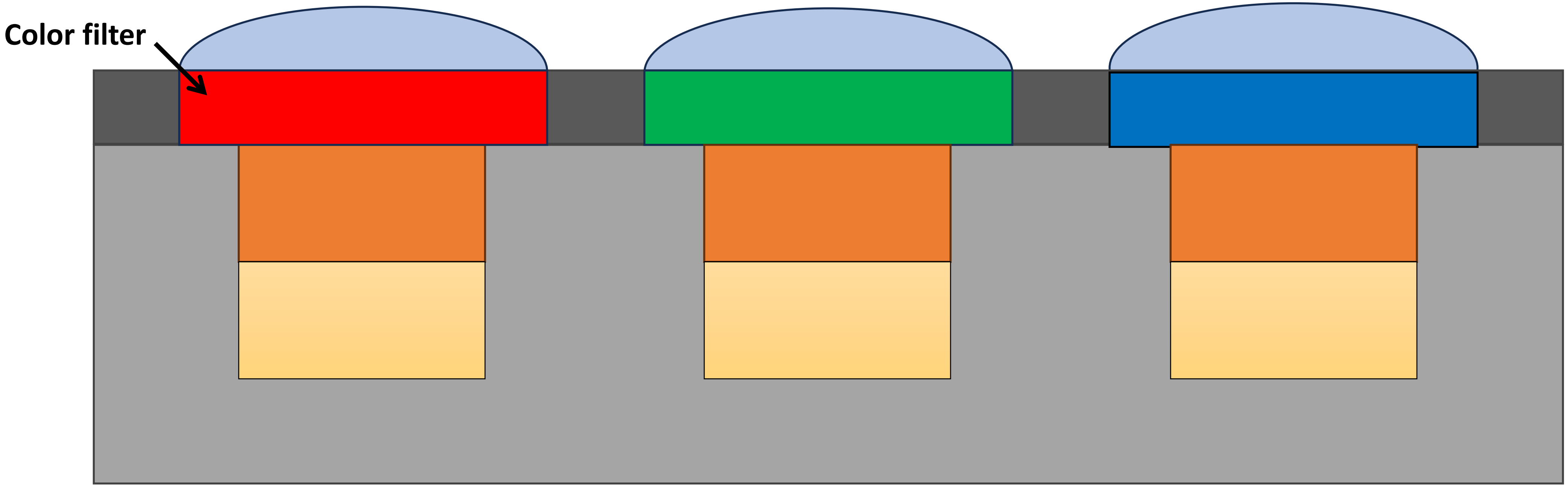
- Camera companies offer identical models with and without OLPF

Nikon D800

Nikon D800E

Color Filter Arrays (CFA)

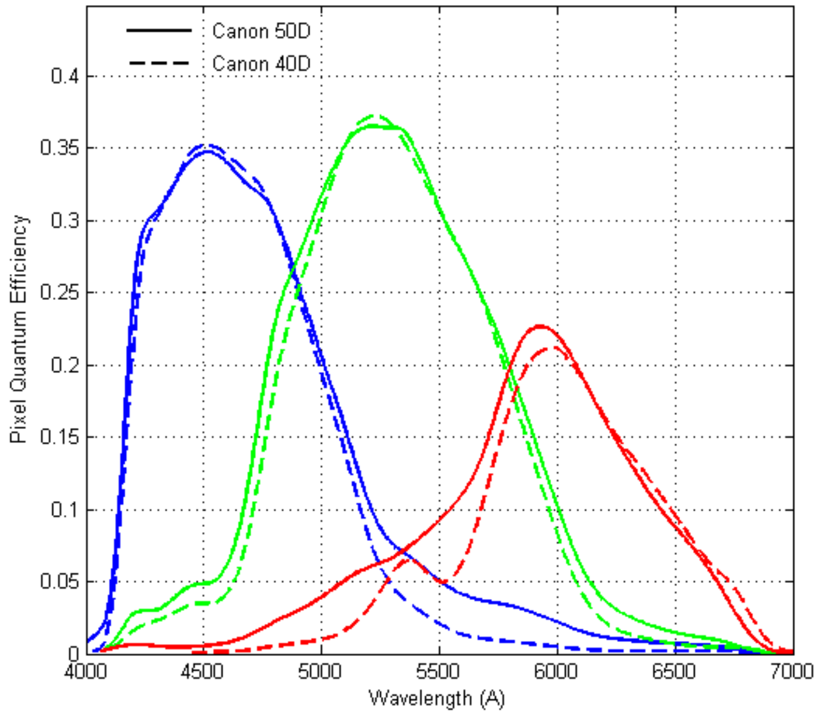
- Mimics cone cells of human vision system
- Allow us to measure color with a digital sensor
- “Pixels” with different filters, each have their own spectral sensitivity



Which color filters to use?

- Which spectral sensitivity functions $f(\lambda)$ to use for each color filter?
  - Do not necessarily match human color perception

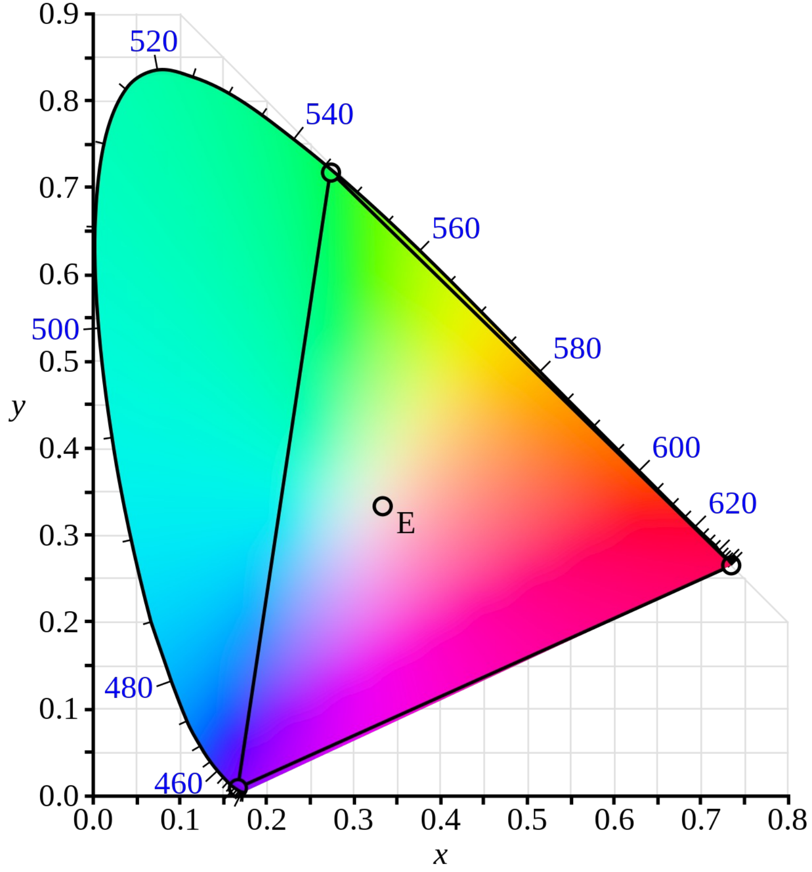
SSF for Canon 50D



$f(\lambda)$

Color space

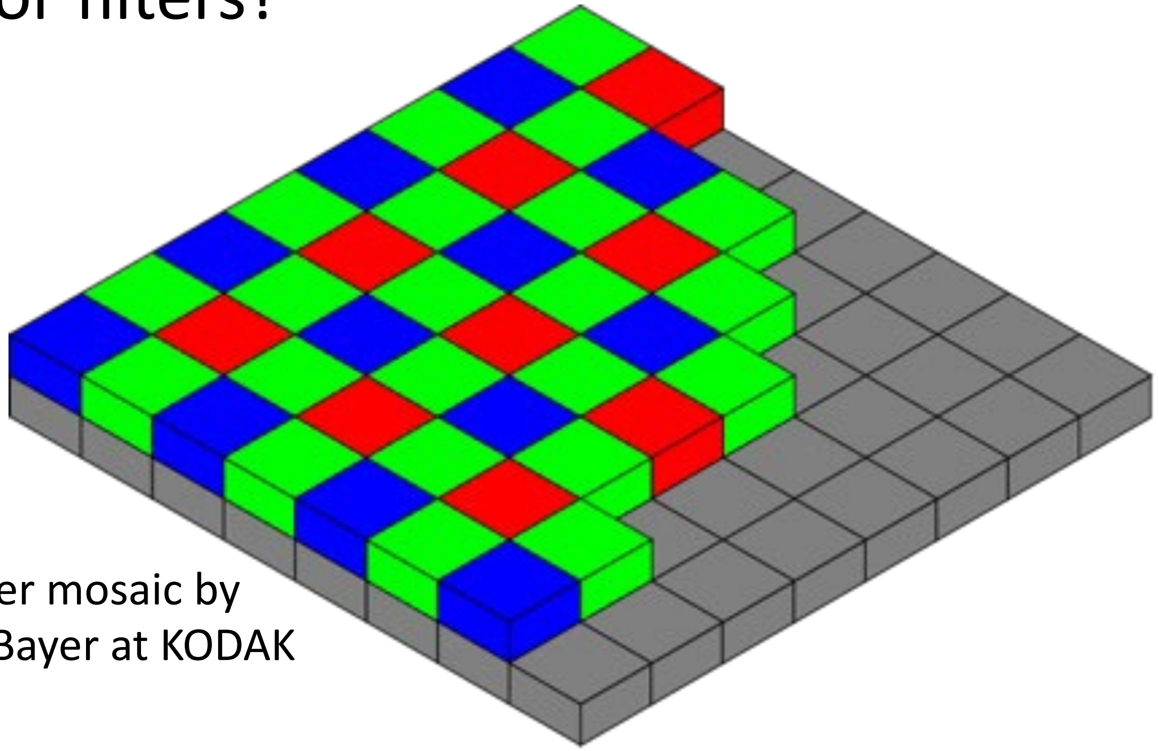
- CIE 1931 color spaces
  - International Commission on Illumincation (CIE)
  - Experiments by William David Wright and John Guild in 1920s
- Links the distribution of wavelengths in the electromagnetic visible spectrum and physiologically perceived colors in human color vision



Gamut of the CIE RGB primaries and location of primaries on the CIE 1931 xy chromaticity diagram. (Wikipedia)

Which color filters to use?

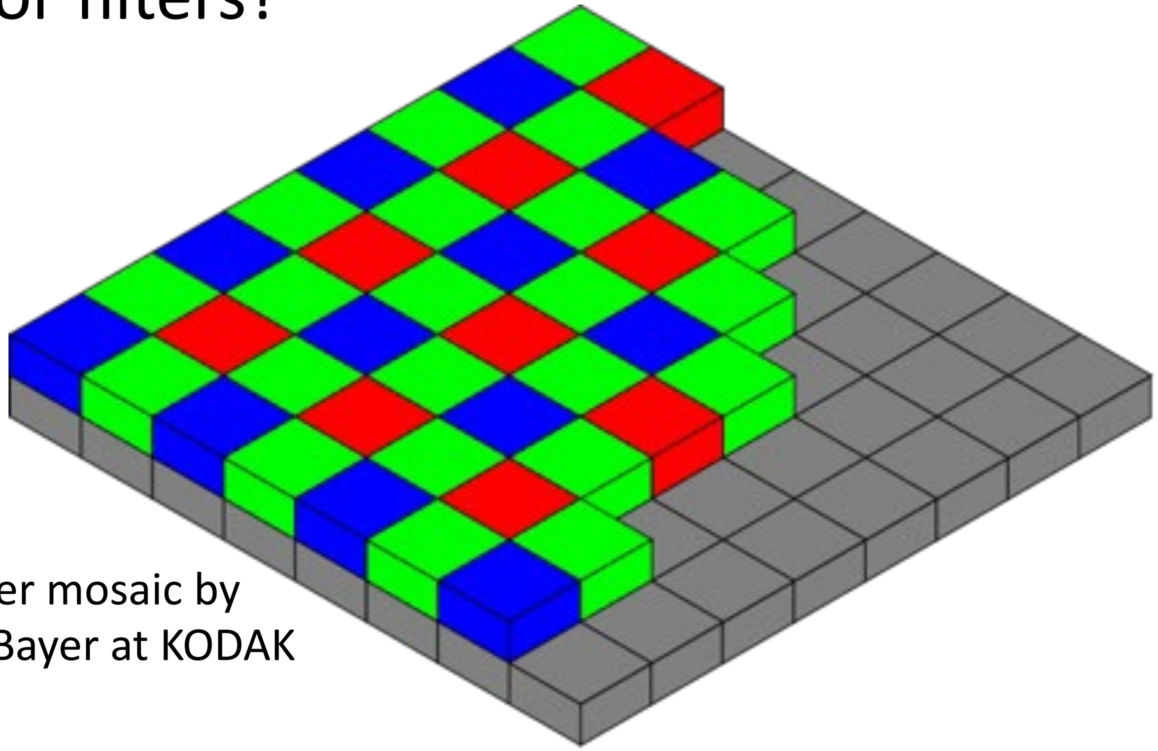
- Which Spectral Sensitivity Functions (SSF) to use for each color filter?
  - Do not necessarily match human color perception
- How to spatially arrange different color filters?
  - Looks like a mosaic



Bayer mosaic by
Bryce Bayer at KODAK

Which color filters to use?

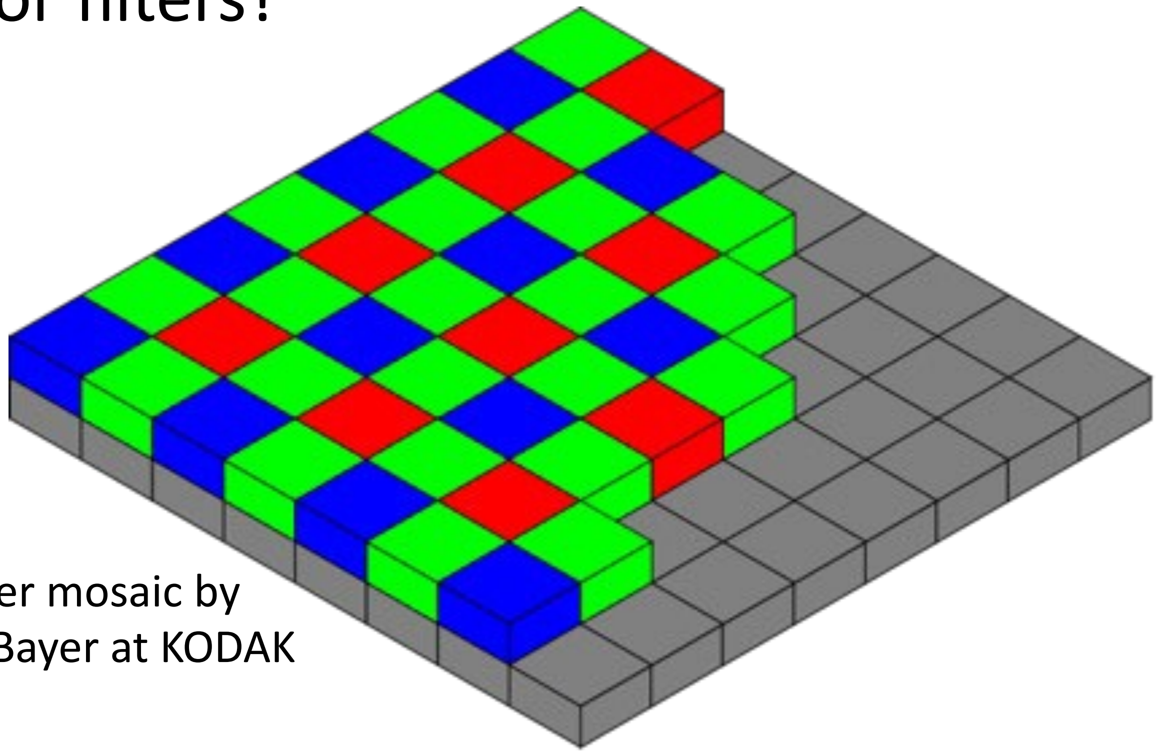
- Which Spectral Sensitivity Functions (SSF) to use for each color filter?
  - Do not necessarily match human color perception
- How to spatially arrange different color filters?
  - Looks like a mosaic
- Bayer mosaic
  - Why more green pixels than red or blue pixels?



