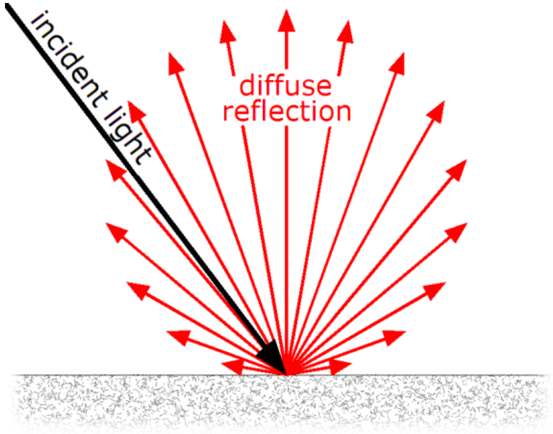
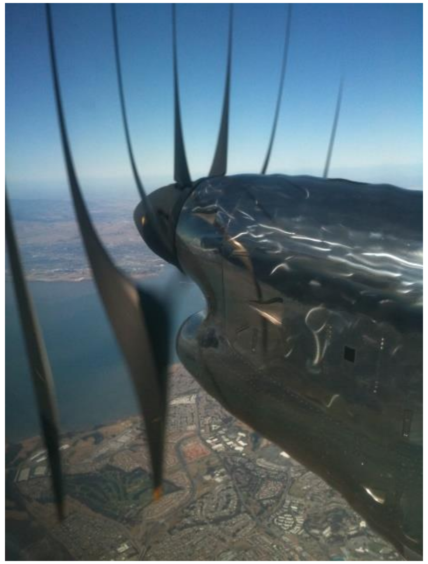

S BR 8167

Recap of Light and Cameras

- Light transport is complex, but in computer vision we often assume their interactions follow the Lambertian reflectance model.



- Time, Intensity, and Space are all continuous, but they are all *discretized* when capturing images. Sometimes this leads to artifacts.

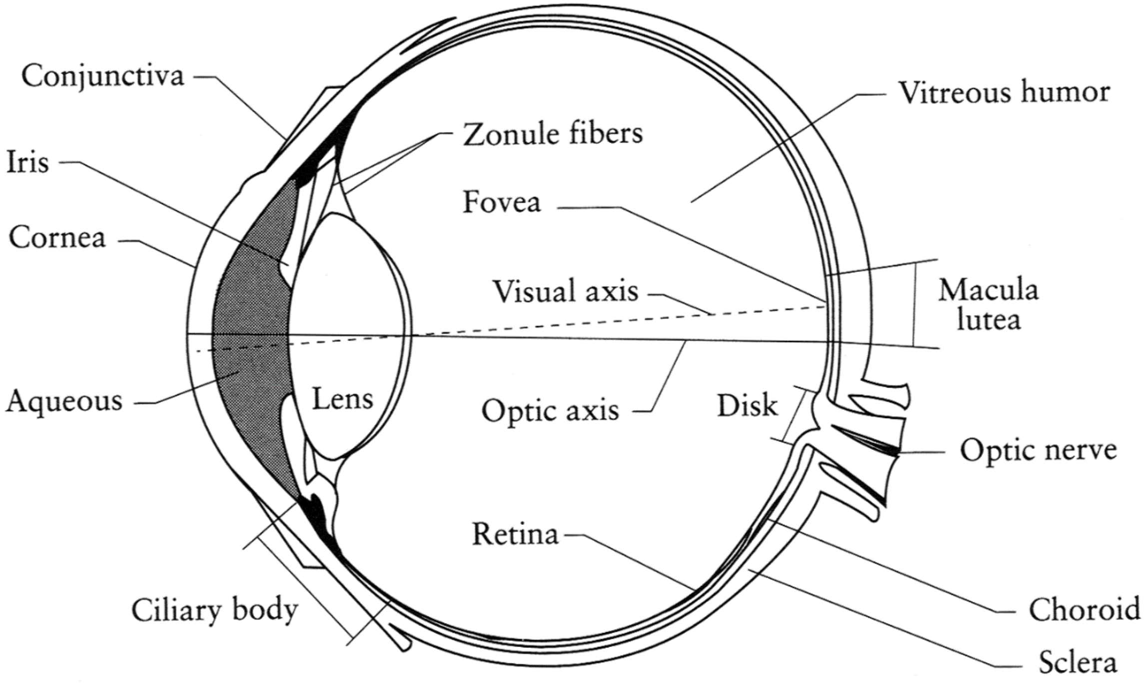


This lecture

- Eyes (your cameras)
- Light and Color

- Next lecture: Interest Points

The Eye



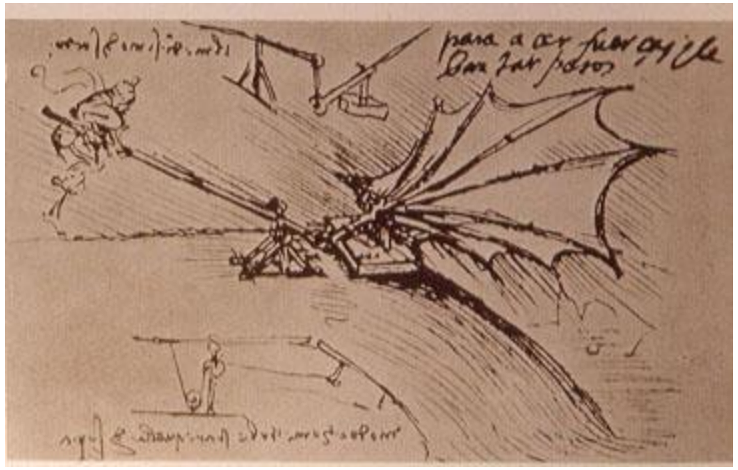
- The human eye is a camera!
  - **Iris** - colored annulus with radial muscles
  - **Pupil** - the hole (aperture) whose size is controlled by the iris
  - What's the "film"?
    - photoreceptor cells (rods and cones) in the **retina**

Slide by Steve Seitz

Why do we care about human vision in this class?

- We don’t, necessarily.

Ornithopters



Why do we care about human vision in this class?

- We don’t, necessarily.
- But cameras necessarily imitate the frequency response of the human eye, so we should know that much.
- Also, computer vision probably wouldn’t get as much scrutiny if biological vision (especially human vision) hadn’t proved that it was possible to make important judgements from 2D images.

Does computer vision "understand" images?

"Can machines fly?" The answer is yes, because airplanes fly.

"Can machines swim?" The answer is no, because submarines don't swim.

"Can machines think?" Is this question like the first, or like the second?

Source: Norvig

The Spectrum of Biological Inspiration

Radar

Lidar

Traditional cameras

Tesla autopilot

Humanoid and Quadruped robots

Less biologically inspired

More biologically inspired

Inside Tesla as Elon Musk Pushed an Unflinching Vision for Self-Driving Cars

The automaker may have undermined safety in designing its Autopilot driver-assistance system to fit its chief executive's vision, former employees say.

Hardware choices have also raised safety questions. Within Tesla, some argued for pairing cameras with radar and other sensors that worked better in heavy rain and snow, bright sunshine and other difficult conditions. For several years, Autopilot incorporated radar, and for a time Tesla worked on developing its own radar technology. But three people who worked on the project said Mr. Musk had repeatedly told members of the Autopilot team that humans could drive with only two eyes and that this meant cars should be able to drive with cameras alone.

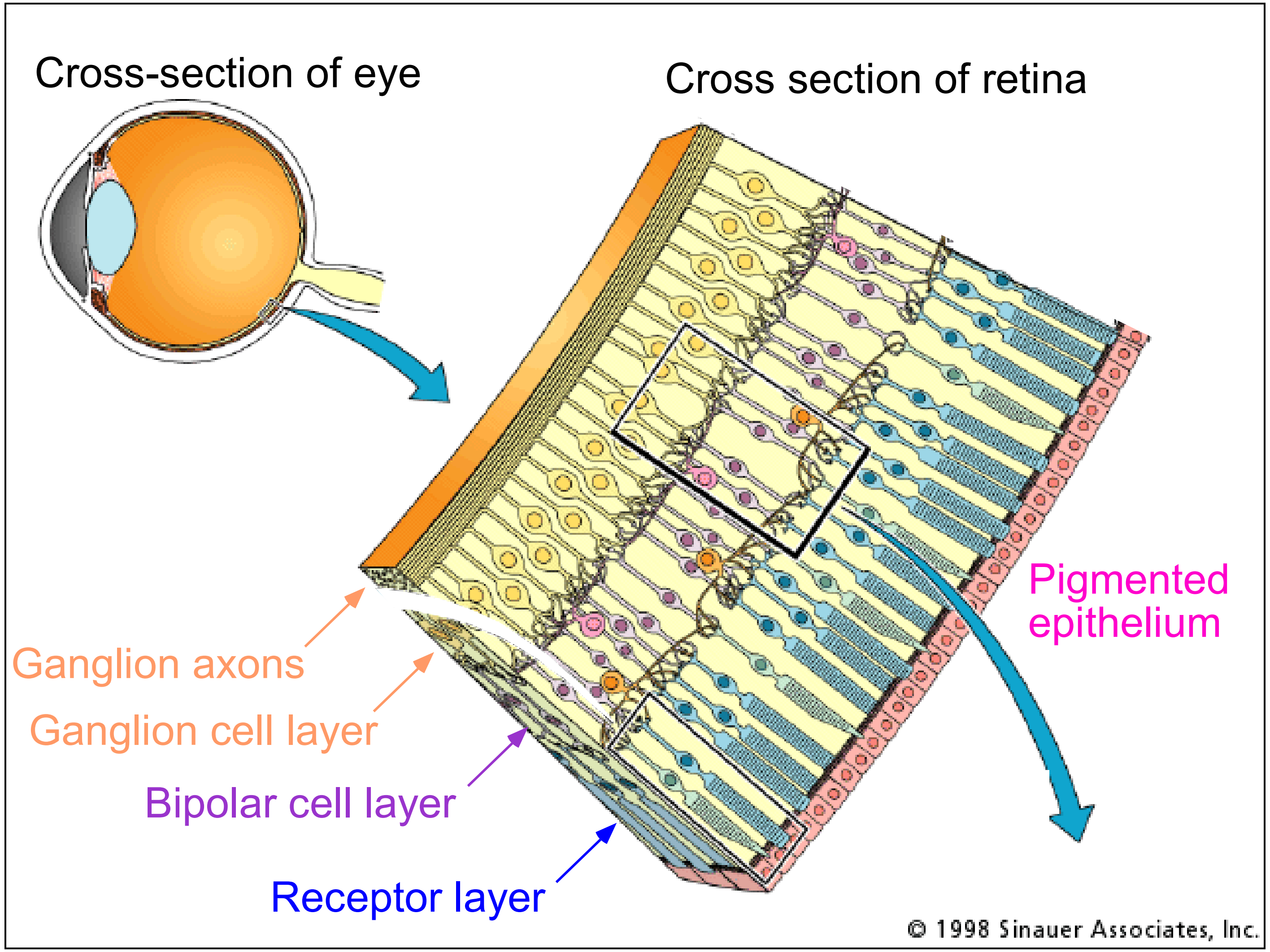
Earlier this week, Musk took a jab at Waymo and claimed that “lidar and radar reduce safety”:

> **Lidar and radar reduce safety due to sensor contention. If lidars/radars disagree with cameras, which one wins? This sensor ambiguity causes increased, not decreased, risk. That’s why Waymos can’t drive on highways.We turned off radars in Teslas to increase safety. Cameras ftw.**

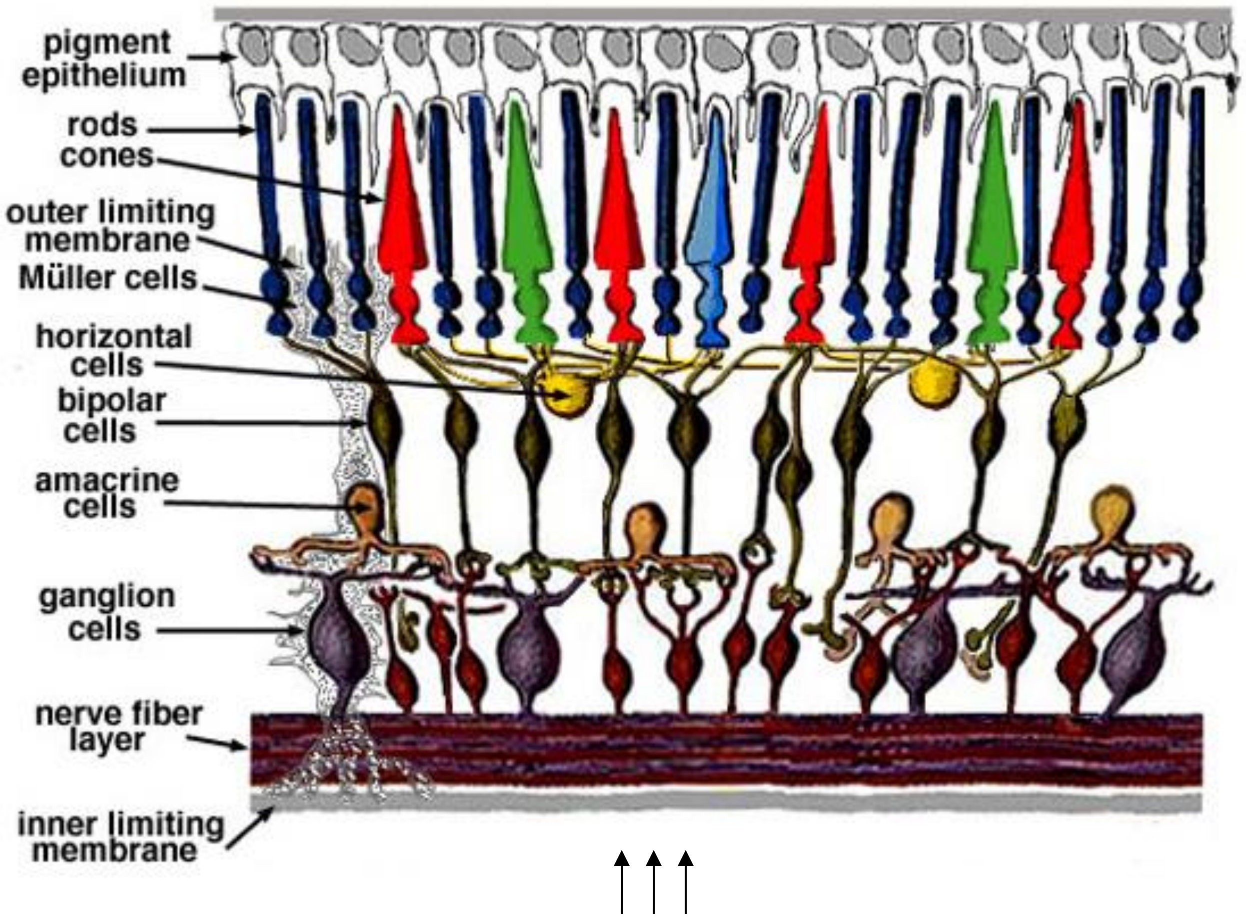
August 2025


TURN
ON RED
ONE
WAY
Spring St NW
FDC

The Retina


Cross-section of eye
Cross section of retina
Ganglion axons
Ganglion cell layer
Bipolar cell layer
Receptor layer
Pigmented epithelium
© 1998 Sinauer Associates, Inc.