Bayer mosaic by
Bryce Bayer at KODAK

Which color filters to use?

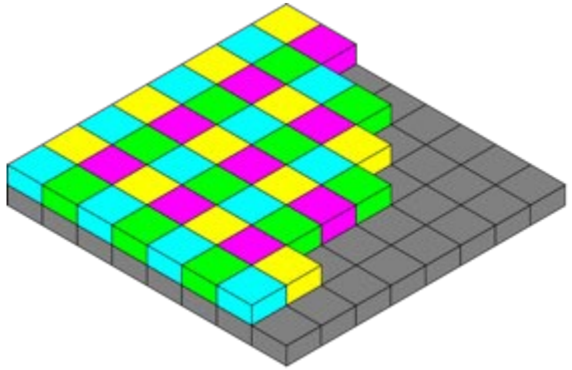
- Which Spectral Sensitivity Functions (SSF) to use for each color filter?
  - Do not necessarily match human color perception
- How to spatially arrange different color filters?
  - Looks like a mosaic
- Bayer mosaic
  - Why more green pixels than red or blue pixels?
    - Because human eye is more sensitive to green colors



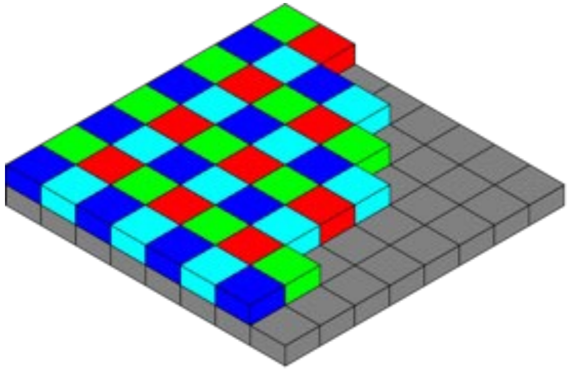
Bayer mosaic by
Bryce Bayer at KODAK

Many other options for CFA

- Many other options

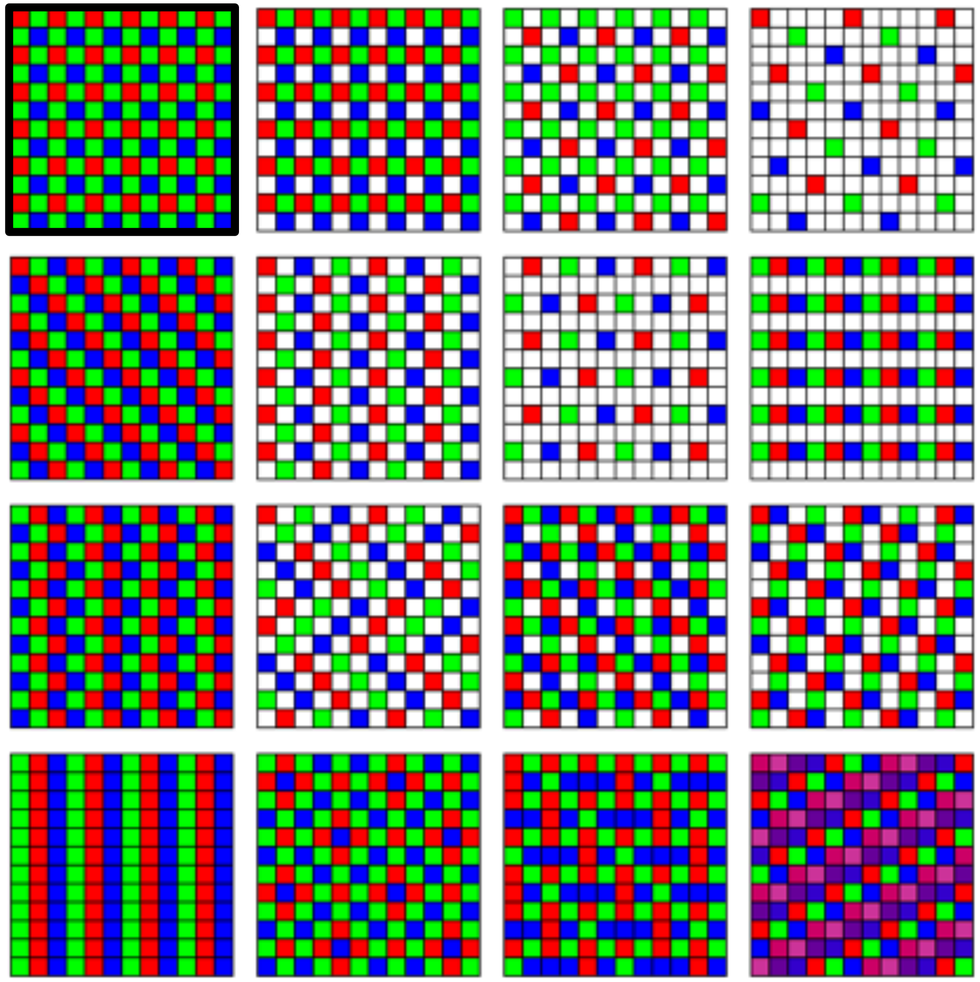


CYGM
Canon IXUS, Powershot



RGBE
Sony Cyber-shot

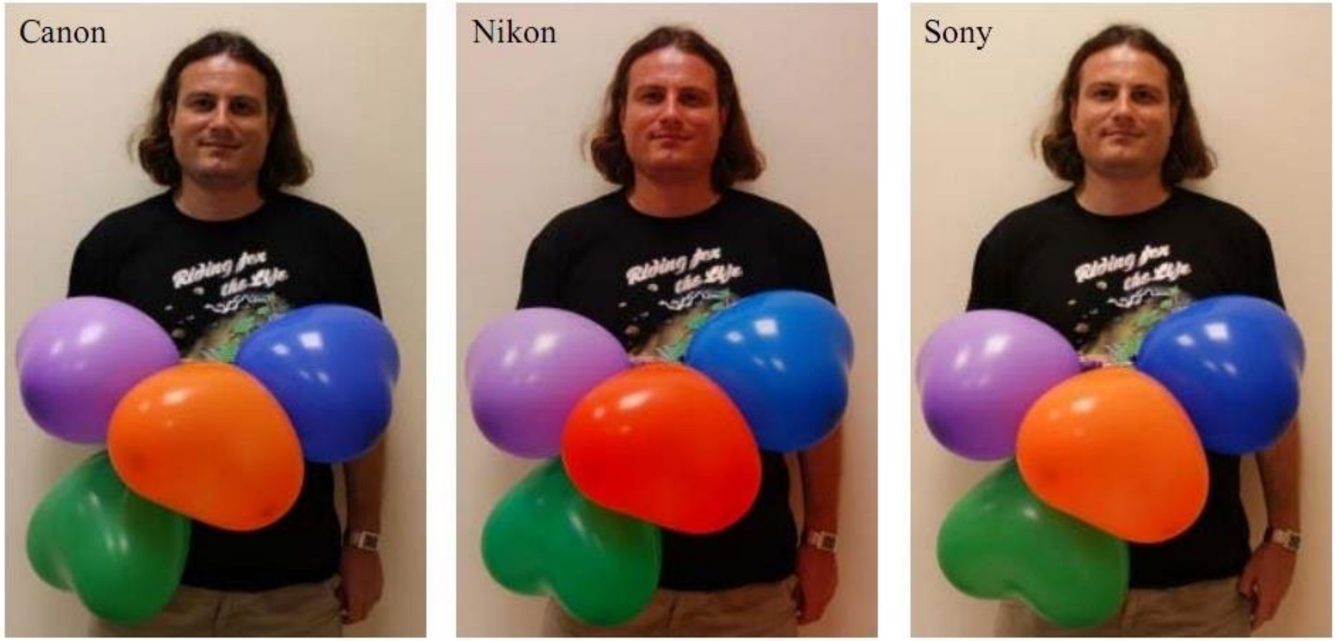
- Find the best CFA is an active area of research



Many different spectral sensitivity functions

- Each camera has its more or less unique, and most of the time secret, SSF.
- Makes it very difficult to correctly reproduce the color of sensor measurements.

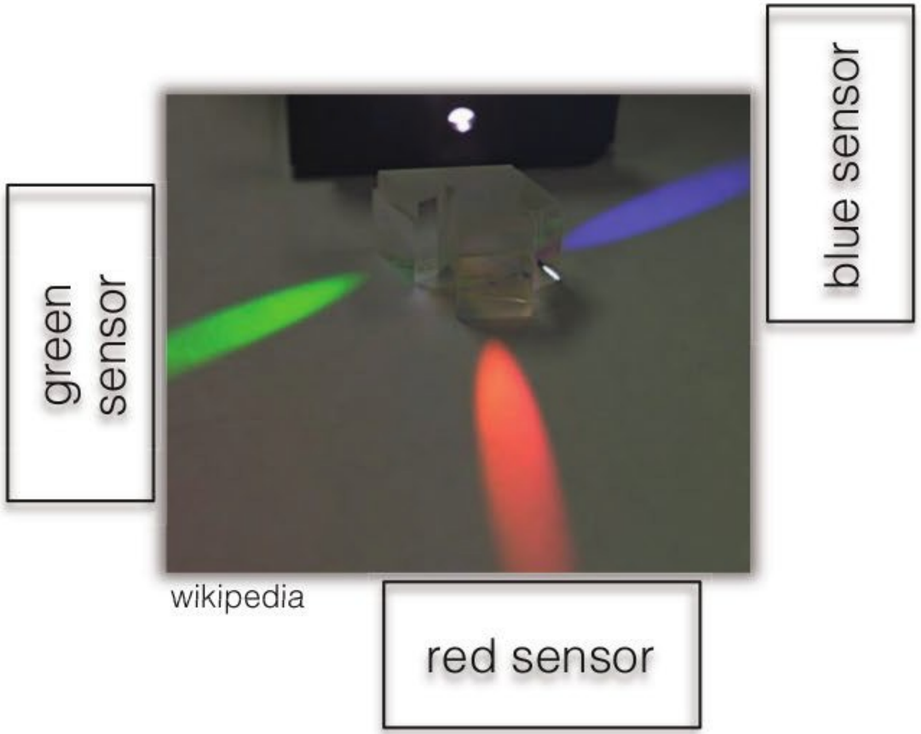
Identical scene captured with three cameras