Retina up-close



Light

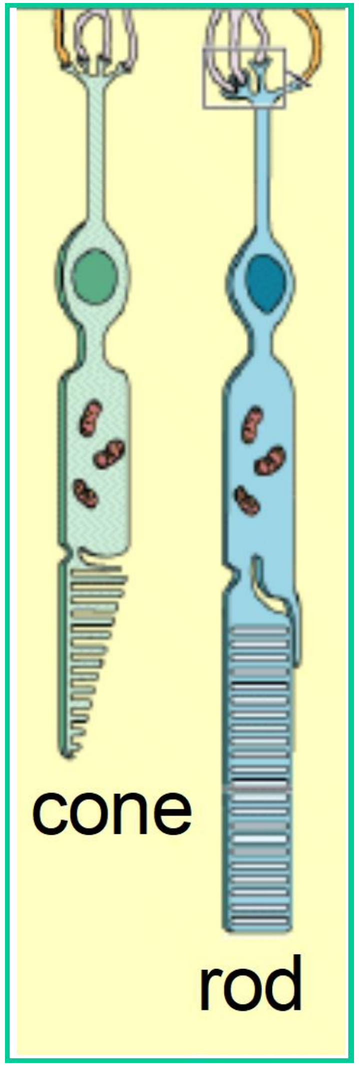
Two types of light-sensitive receptors

Cones

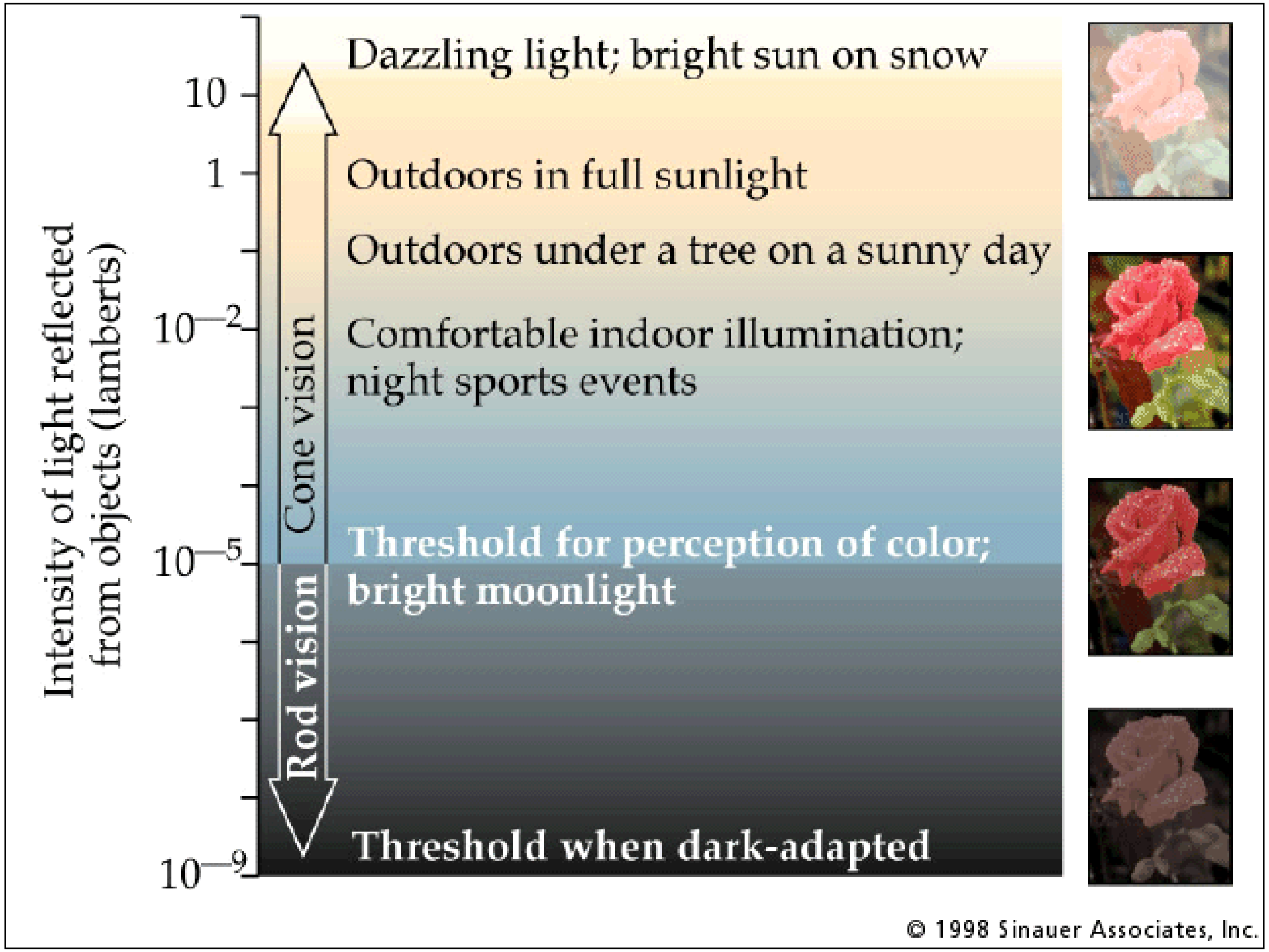
cone-shaped
less sensitive
operate in high light
color vision

Rods

rod-shaped
highly sensitive
operate at night
gray-scale vision



Rod / Cone sensitivity



Camera Gamma Curve

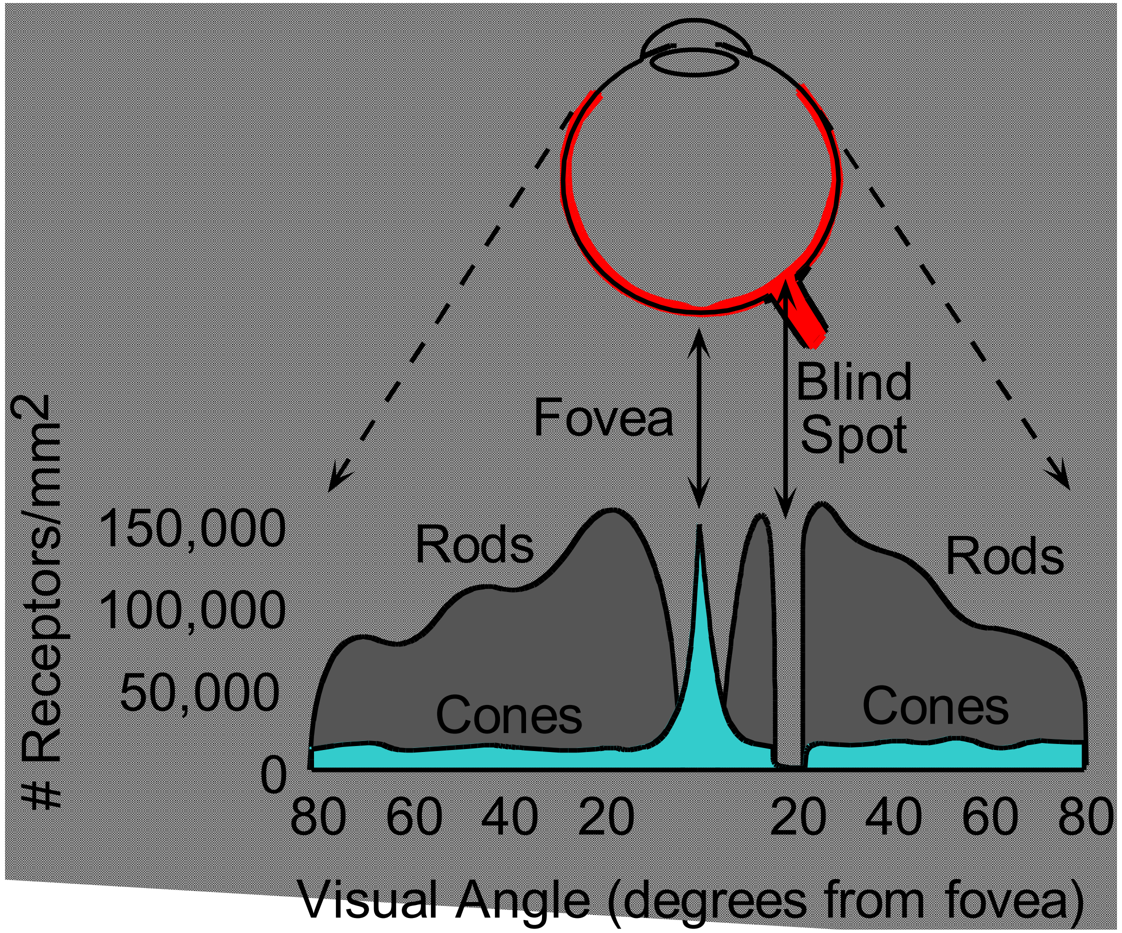
$$encoded\ brightness = measured\ brightness^{\gamma}$$