Other ways for capturing color

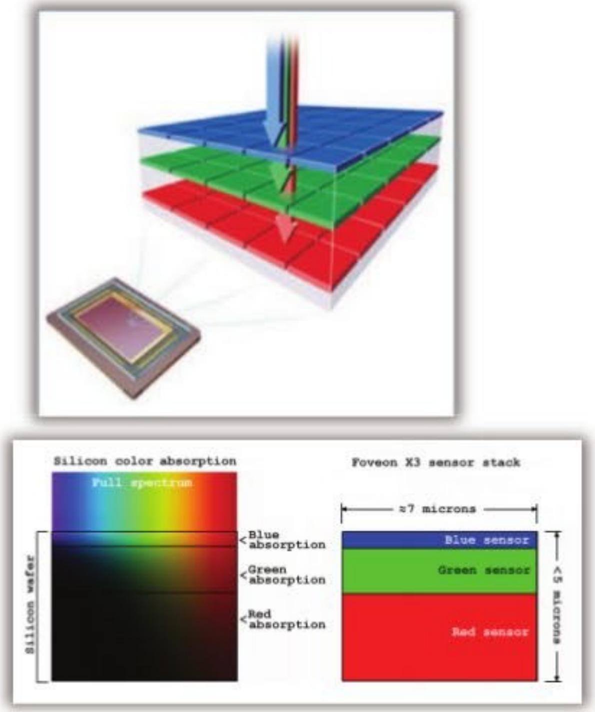
Prokudin-Gorsky

Sequential



wikipedia

Multiple sensors



Foveon X3

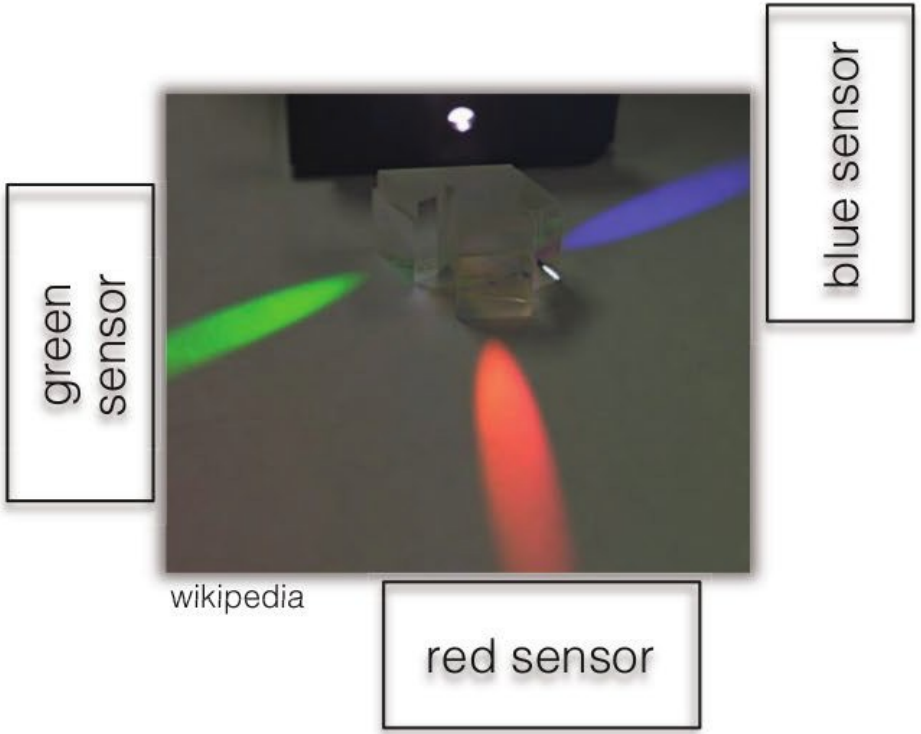
Vertically Stacked

Other ways of capturing color

field sequential

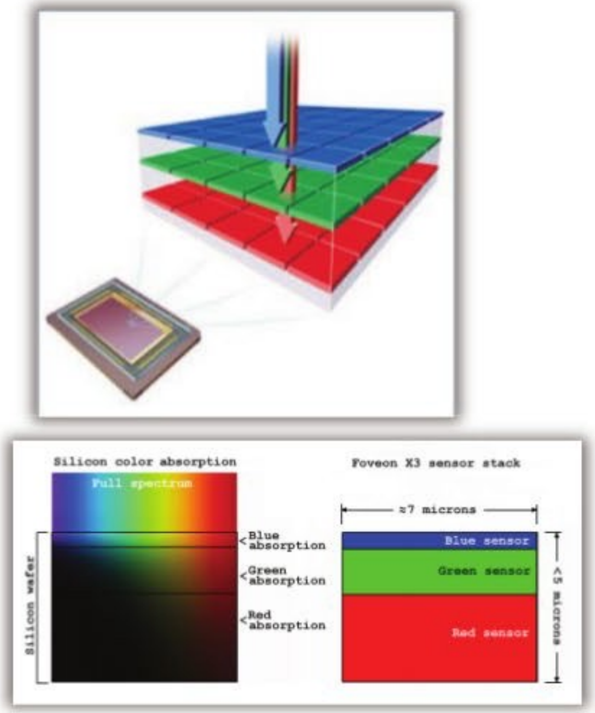
Prokudin-Gorsky

multiple sensors



wikipedia

vertically stacked

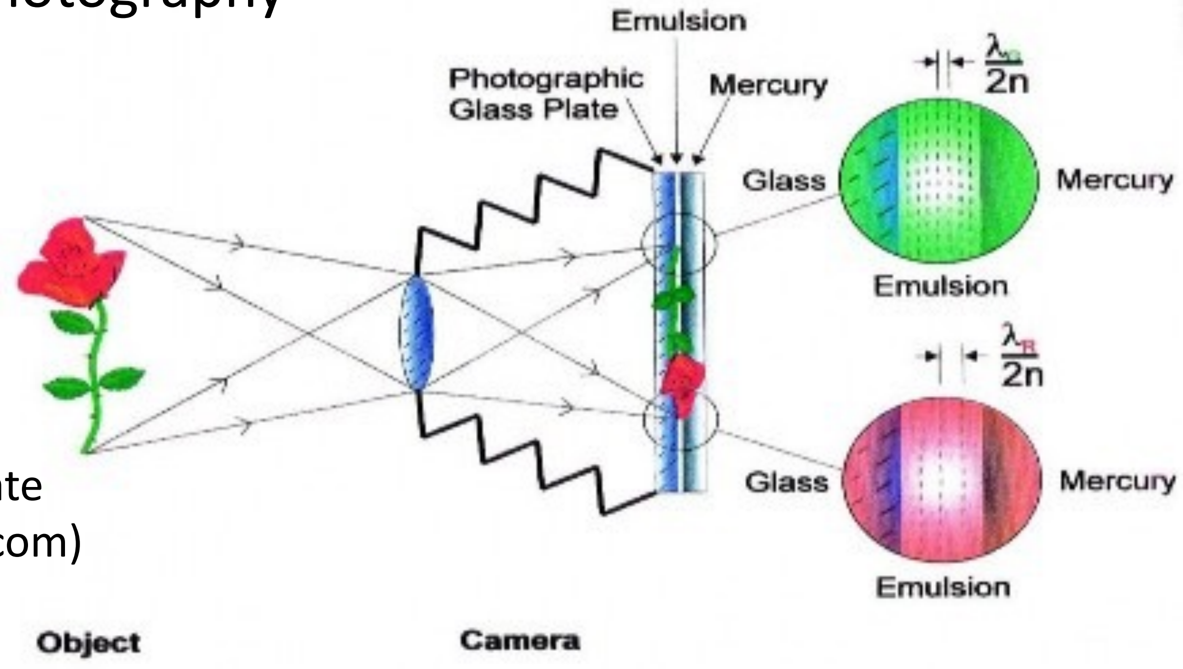


Foveon X3

[Slide credit: Gordon Wetzstein]

Other ways of capturing color

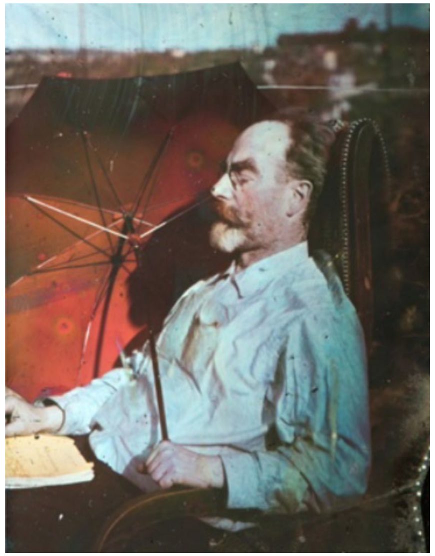
- Interferential Photography
  - Also called Lippmann Photography
  - Gabriel Lippmann was awarded the Nobel Prize in Physics in 1908 for his discovery of interference photography



The principle of Lippmann plate
(From AlternativePhotography.com)

Other ways of capturing color

- Interferential Photography
    - Also called Lippmann Photography
    - Gabriel Lippmann was awarded the Nobel Prize in Physics in 1908 for his discovery of interference photography



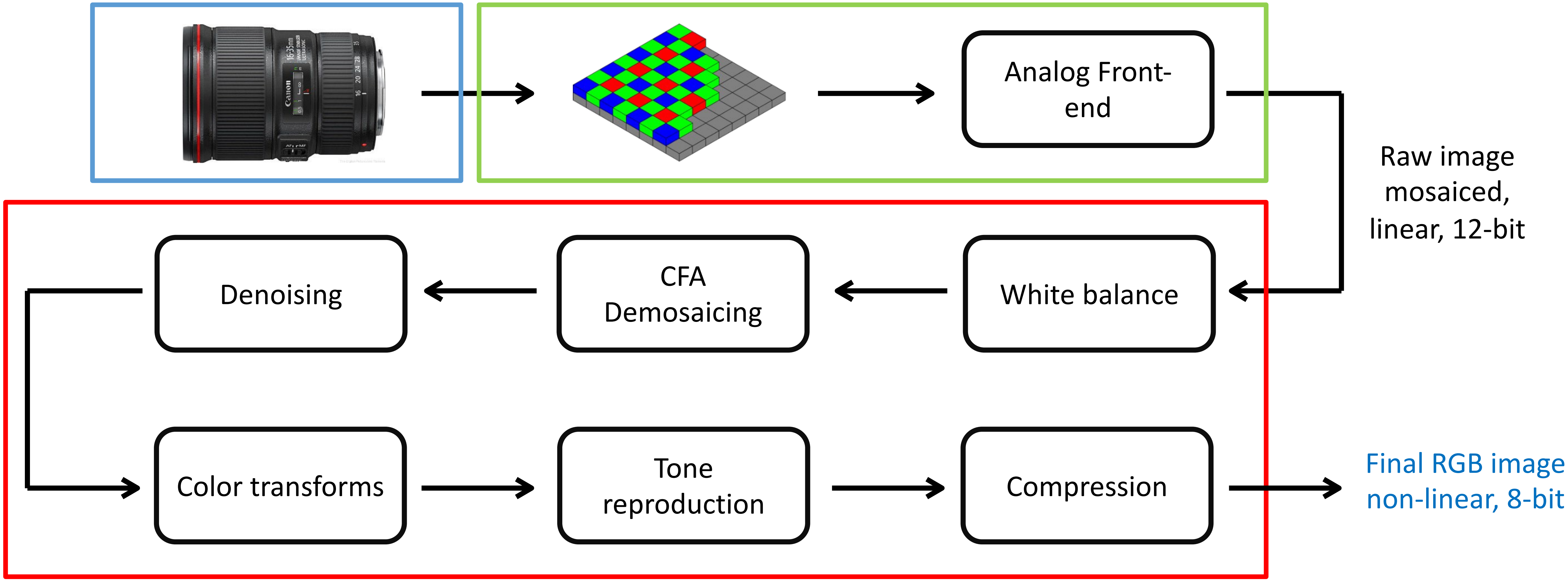
Lippmann Self-Portrait

Lippmann photograph by Neuhauss, recorded in 1899
(Collection: Royal Photographic Society)

The (in-camera) image processing pipeline

- The sequence of image processing operations needed to convert a raw image to a conventional image

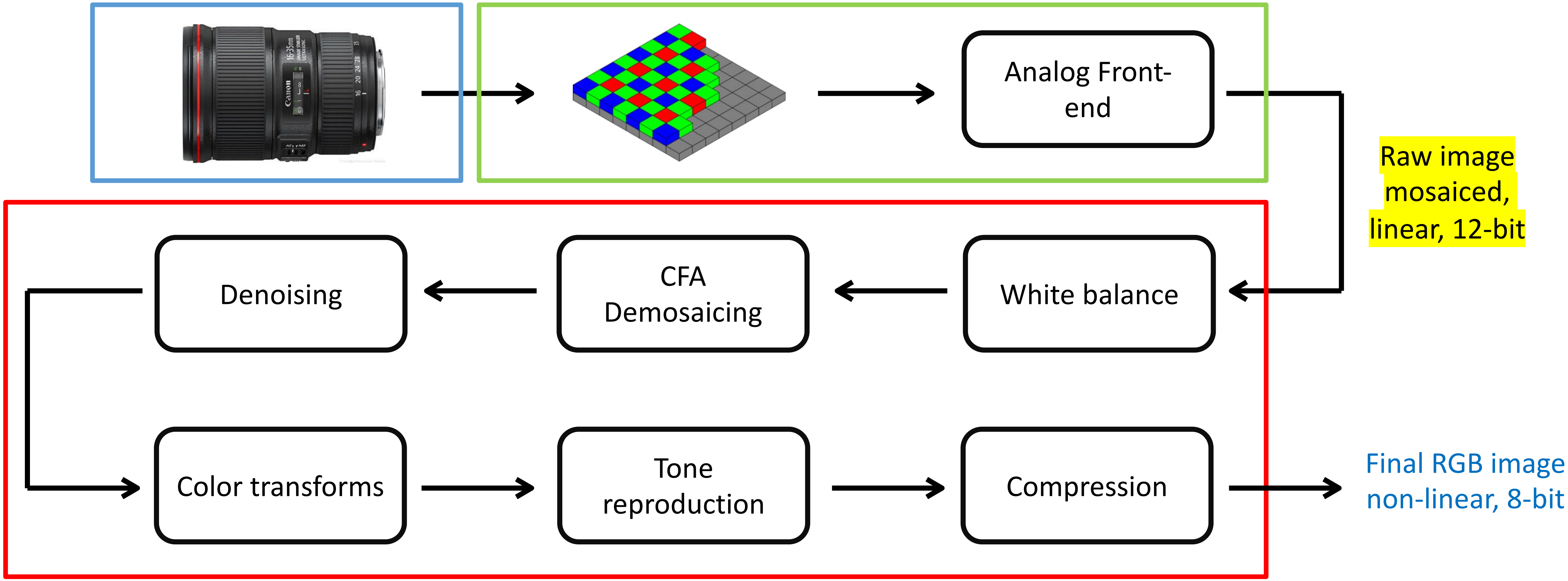
The (in-camera) image processing pipeline



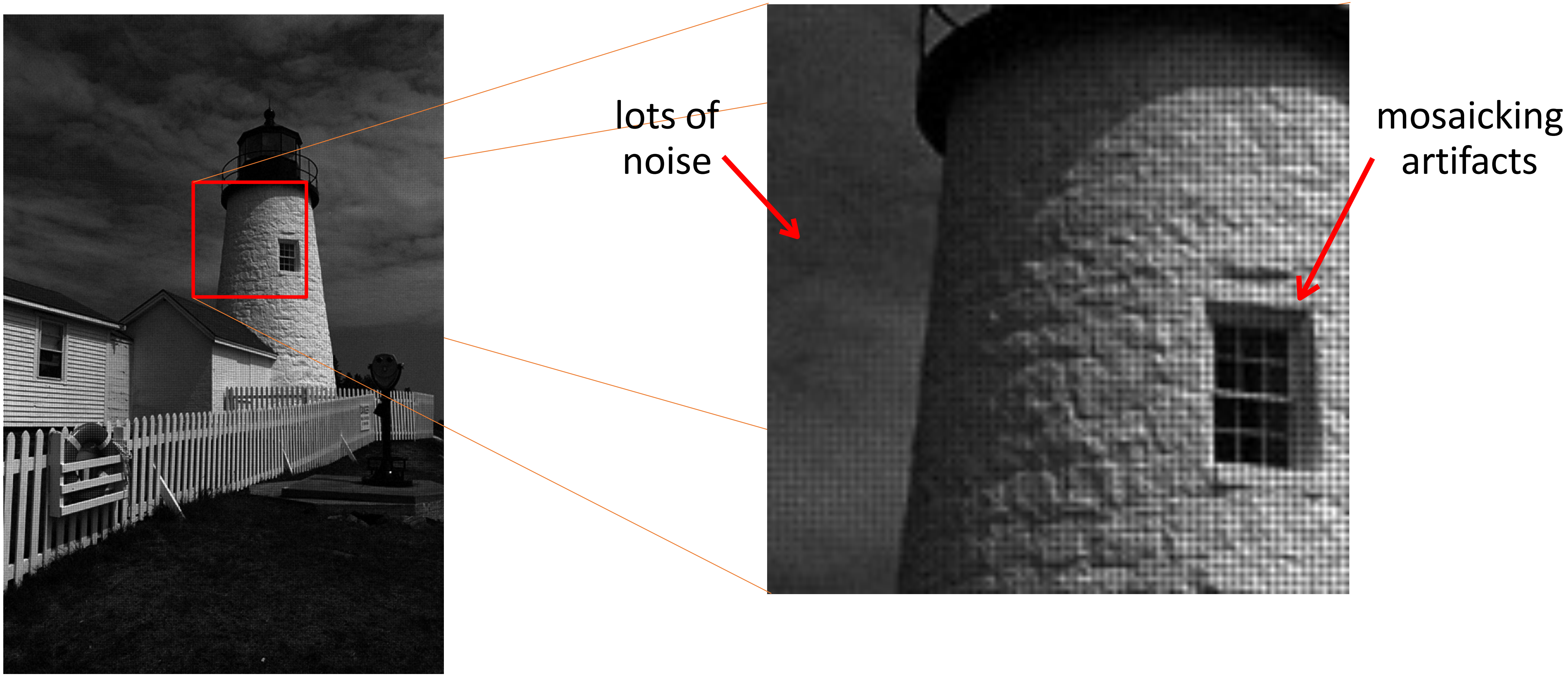
The (in-camera) image processing pipeline

- The sequence of image processing operations needed to convert a raw image to a conventional image
- The process of converting a RAW image to a “conventional” image is sometimes called *rendering*
  - Not the same rendering that you see in computer graphics
- The inverse process of going from a “conventional” RGB image back to a RAW image is called *derendering*

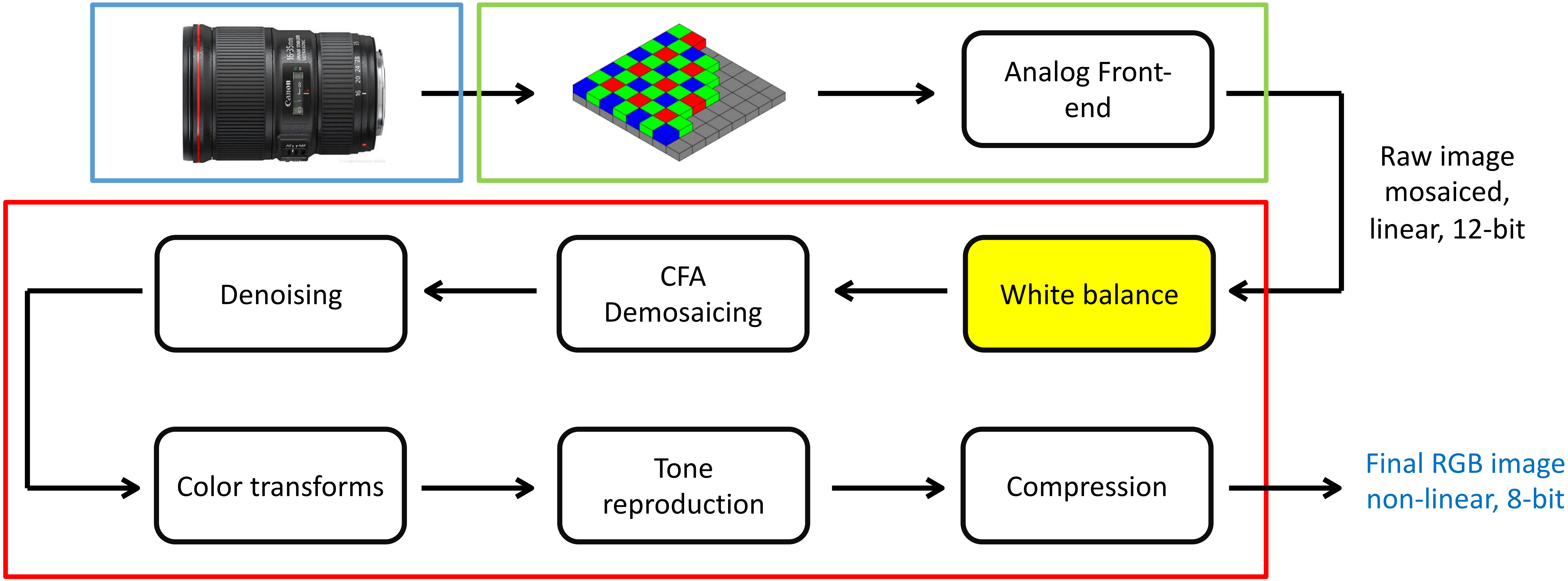
The (in-camera) image processing pipeline



Raw image



The (in-camera) image processing pipeline



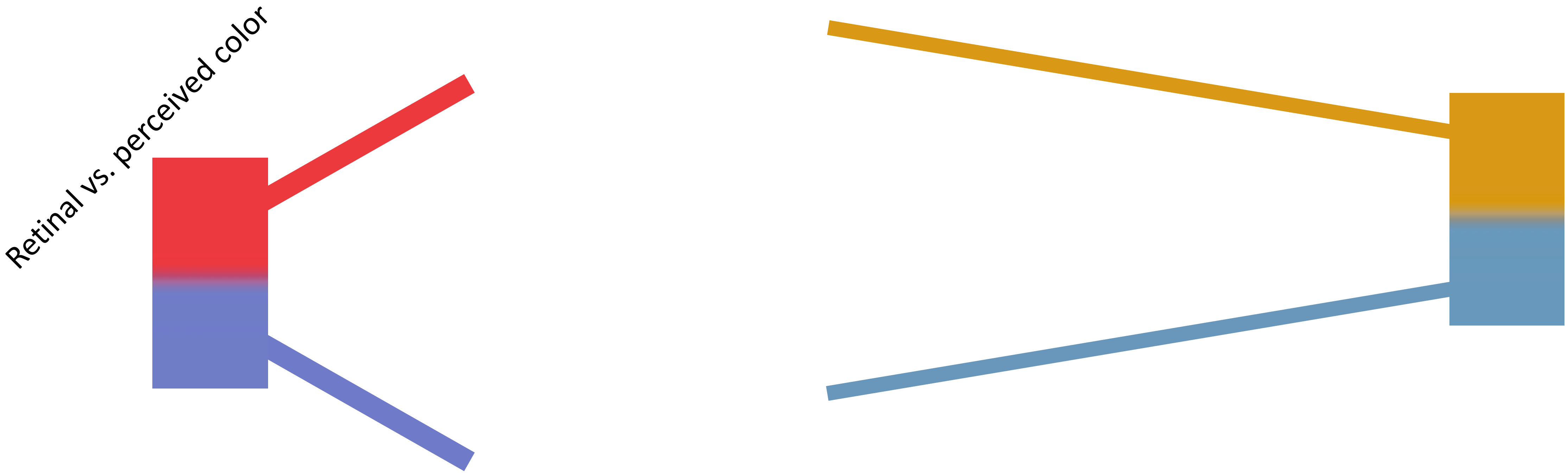
White balancing

- Human visual system has *chromatic adaptation*
  - We can perceive white (and other colors) correctly under different light sources.

[Slide credit: Todd Zickler]

White balancing

- Human visual system has *chromatic adaptation*
  - We can perceive white (and other colors) correctly under different light sources.



[Slide credit: Todd Zickler]

White balancing

- Human visual system has *chromatic adaptation*
  - We can perceive white (and other colors) correctly under different light sources
- Retinal vs. perceived color
- Cameras cannot do it, since these do not have perception

different whites

image captured under fluorescent

image white- balanced to daylight

White balancing

- The process of removing color casts so that colors that we would *perceive* as white are *rendered* as white in final image.

different whites

image captured under fluorescent

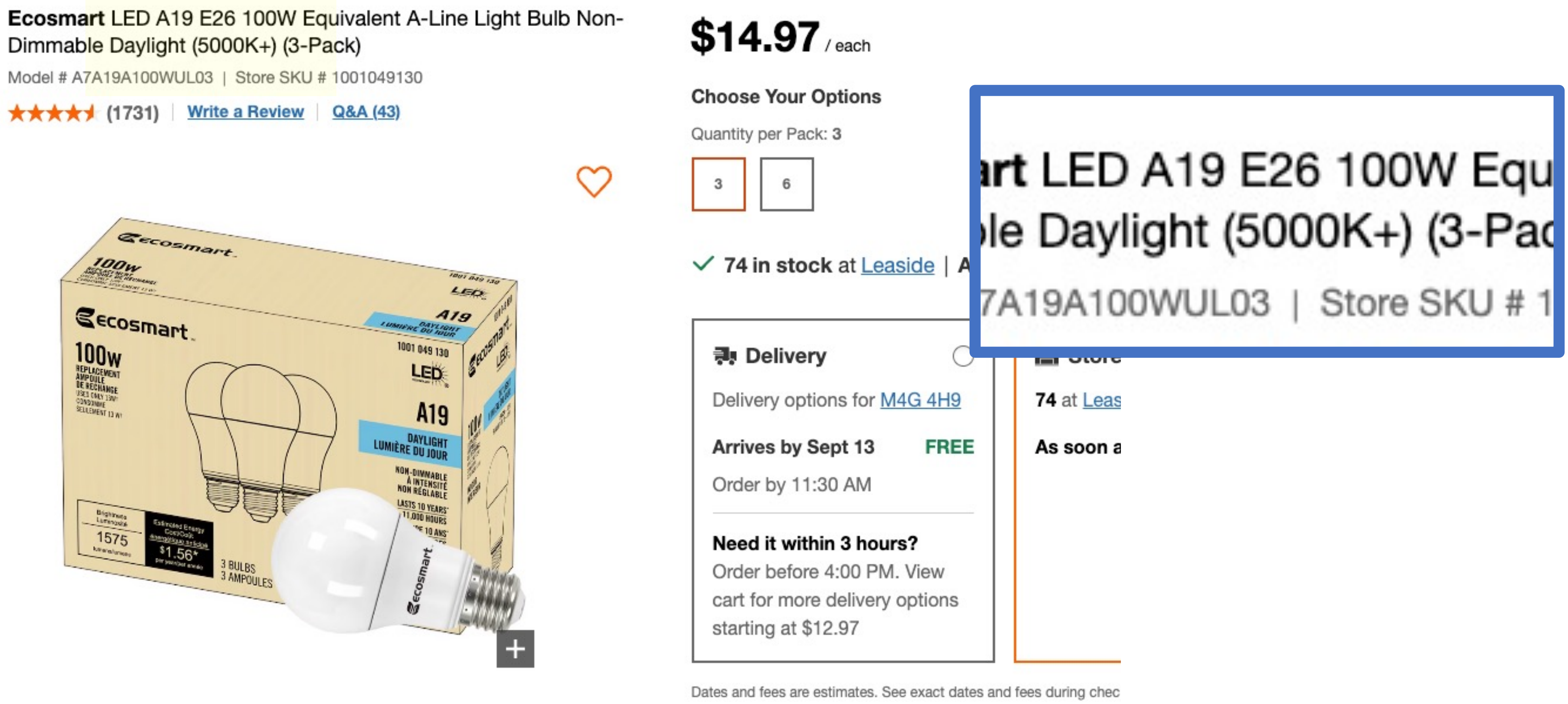
image white- balanced to daylight

White balancing presets

- Cameras nowadays come with a large number of presets: You can select which light you are taking images under, and the appropriate white balancing is applied.

WB SETTINGS	COLOR TEMPERATURE	LIGHT SOURCES
	10000 - 15000 K	Clear Blue Sky
	6500 - 8000 K	Cloudy Sky / Shade
	6000 - 7000 K	Noon Sunlight
	5500 - 6500 K	Average Daylight
	5000 - 5500 K	Electronic Flash
	4000 - 5000 K	Fluorescent Light
	3000 - 4000 K	Early AM / Late PM
	2500 - 3000 K	Domestic Lightning
	1000 - 2000 K	Candle Flame

What type of lighting do we have?



Manual vs. automatic white balancing

- Manual
  - Select a camera preset based on lighting.
  - Manually select object in photograph that is color-neutral and use it to normalize.

- Automatic
  - Grey world assumption: force average color of scene to be grey.
  - White world assumption: force brightest object in scene to be white.
  - Sophisticated histogram-based algorithms (what most modern cameras do).

Automatic white balancing

- Grey world assumption

$$R_{ave} = G_{ave} = B_{ave}$$

- Fix the intensities

white-balanced RGB $\longrightarrow$

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

$\longleftarrow$ sensor RGB

- Doesn't hold for all scenes
  - Dominant colors
  - Lighting effects

Automatic white balancing

- "White world assumption" assumes that there should be some white object present in the scene

white-balanced RGB → $$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} G_{max}/R_{max} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & G_{max}/B_{max} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$ ← sensor RGB