Typical encoding gamma value

In this class (e.g. for project 1), you are manipulating *encoded* brightness values. A pixel with value of “100” does not mean twice the light emissions as a pixel with value of “50”.

https://en.wikipedia.org/wiki/Gamma_correction

Distribution of Rods and Cones



Night Sky: why are there more stars off-center?

Averted vision: http://en.wikipedia.org/wiki/Averted_vision

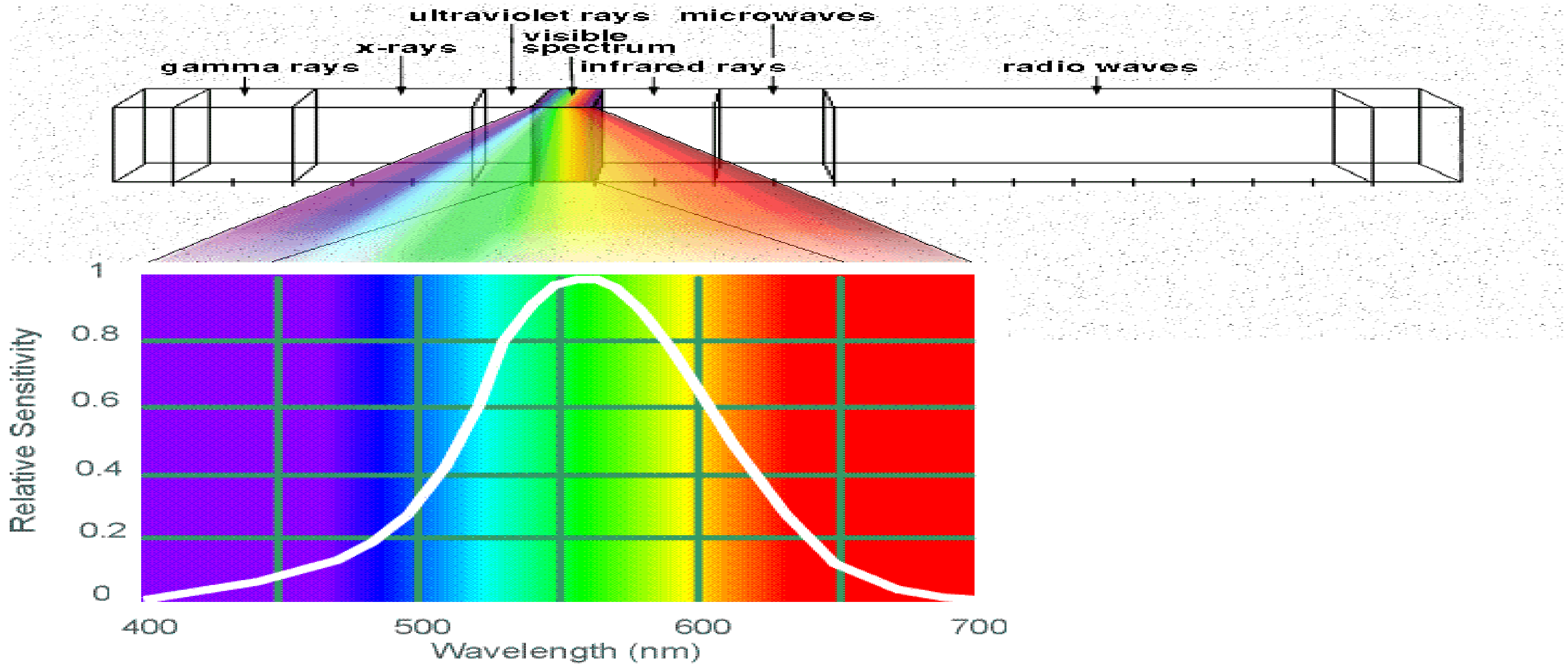
Eye Movements

- Saccades
- Can be consciously controlled. Related to perceptual attention.
- 200ms to initiation, 20 to 200ms to carry out. Large amplitude.
- Microsaccades
- Involuntary. Smaller amplitude. Especially evident during prolonged fixation. Function debated.
- Ocular microtremor (OMT)
- Involuntary. high frequency (up to 80Hz), small amplitude.
- Smooth pursuit – tracking an object