- Doesn't hold for scenes containing very bright colored lights

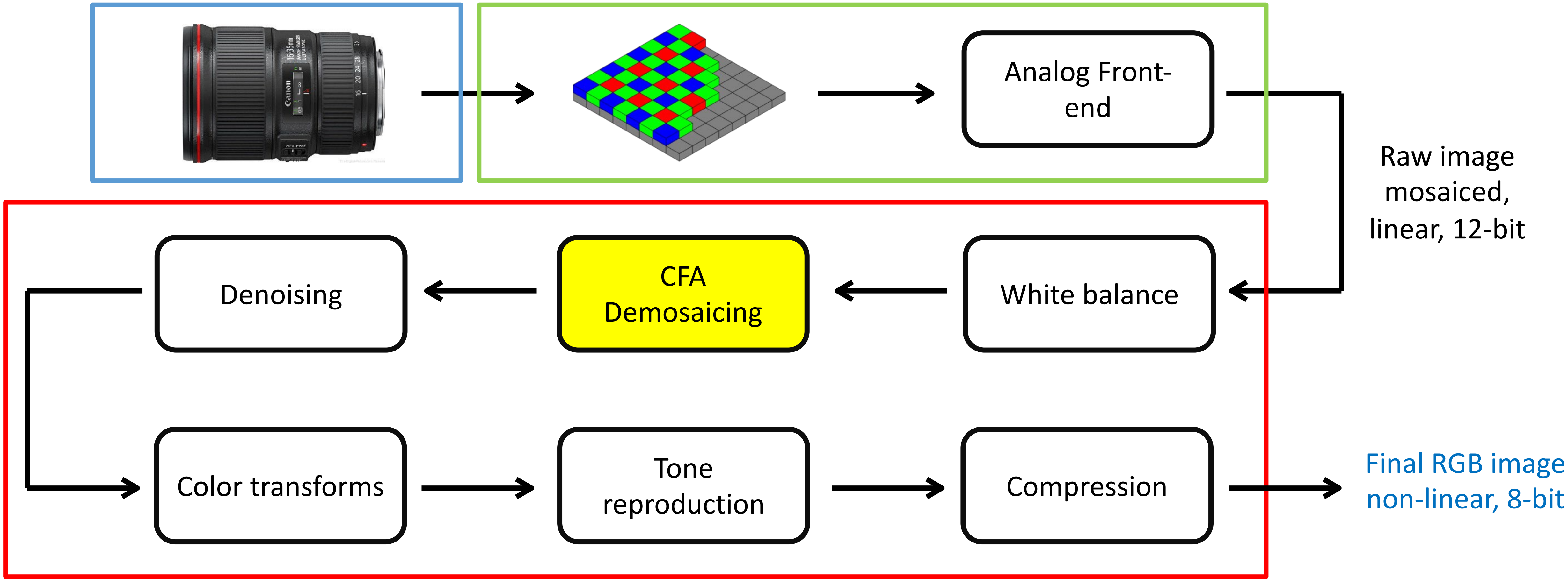
Automatic white balancing example

input image

grey world

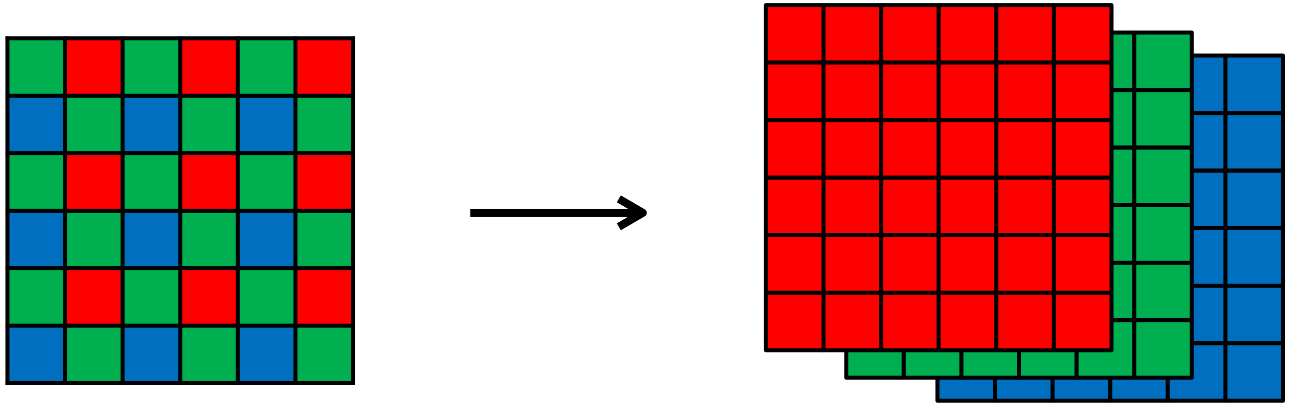
white world

The (in-camera) image processing pipeline



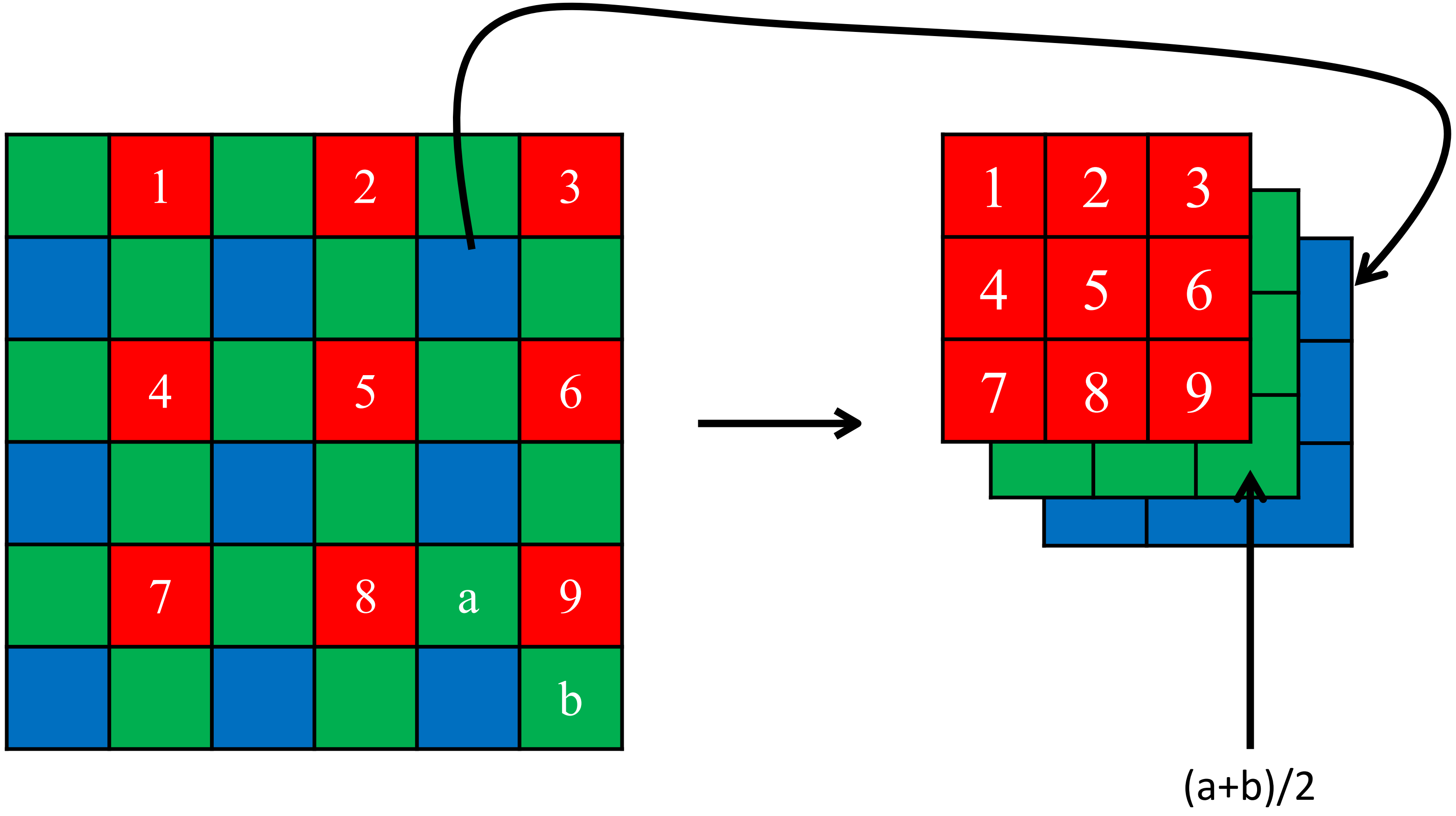
CFA demosaicing

- Produce full RGB image from mosaiced sensor output.



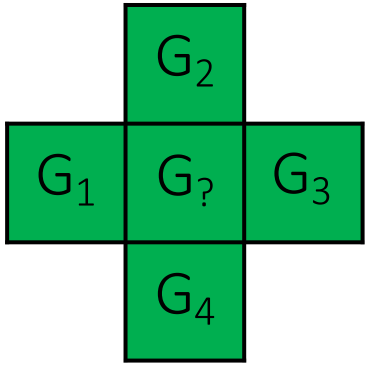
- Interpolate from neighbors
  - Bilinear interpolation (needs 4 neighbors)
  - Bicubic interpolation (needs more neighbors, may overblur).
  - Edge-aware interpolation (more on this later).

Demosaicing via Downsampling



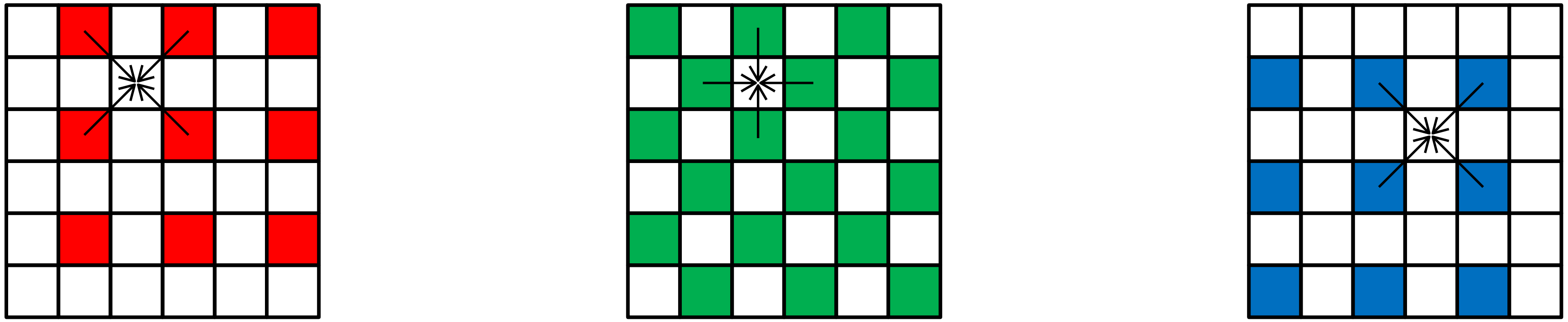
Demosaicing by bilinear interpolation

Bilinear interpolation: Simply average your 4 neighbors.

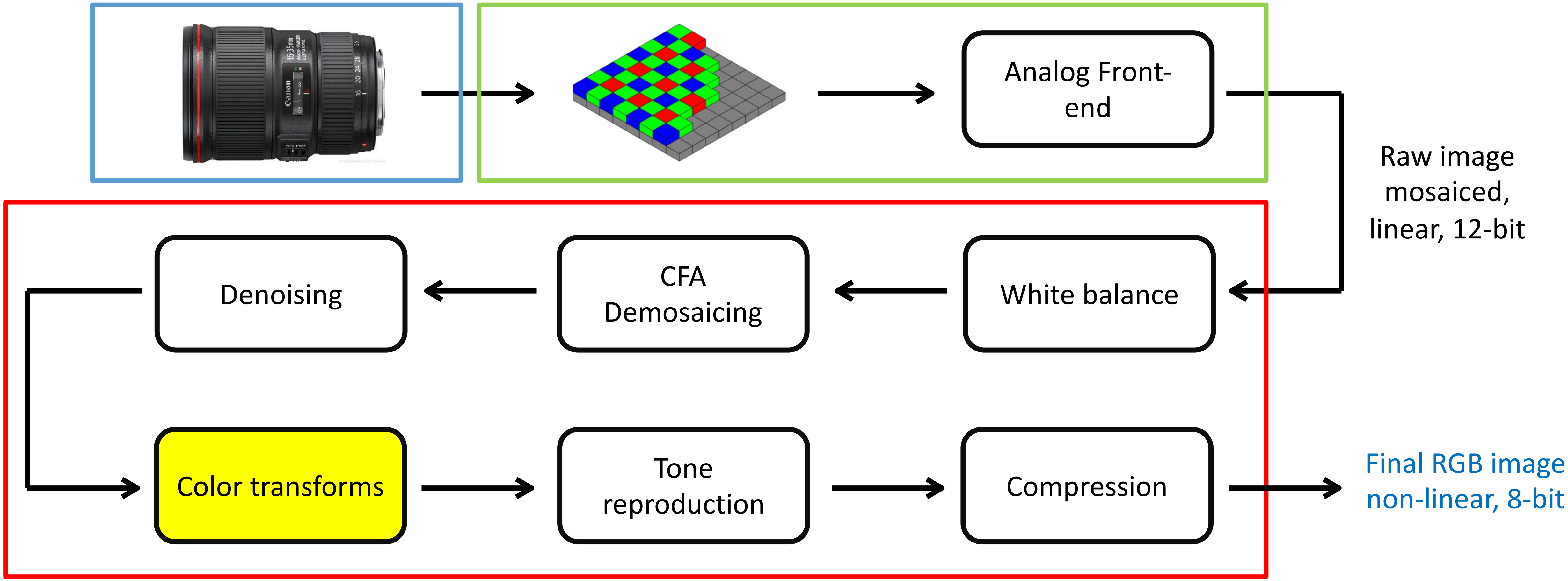


$$G_? = \frac{G_1 + G_2 + G_3 + G_4}{4}$$

Neighborhood changes for different channels:



The (in-camera) image processing pipeline



Color Transform

- Cameras often offer means to convert raw sensor readings to CIE XYZ Color Space.
- This color space was introduced by the CIE in 1931
- The name "XYZ" refers to the three tristimulus values that describe how the human eye responds to light of different wavelengths. Essentially, X, Y, and Z are values that represent the response of an average human observer to light across the visible spectrum.
- The XYZ color space encompasses all the colors that the average human can see, making it a device-independent color space. This means it's not tied to the color capabilities of any particular device, such as a camera, monitor, or printer.
- Because XYZ values can be somewhat unintuitive, they are often normalized into chromaticity values, denoted as x and y, which are more commonly used to represent color in many applications. This chromaticity representation is two-dimensional, making it easier to visualize.
- We can then transform from XYZ to an appropriate RGB color space for display purposes.

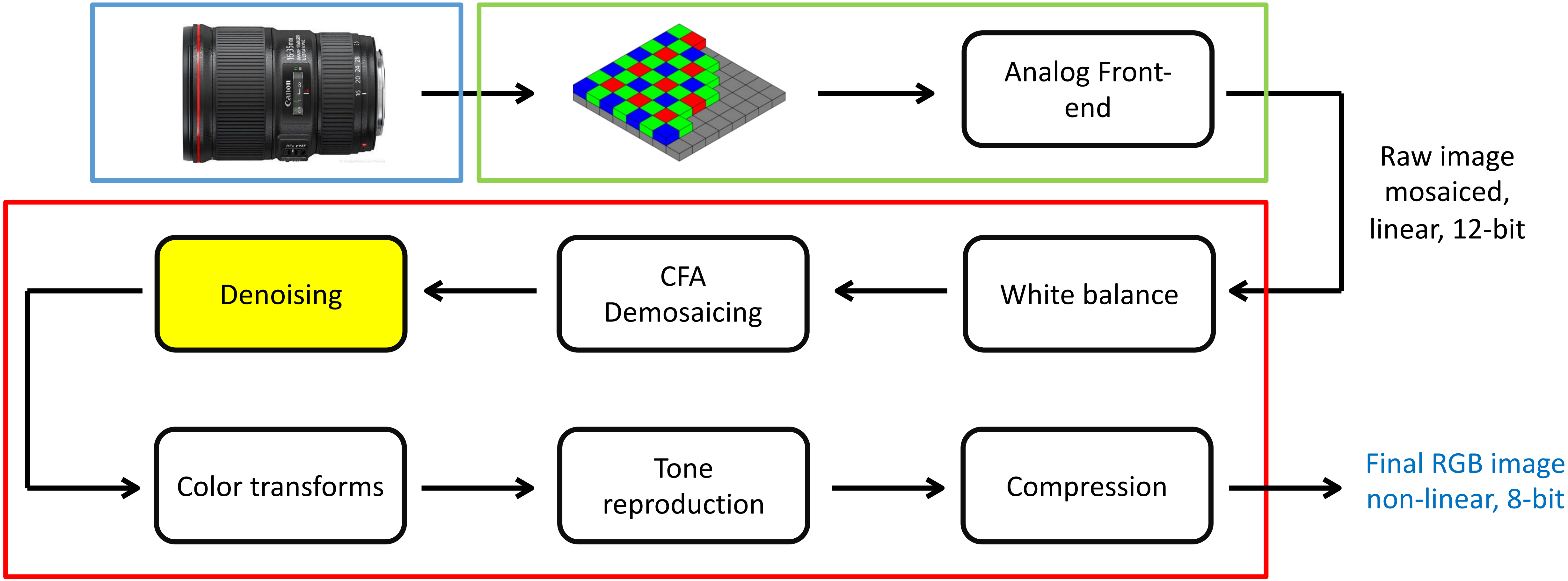
Color Transform

- The following matrix converts from sRGB to XYZ

```
sRGB_to_XYZ = np.array([[0.4124564, 0.3575761, 0.1804375],
                        [0.2126729, 0.7151522, 0.0721750],
                        [0.0193339, 0.1191920, 0.9503041]], dtype=np.double)
```

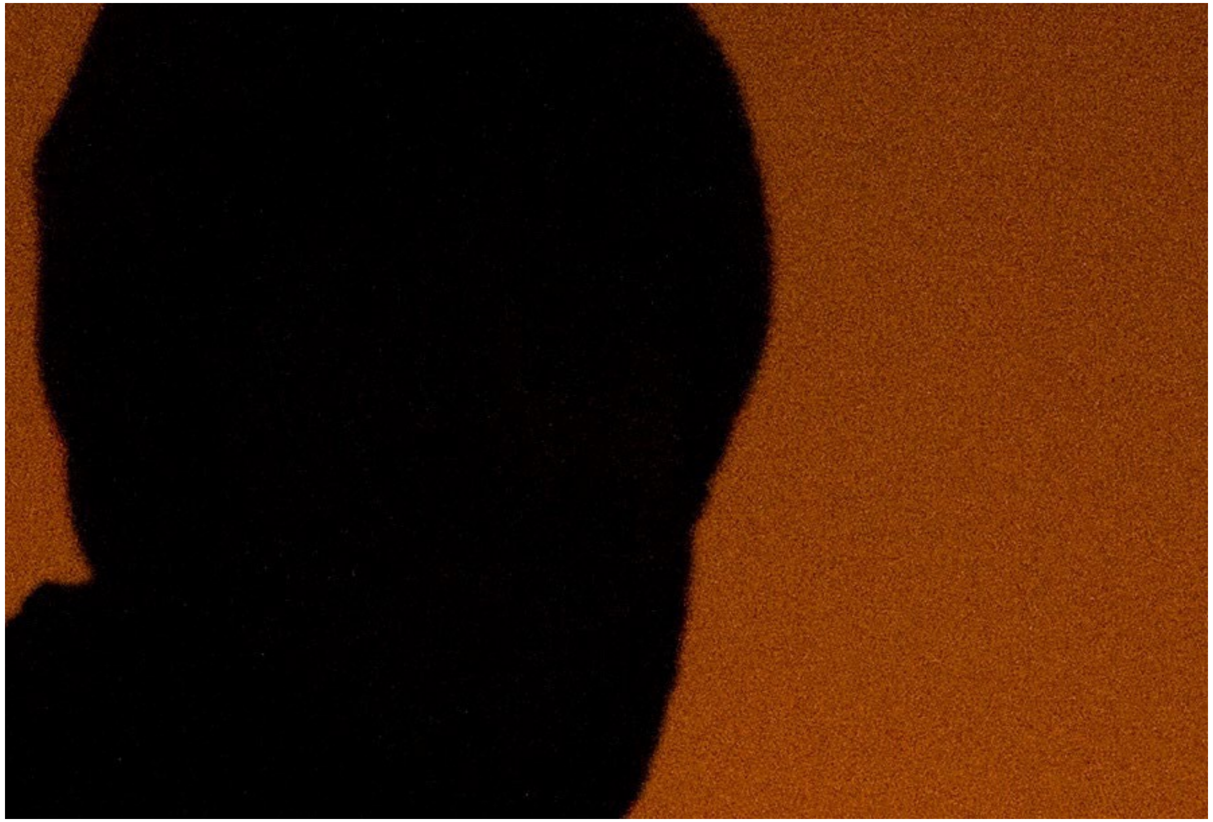
- How do use this matrix to go from XYZ to sRGB color space?
  - Use matrix inversion

The (in-camera) image processing pipeline



Noise in images

- Can be very pronounced in low-light images.

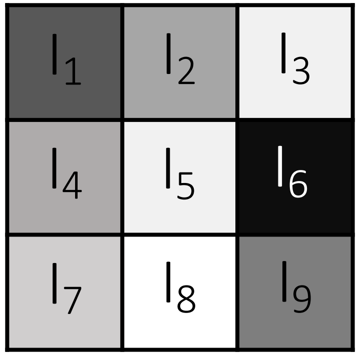


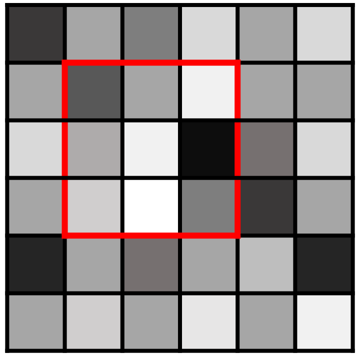
Three types of sensor noise

- (Photon) shot noise
  - Photon arrival rates are a random process (Poisson distribution).
  - The brighter the scene, the larger the variance of the distribution.
- Dark-shot noise
  - Emitted electrons due to thermal activity (becomes worse as sensor gets hotter.)
- Read noise
  - Caused by read-out and AFE electronics (e.g., gain, A/D converter).
  - Bright scene and large pixels: photon shot noise is the main noise source.

How to denoise?

Look at the neighborhood around you.





- Mean filtering (take average):

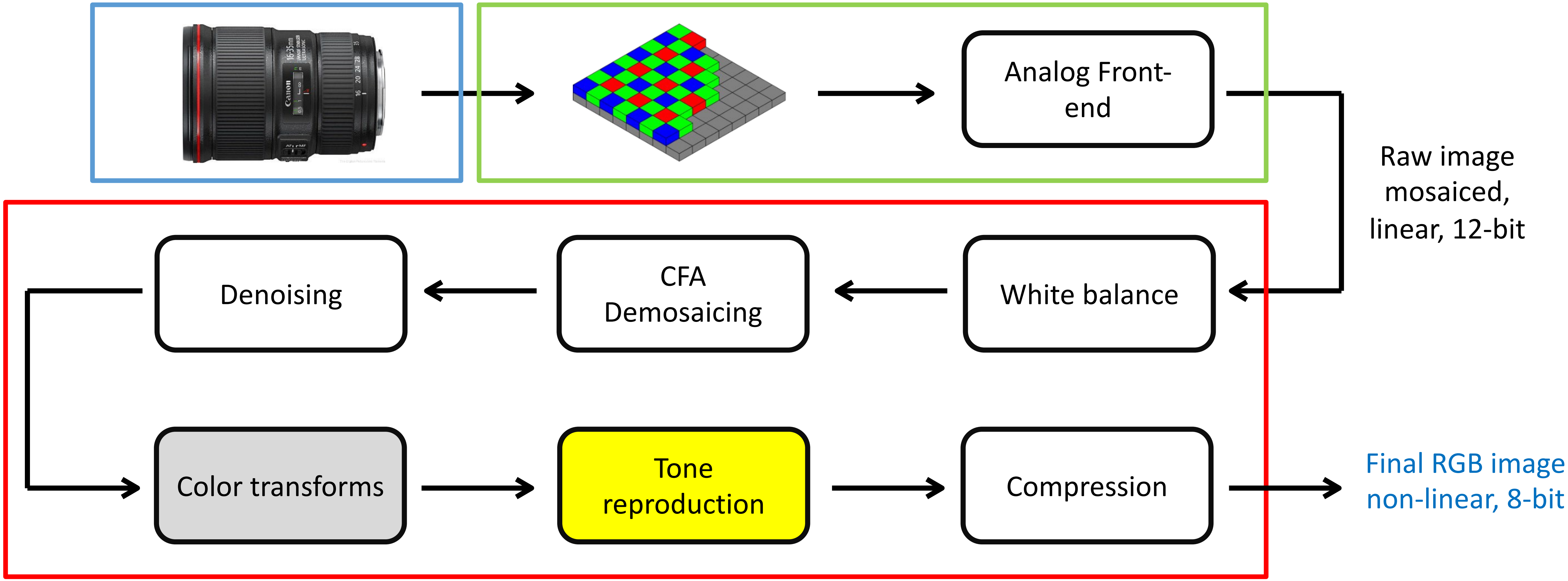
$$I'_5 = \frac{I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7 + I_8 + I_9}{9}$$

- Median filtering (take median):

$$I'_5 = \text{median}(I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8, I_9)$$

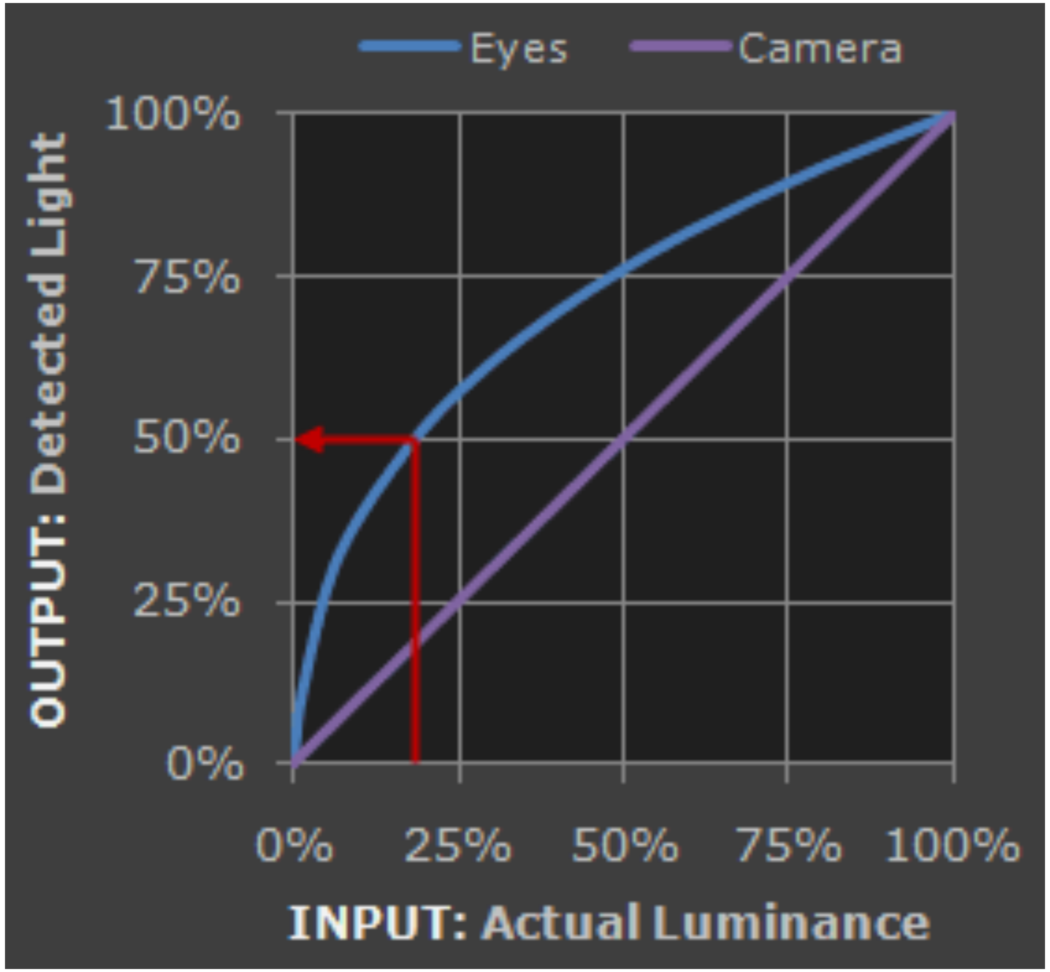
Week 2 – Lecture 1

The (in-camera) image processing pipeline



Perceived vs. Measured Brightness in Human Eye

- Sensor response is *mostly* linear
- Human-eye response (measured brightness) is also linear
- Human-eye perception (perceived brightness) is non-linear
  - More sensitive to darker tones
  - Approximately equal to the square-root function



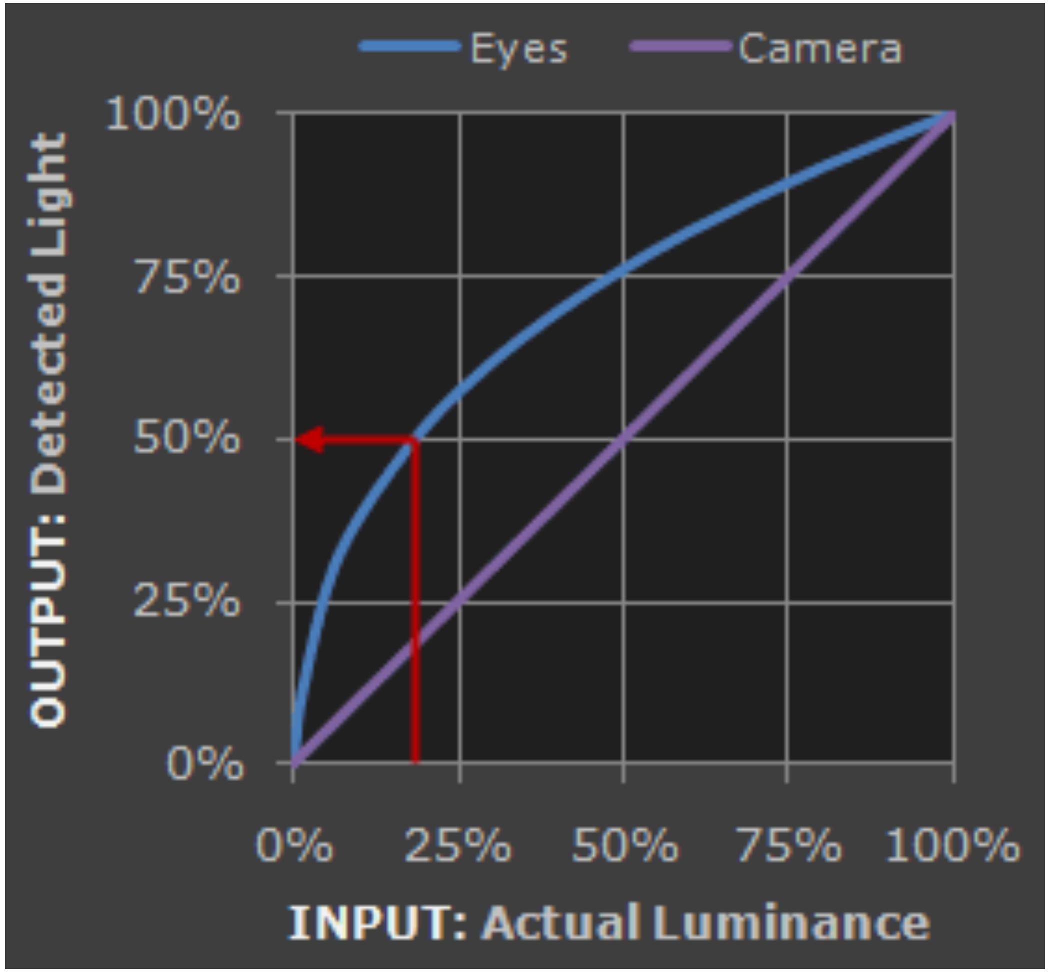
Perceived vs. Measured Brightness in Human Eye

- Gamma curve

$$L = Y^{\gamma}$$

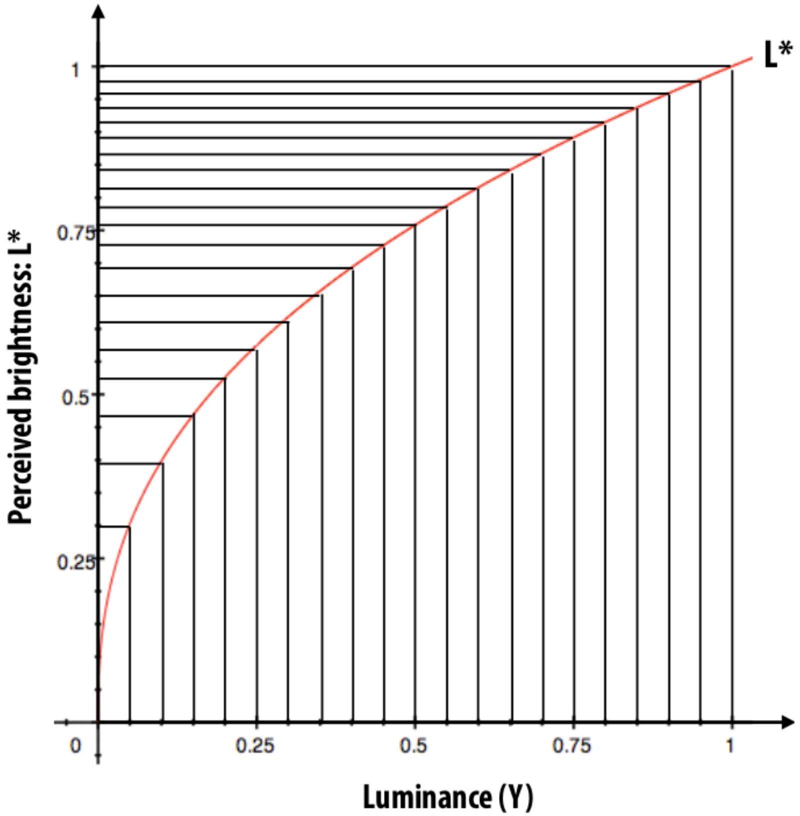
Perceived intensity — Intensity — Gamma value

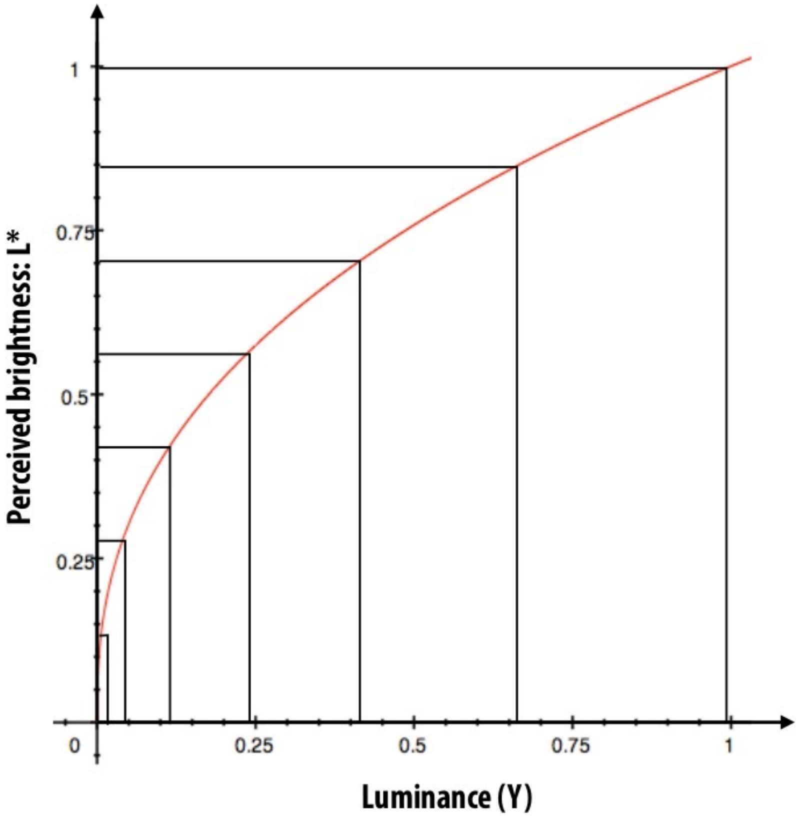
- Typically, $\gamma = \frac{1}{2.2}$



Gamma encoding

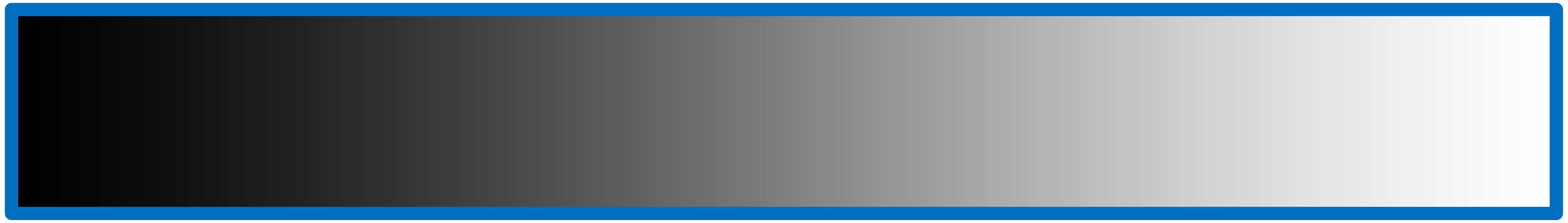
- After this stage, we perform compression, which includes changing from 12 to 8 bits.
- Apply non-linear curve to use available bits to better encode the information human vision is more sensitive to.