Slow mo guys – Saccades

- https://youtu.be/Fmg9ZOHESgQ?t=21s

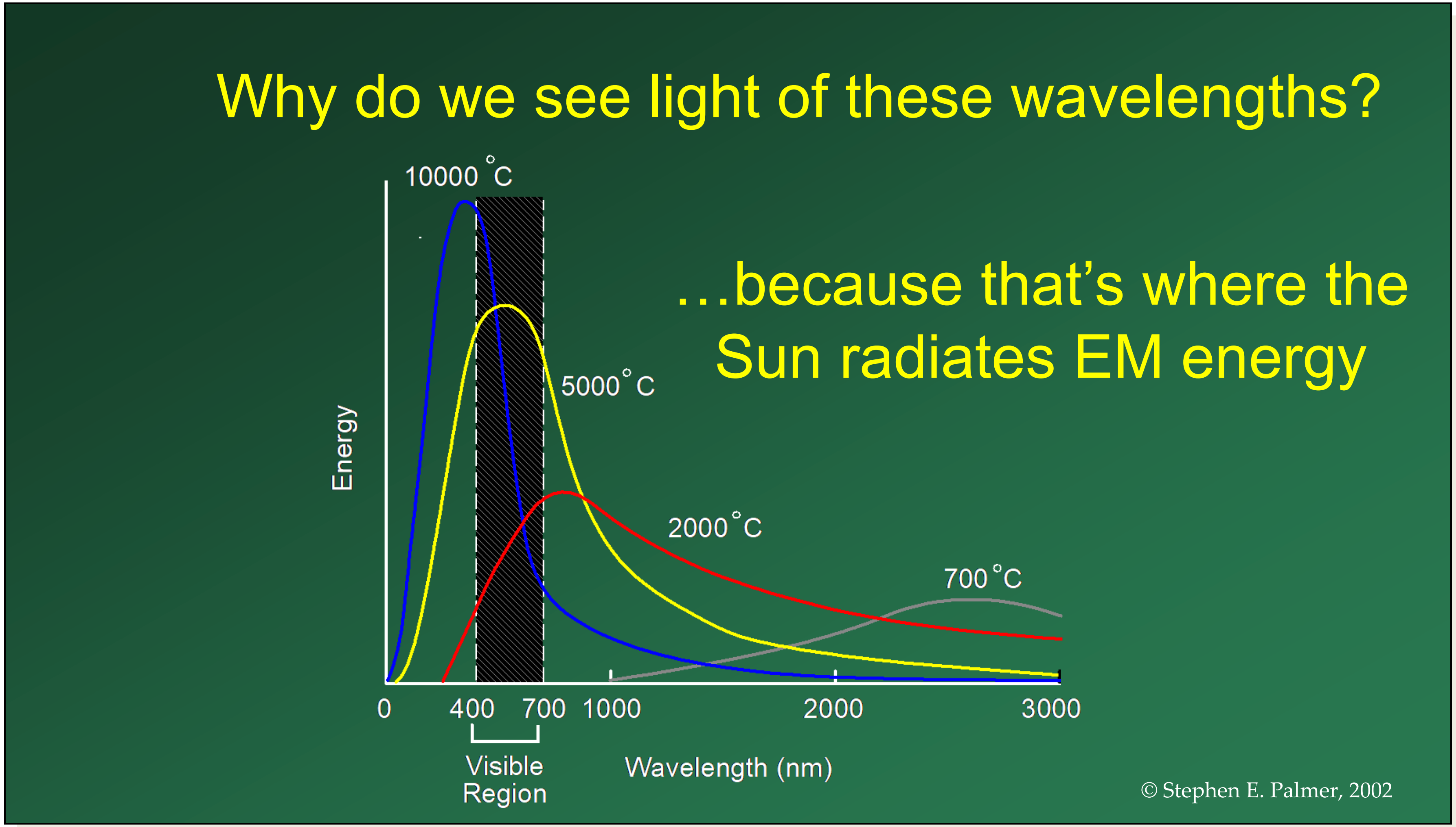
Electromagnetic Spectrum

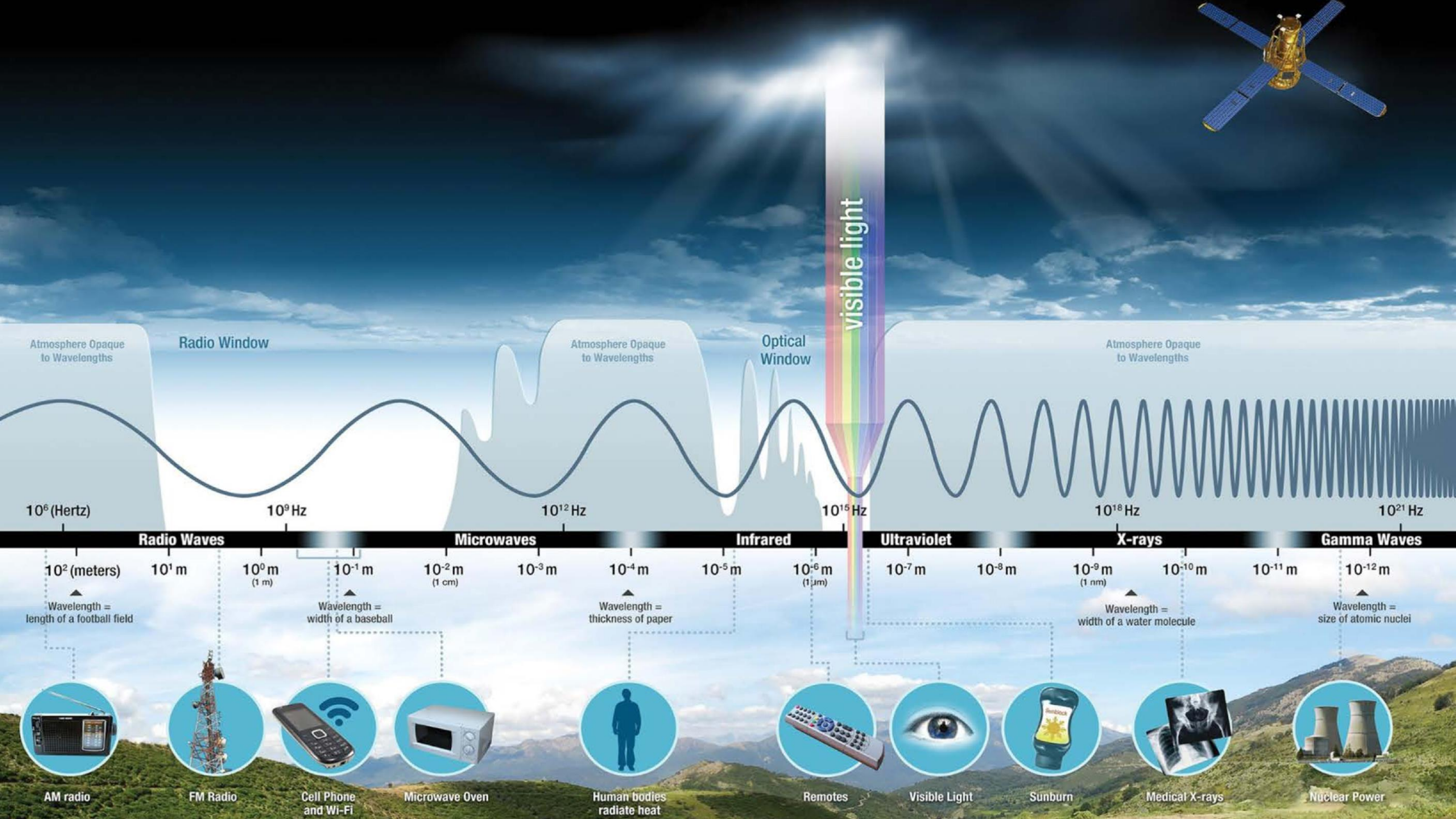


Human Luminance Sensitivity Function

http://www.yorku.ca/eye/photopik.htm

Visible Light

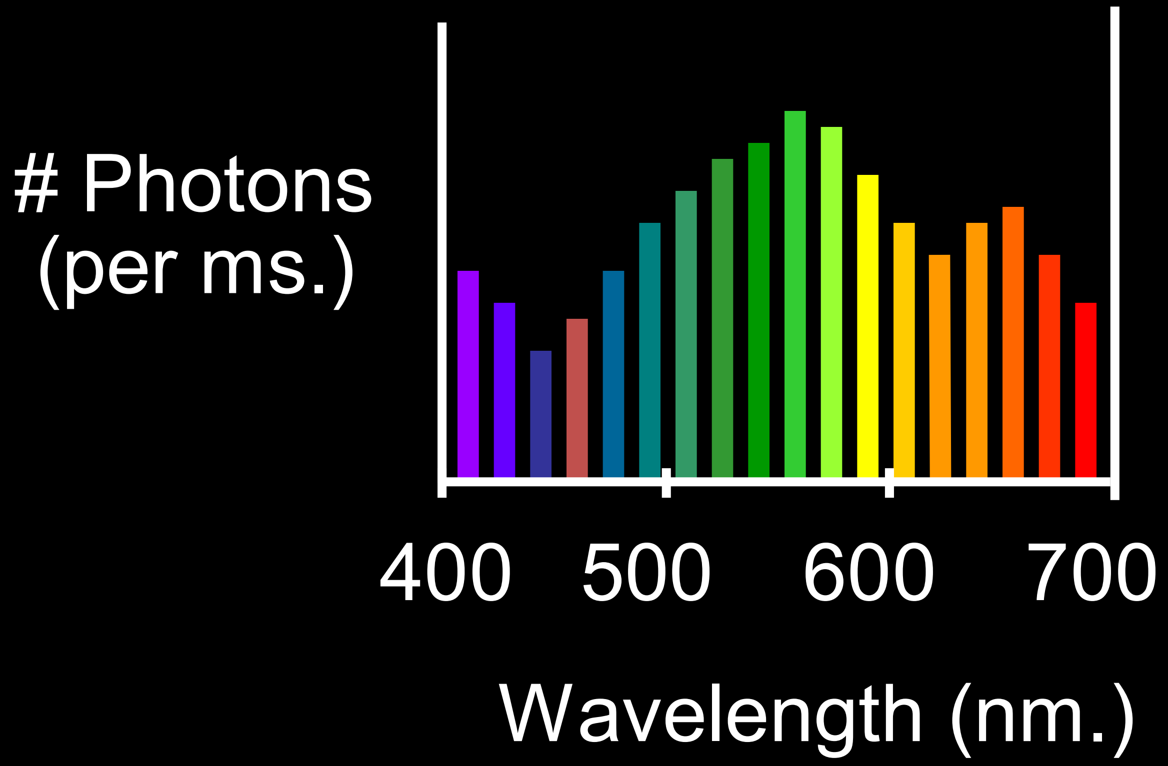




visible light
Atmosphere Opaque to Wavelengths
Radio Window
Atmosphere Opaque to Wavelengths
Optical Window
Atmosphere Opaque to Wavelengths
10^6 (Hertz)
10^9 Hz
10^12 Hz
10^15 Hz
10^18 Hz
10^21 Hz
Radio Waves
Microwaves
Infrared
Ultraviolet
X-rays
Gamma Waves
10^2 (meters)
10^1 m
10^0 m (1 m)
10^-1 m
10^-2 m (1 cm)