Demonstration

original (8-bits, 256 tones)



Can you predict what will happen if we linearly encode this tone range with only 5 bits?

Can you predict what will happen if we gamma encode this tone range with only 5 bits?

Demonstration

original (8-bits, 256 tones)

linear encoding (5-bits, 32 tones)

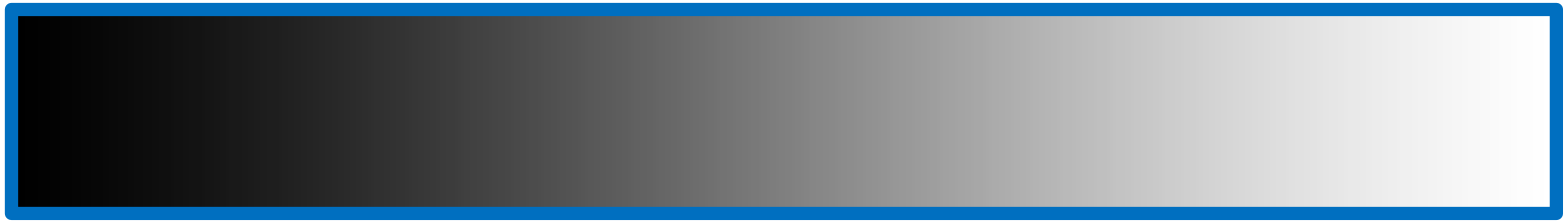
all of this range gets mapped to just one tone

all of these tones look the same

Can you predict what will happen if we gamma encode this tone range with only 5 bits?

Demonstration

original (8-bits, 256 tones)



linear encoding (5-bits, 32 tones)

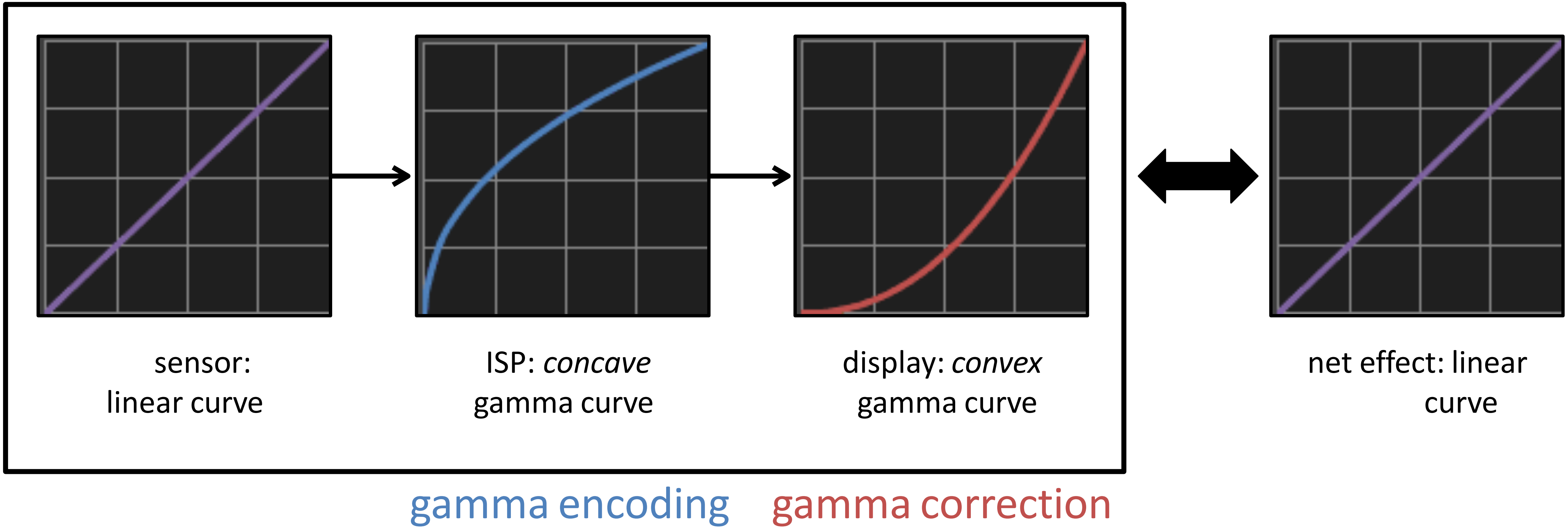
all of this range gets mapped to just one tone

all of these tones look the same

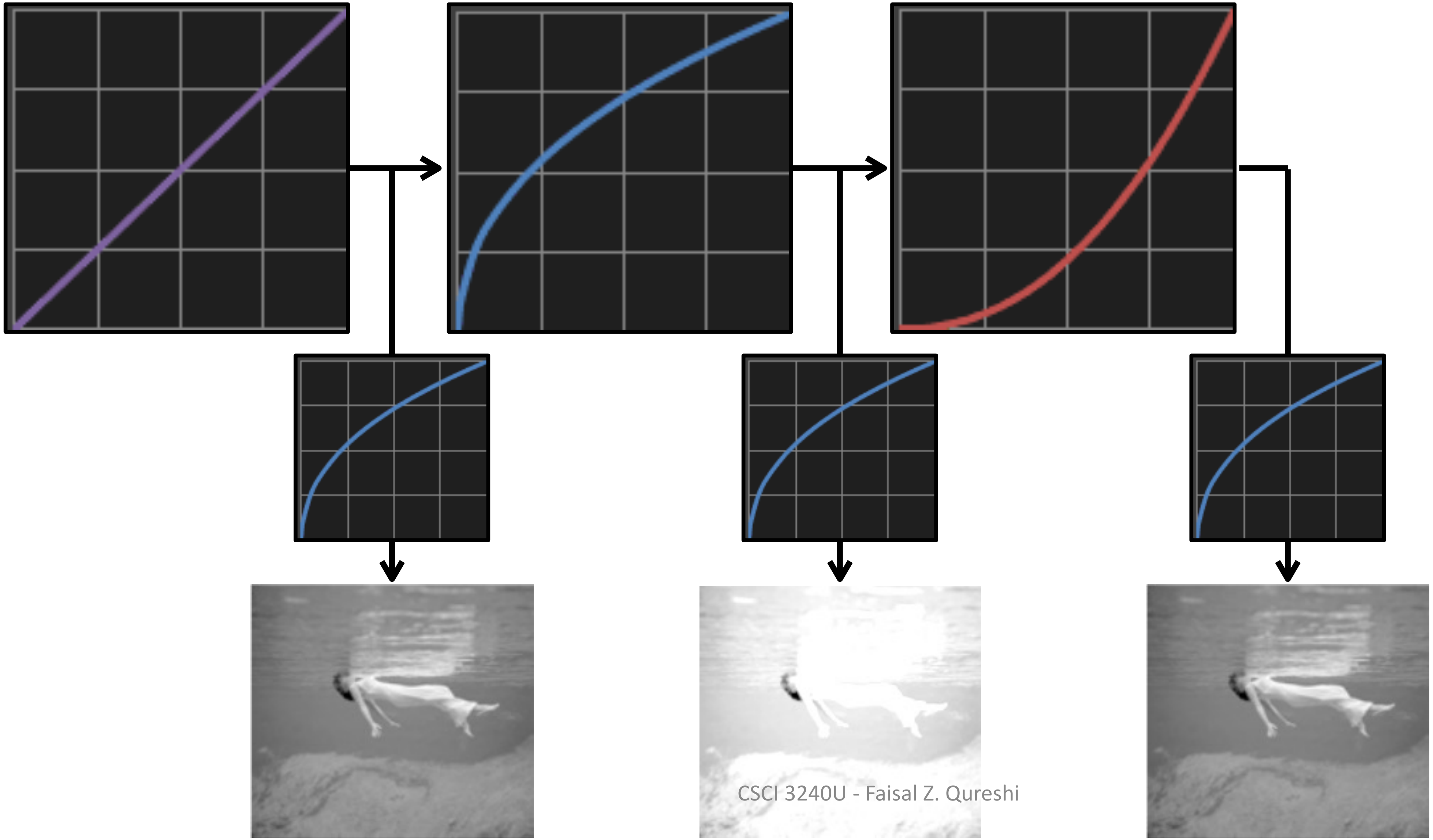
gamma encoding (5-bits, 32 tones)

tone encoding becomes a lot more perceptually uniform

Tone reproduction pipeline



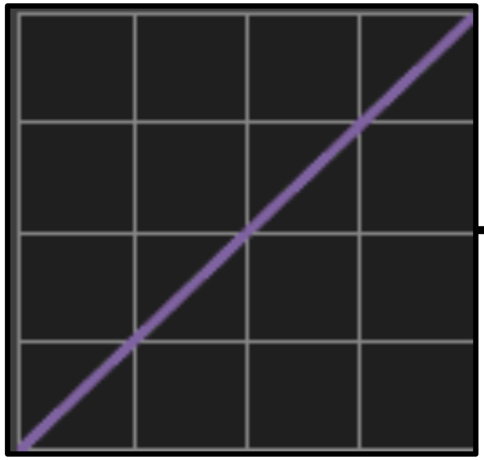
Tone reproduction pipeline



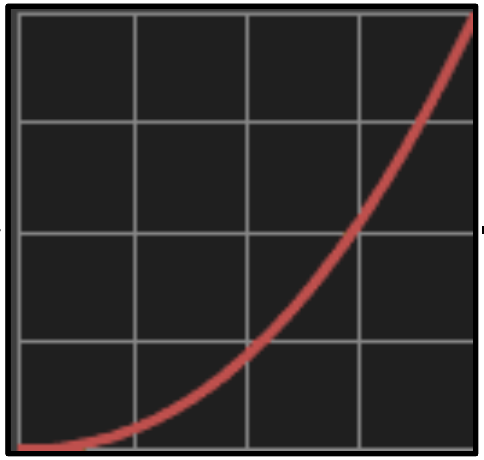
human visual system: *concave* gamma curve

image a human would see at different stages of the pipeline

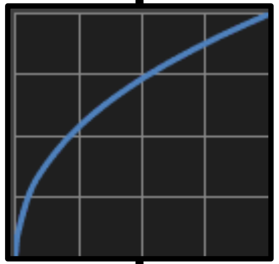
RAW pipeline

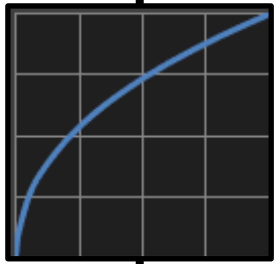


gamma encoding is skipped!



display still applies gamma correction!





human visual system: *concave* gamma curve

RAW image appears very dark! (Unless you are using a RAW viewer)

image a human would see at different stages of the pipeline

Historical note

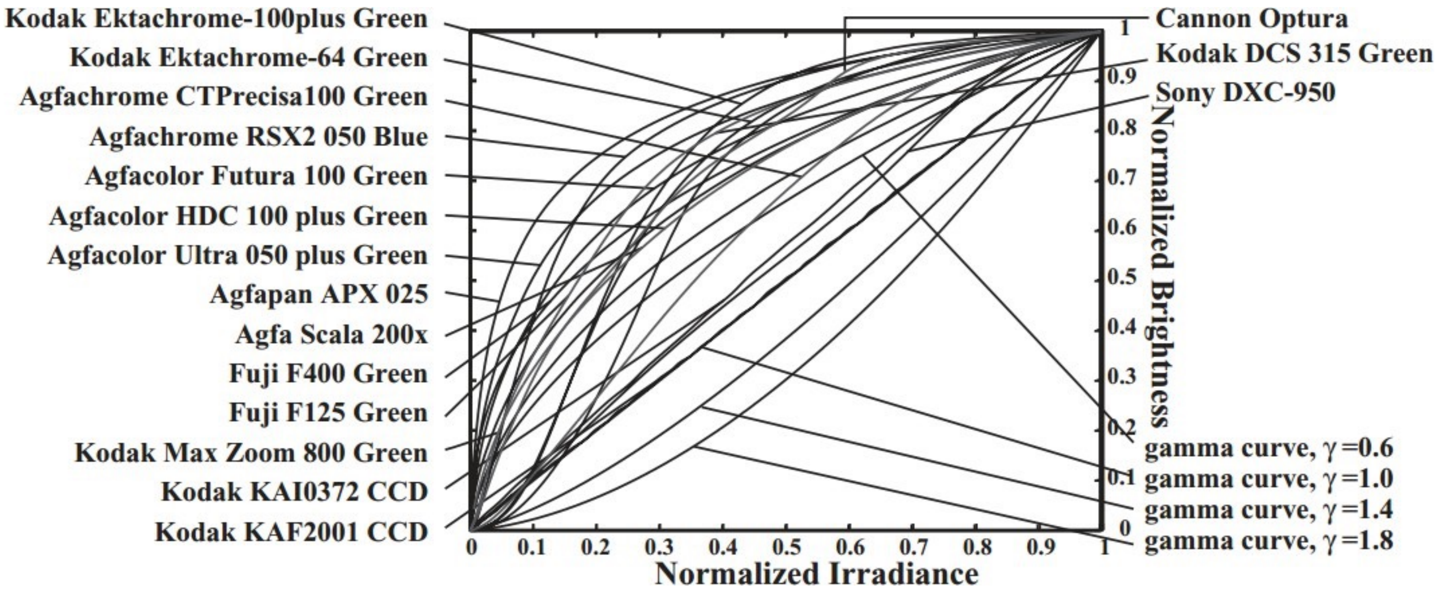
- CRT displays used to have a response curve that was (almost) exactly equal to the inverse of the human sensitivity curve. Therefore, displays could skip gamma correction and display directly the gamma-encoded images.
- It is sometimes mentioned that gamma encoding is done to undo the response curve of a display. This used to (?) be correct, but it is not true nowadays. Gamma encoding is performed to ensure a more perceptually-uniform use of the final image's 8 bits.