10^-3 m
10^-4 m
10^-5 m
10^-6 m (1 µm)
10^-7 m
10^-8 m
10^-9 m (1 nm)
10^-10 m
10^-11 m
10^-12 m
Wavelength = length of a football field
Wavelength = width of a baseball
Wavelength = thickness of paper
Wavelength = width of a water molecule
Wavelength = size of atomic nuclei
AM radio
FM Radio
Cell Phone and Wi-Fi
Microwave Oven
Human bodies radiate heat
Remotes
Visible Light
Sunburn
Medical X-rays
Nuclear Power

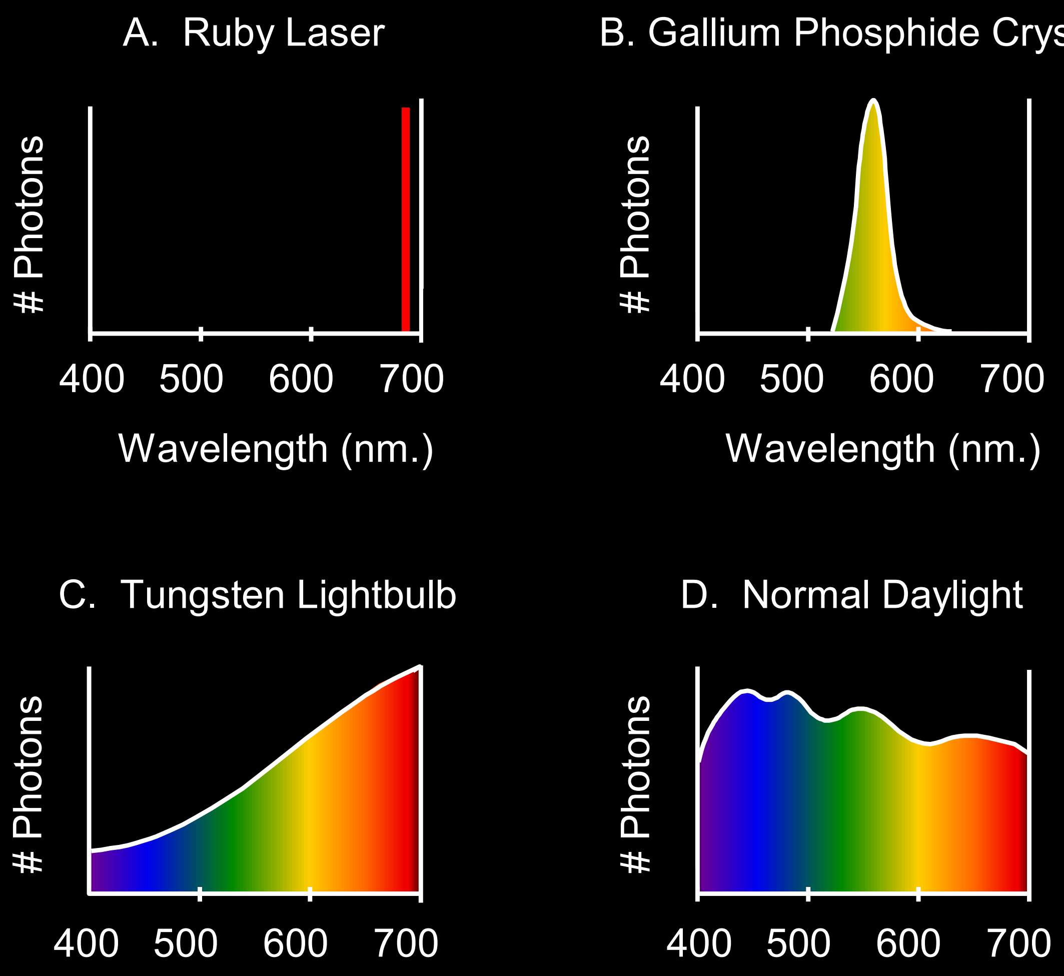
The Physics of Light

Any patch of light can be completely described physically by its spectrum: the number of photons (per time unit) at each wavelength 400 - 700 nm.