Gamma encoding curves

- The exact gamma encoding curve depends on the camera.
- Often well approximated as Lγ, for different values of the power γ ("gamma").
  - A good default is γ = 1/2.2 (approximately the square root).

before gamma

after gamma



Warning: Our values are no longer linear relative to scene radiance!

Do I ever need to use RAW?

- Emphatic yes!
- Every time you use a physics-based computer vision algorithm, you need linear measurements of radiance.
  - Examples: photometric stereo, shape from shading, image-based relighting, illumination estimation, anything to do with light transport and inverse rendering, etc.
  - Applying the algorithms on non-linear (i.e., not RAW) images will produce completely invalid results.
- Even when physics is not your concern, you may want to still use raw
  - If you like re-finishing your photos (e.g., on Photoshop), RAW makes your life much easier and your edits much more flexible.

Are there any downsides to using RAW?

- Image files are *a lot* bigger.
  - You burn through multiple memory cards.
  - Your camera will buffer more often when shooting in burst mode.
  - Your computer needs to have sufficient memory to process RAW images.

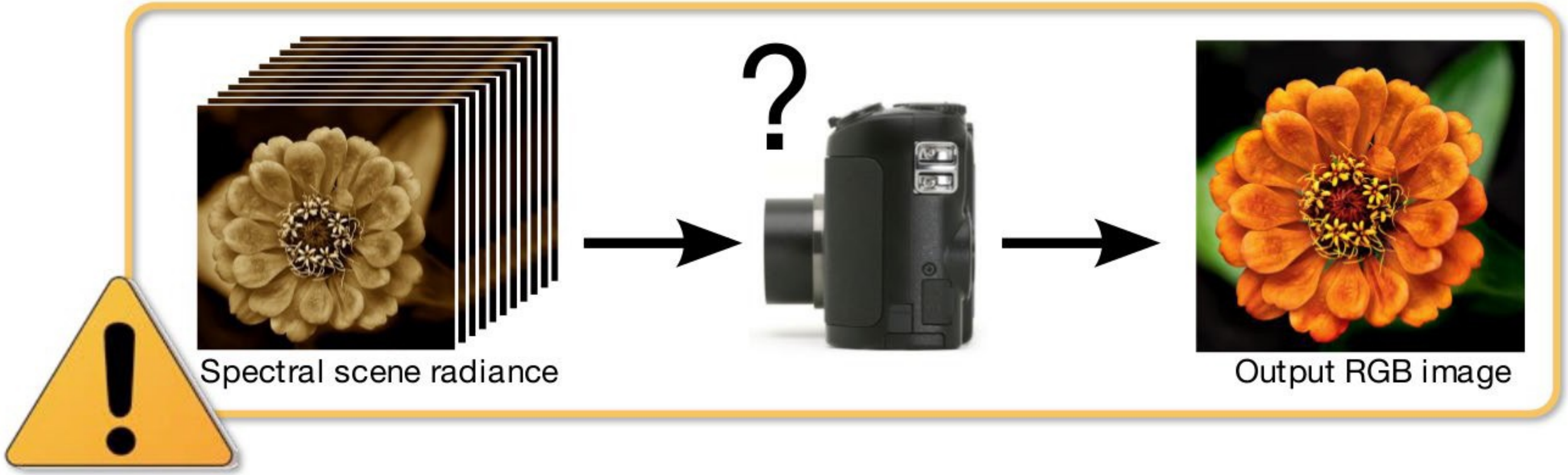
Is it even possible to get access to RAW images?

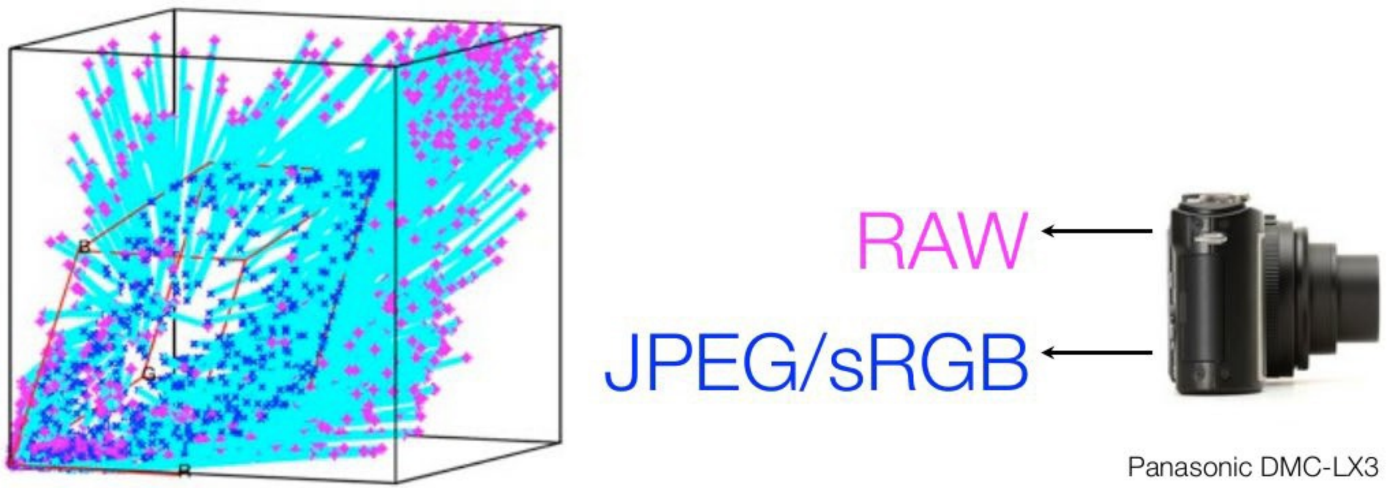
- Quite often yes!
- Most high-end cameras provide an option to store RAW image files
- Certain phone cameras allow, directly or indirectly, access to RAW.
- Sometimes, it may not be “fully” RAW.
  - The Lightroom app provides images after demosaicking but before tone reproduction.

I forgot to set my camera to RAW, can I still get the RAW file?

- Nope, tough luck.
- The image processing pipeline is lossy: After all the steps, information about the original image is lost.
  - Sometimes we may be able to reverse a camera's image processing pipeline *if we know exactly what it does* (e.g., by using information from other similar RAW images).
  - The conversion of PNG/JPG back to RAW is known as "derendering" and is an active research area.

Derendering





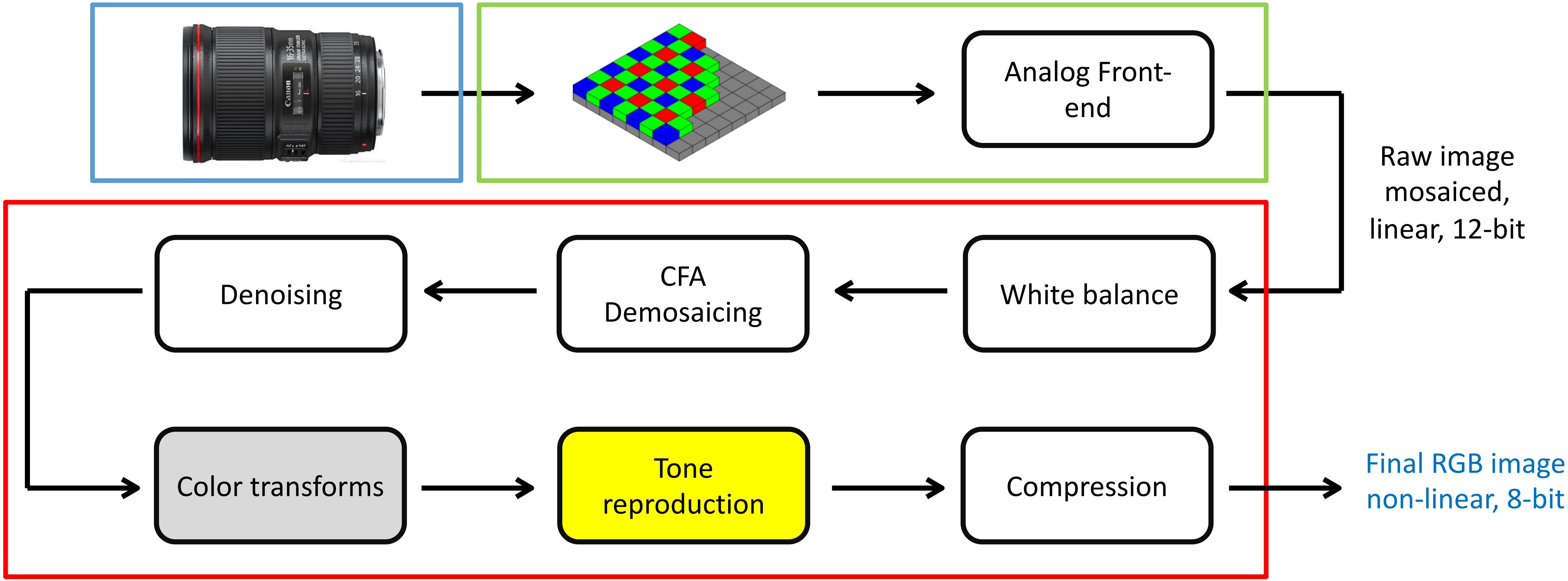
Important to remember

- What I described today is an "idealized" version of what we *think* commercial cameras do.
  - Almost all of the steps in both the sensor and image processing pipeline I described earlier are camera-dependent.
  - Even if we know the basic steps, the implementation details are proprietary information that companies actively try to keep secret.
  - I will go back to a few of my slides to show you examples of the above.

Summary

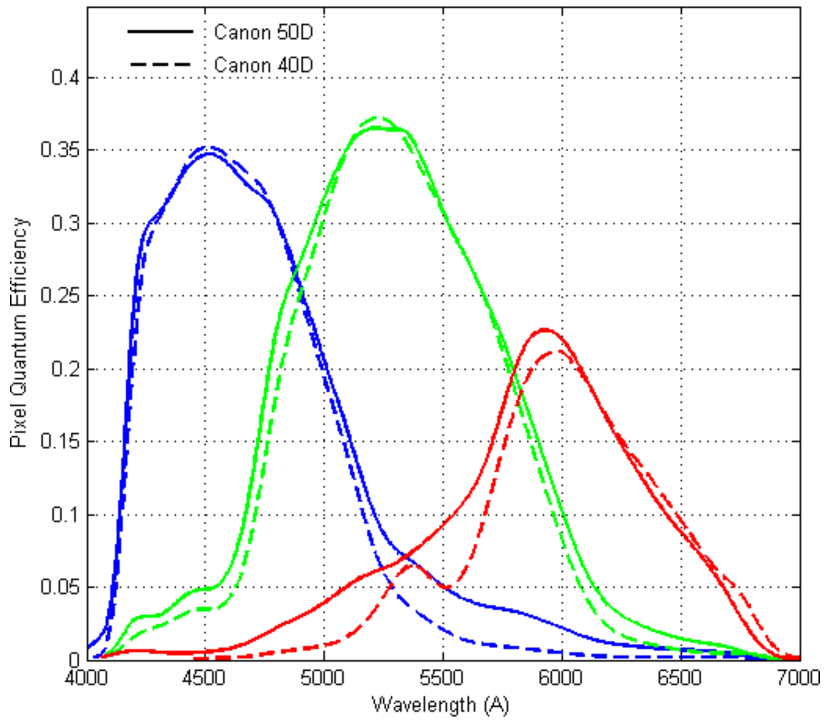
- Image processing pipeline
  - Whitebalancing
  - Demoisacing
  - Denoising
  - Gamma correction

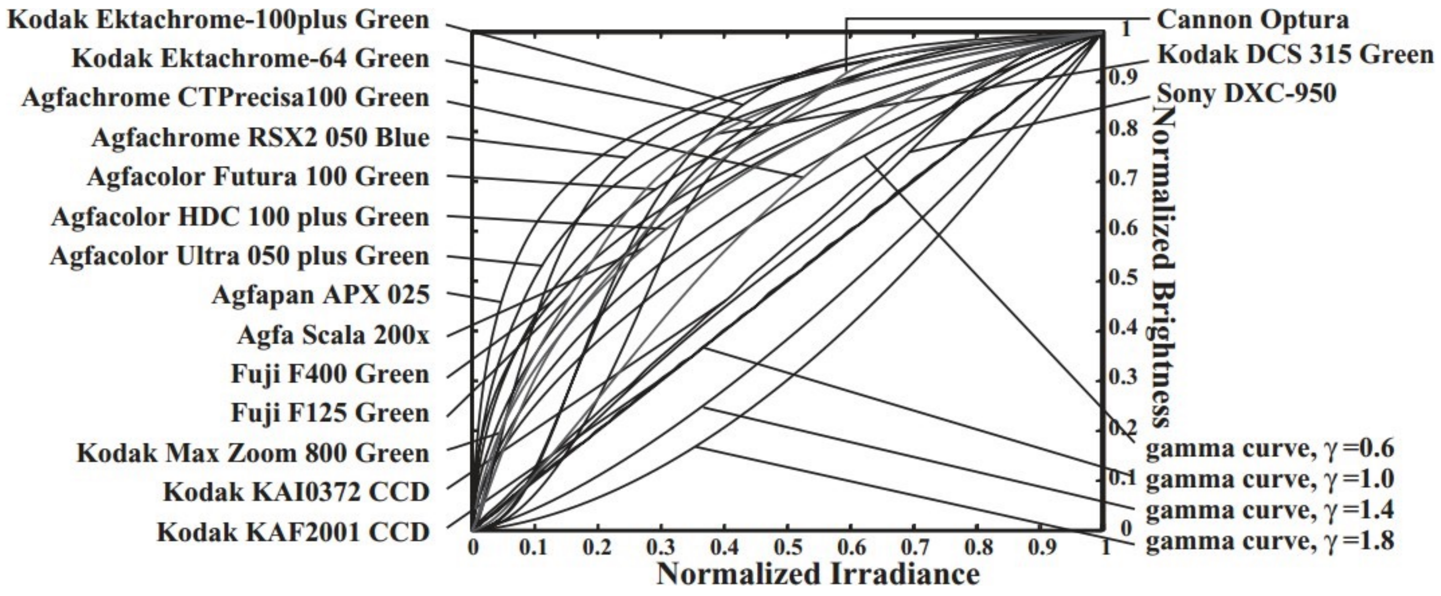
The (hypothetical) image processing pipeline



The Many Sensitivity Curves

- All of these sensitivity curves are different from camera to camera and kept secret.





Serious inhibition for research

- Very difficult to get access to ground-truth data at intermediate stages of the pipeline.
- Very difficult to evaluate effect of new algorithms for specific pipeline stages.

How do we open a raw file in Python?

- Python `rawpy` package

Is this the best image processing pipeline?

- It depends on how you define “best.” This definition is task-dependent.
- The standard image processing pipeline is designed to create “nice-looking” images.
- If you want to do physics-based vision, the best image processing pipeline is no pipeline at all (use RAW).
- What if you want to use images for, e.g., object recognition? Tracking? Robotics SLAM? Face identification? Forensics?
- Developing task-adaptive image processing pipelines is an active area of research.

Take Home Message

- The values of pixels in a photograph and the values output by your camera's sensor are two very different things.
- The relationship between the two is complicated and unknown.