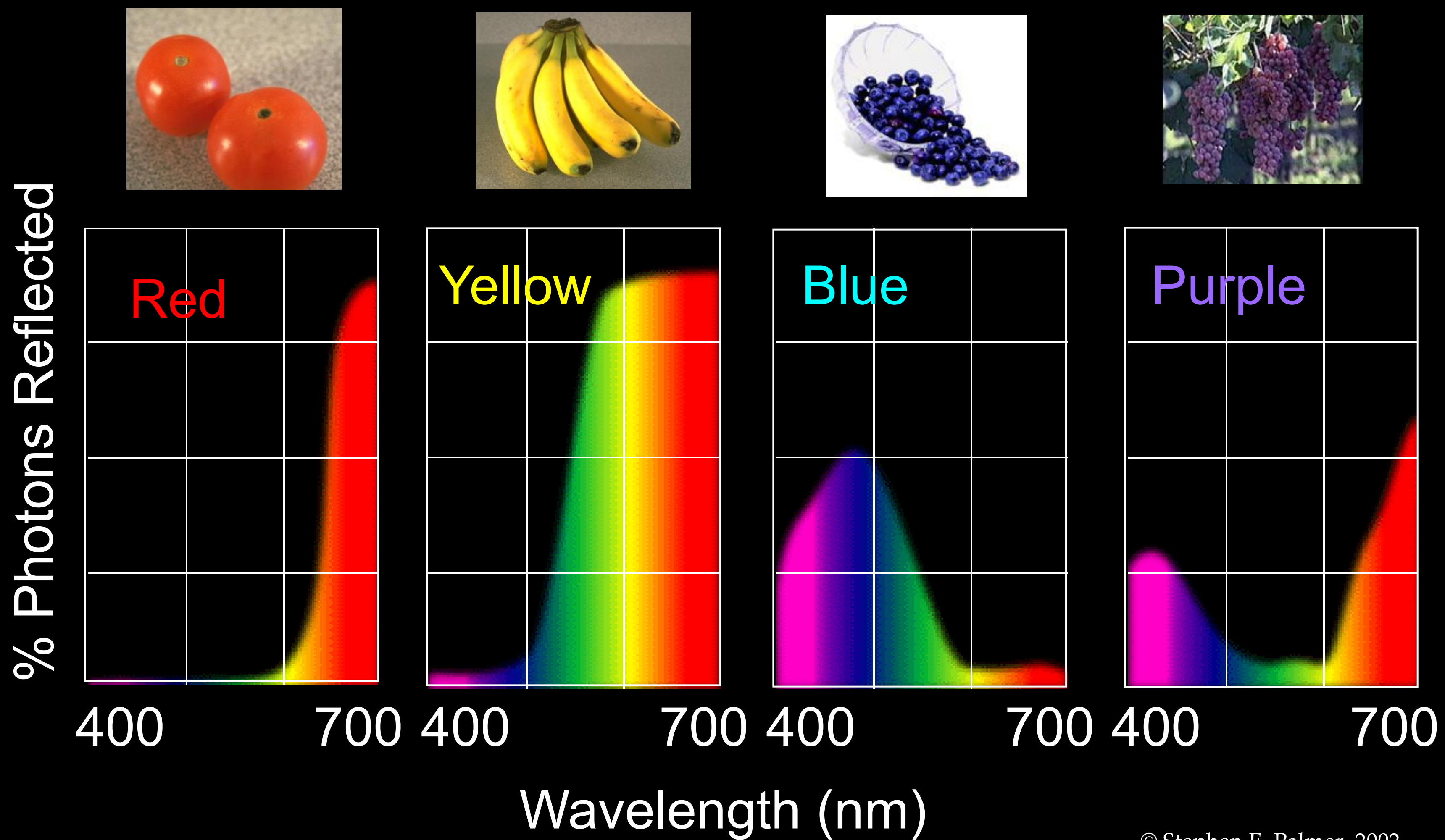
The Physics of Light

Some examples of the spectra of light sources



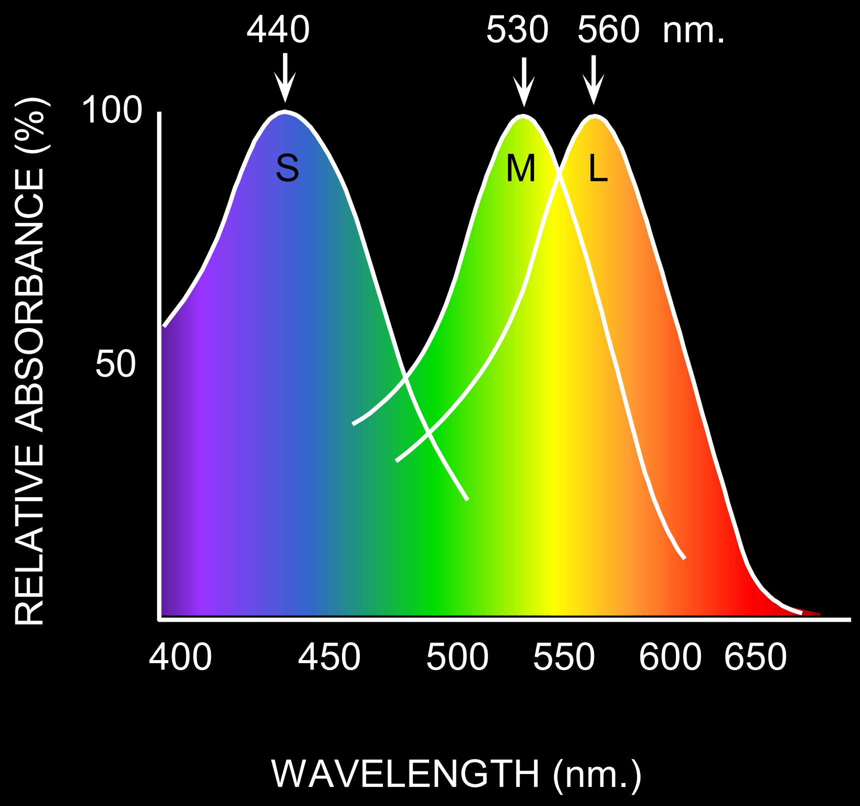
The Physics of Light

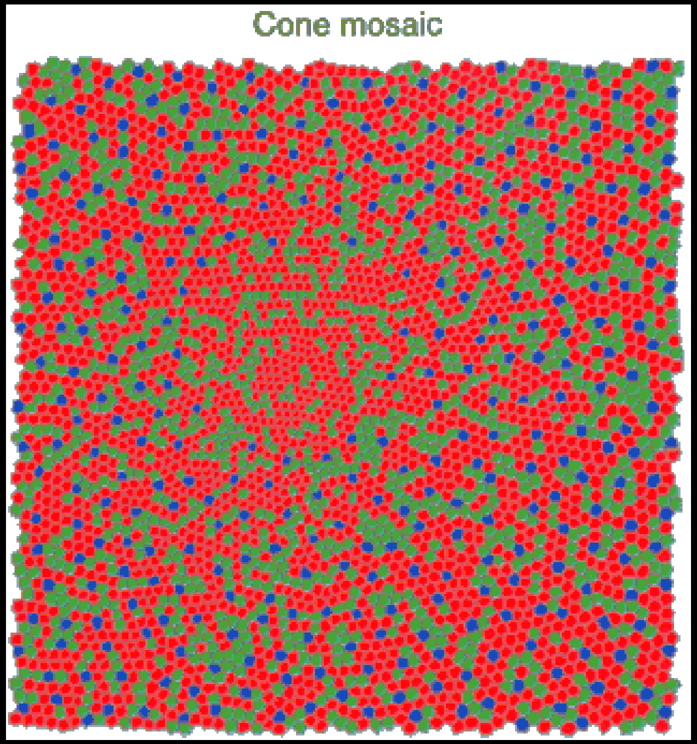
Some examples of the reflectance spectra of surfaces



Physiology of Color Vision

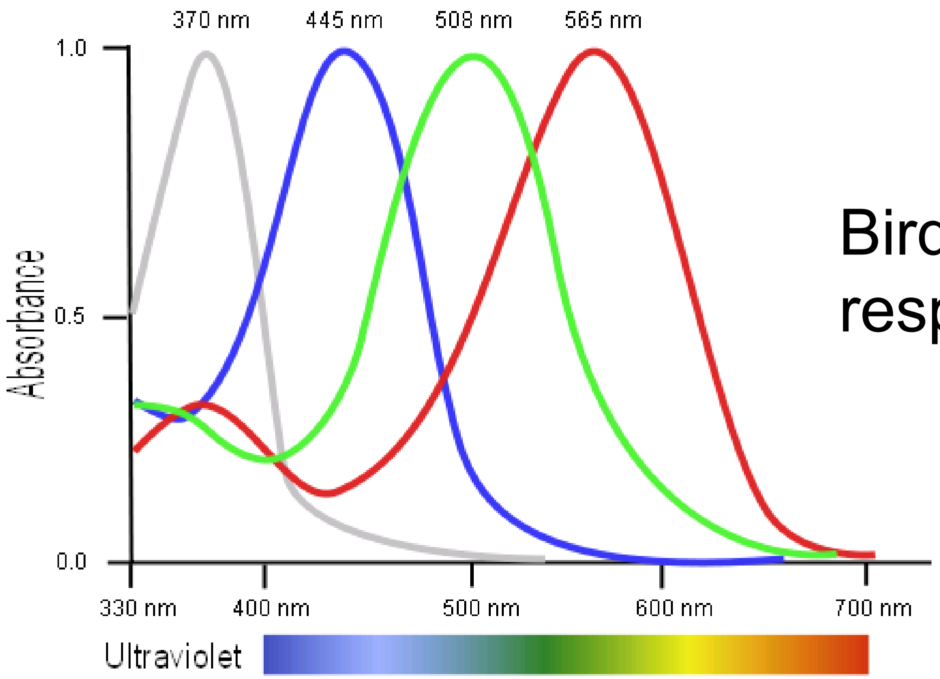
Three kinds of cones:





- Why are M and L cones so close?
- Why are there 3?

Tetrachromacy



Bird cone responses

- Most birds, and many other animals, have cones for ultraviolet light.
- Some humans, mostly female, seem to have slight tetrachromatism.

Table 1.

Cone pigment complements in some eutherian mammals.

order	exemplars	SWS1 pigment[a]	LWS pigment[a]	reference
Rodentia	*Mus* (mouse)	UV	M	Jacobs *et al.* (1991)
	Rattus (rat)	UV	M	Jacobs *et al.* (1991)
	Geomys (gopher)	UV	M	Williams *et al.* (2005)
	Cavia (guinea pig)	S	M	Parry & Bowmaker (2002)
	Spermophilus (squirrel)	S	M	Jacobs *et al.* (1985)
	Cricetomys (African rat)	absent	M/L[b]	Peichl & Moutairou (1998)
Lagomorpha	*Oryctolagus* (rabbit)	S	M	Nuboer *et al.* (1983)
primate	*Macaca* (macaque monkey)	S	M+L	Schnapf *et al.* (1988)
	Saimiri (squirrel monkey)	S	poly (3)	Mollon *et al.* (1984)
	Aotus (owl monkey)	absent	L	Jacobs *et al.* (1993*b*)
	Alouatta (howler monkey)	S	M+L	Jacobs *et al.* (1996*a*)
	Galago (bushbaby)	absent	L	Deegan II & Jacobs (1996)
	Lemur (ring-tailed)	S	L	Jacobs & Deegan II (1993)
	Propithecus (sifaka)	S	poly (2)	Tan & Li (1999)
Scandentia	*Tupaia* (tree shrew)	S	L	Jacobs & Neitz (1986)
Cetacea	*Eschrichtius* (whale)	absent	L	Levenson & Dizon (2003)
	Tursiops (dolphin)	absent	L	Fasick *et al.* (1998)
Artiodactyla	*Bos* (cow)	S	L	Jacobs *et al.* (1994)
	Odocoileus (deer)	S	M	Jacobs *et al.* (1994)
	Sus (pig)	S	L	Neitz & Jacobs (1989)
Perissodactyla	*Equus* (horse)	S	L	Carroll *et al.* (2001)
Carnivora	*Felis* (cat)	S	L	Loop *et al.* (1987)
	Canis (dog)	S	L	Jacobs *et al.* (1993*a*)
	Mustela (ferret)	S	L	Calderone & Jacobs (2003)
	Ursus (bear)	S	L	Levenson *et al.* (2006)
	Crocuta (hyena)	UV/S	M/L	Calderone *et al.* (2003)
	Phoca (seal)	absent	L	Levenson *et al.* (2006)
	Enhydra (otter)	S	L	Levenson *et al.* (2006)

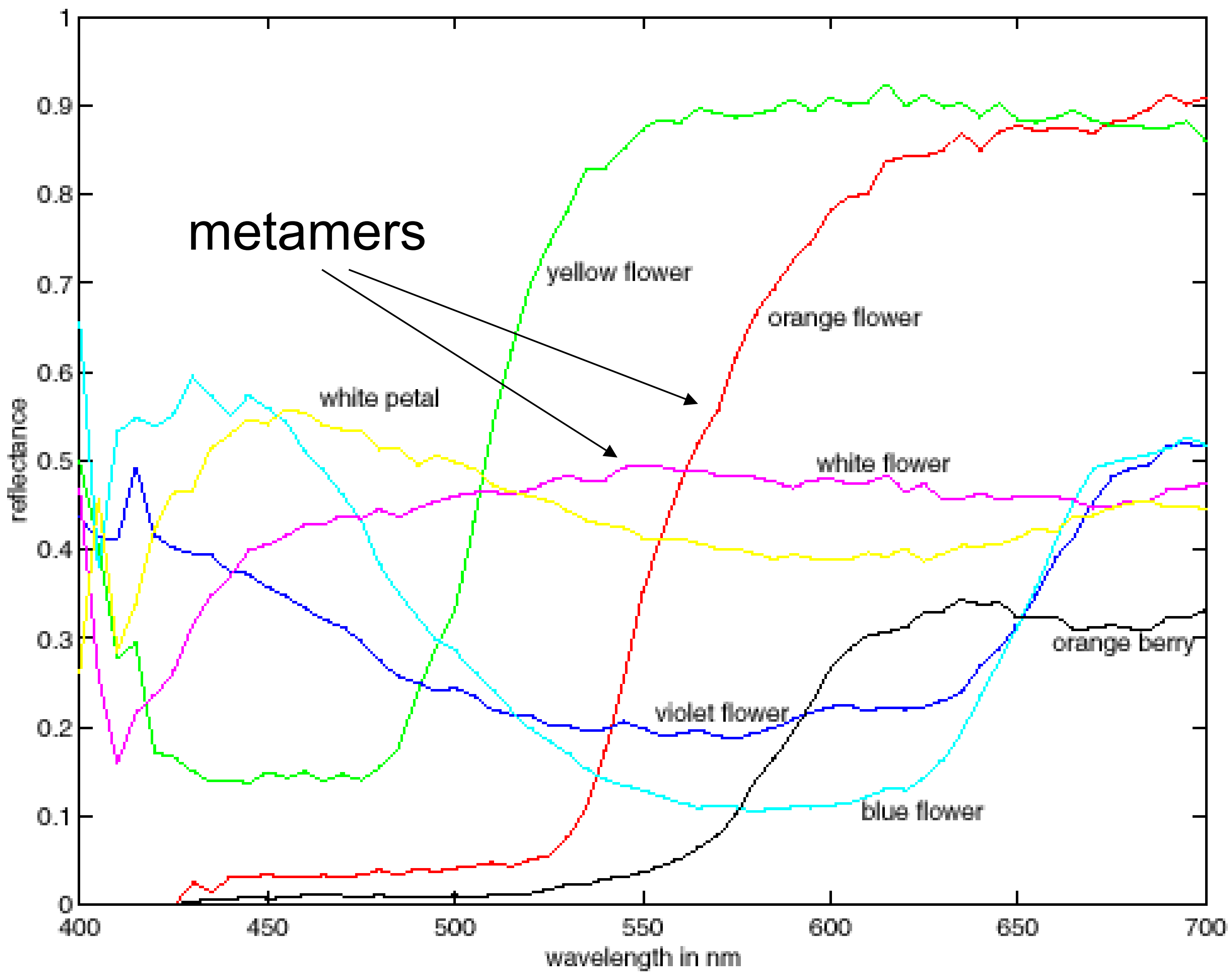
(b) Eutherians

Representatives from two cone opsin gene families appear in contemporary eutherian mammals and, with the exception of some primates, none of these animals derive more than a single photopigment type from each of their two gene families (*SWS1* and *LWS*; figure 1). Given that commonality, what cone pigments may have been present in the retinas of the early eutherians? Sequence comparisons of cone opsin genes have suggested that the ancestral mammalian pigments drawn from these two gene families peaked in the UV, at about 360 nm (Hunt *et al.* 2001), and in the long wavelengths at 560 nm (Yokoyama *et al.* 2008). Assuming the visual pigment chromophore of early mammals was 11-*cis*-retinal, the same as that of contemporary mammals, the cone pigments of early eutherian mammals would have had absorption properties similar to those sketched at the bottom of figure 1.

If these deductions are correct, the retinas of the early eutherians were similar to those of the majority of contemporary mammals in containing two types of cone pigment. Such an arrangement could support dichromatic colour vision. Whether it did would additionally depend on there having been at least some degree of selective expression of the two pigment types in separate receptor classes, on these early mammals having a nervous system organized to allow a contrast of signals from the two types of cone, and on them at least occasionally encountering photic environments sufficiently intense to activate neural comparison circuits. The fact that this basic two cone pigment arrangement is largely conserved among contemporary mammals strongly attests to its adaptive utility in our eutherian ancestors and to its probable role in supporting dichromatic colour vision in these early mammals.

Jacobs GH. Evolution of colour vision in mammals. *Philos Trans R Soc Lond B Biol Sci*. 2009;364(1531):2957-2967. doi:10.1098/rstb.2009.0039

More Spectra



Multispectral / Hyperspectral Imaging

- Astronomy / Earth Observation have long relied on multi-spectral imaging.

Hubble Space Telescope image of the Pillars of Creation, taken in 2014

False-colour images of nebulae are essentially RGB images whose colour channels have been mapped to specific emission lines. In these images, each colour can represent a specific element. In other words, a false-colour image of a nebula tells us exactly what it's made of. There are many emission lines, but the three most commonly photographed by astronomers are hydrogen-alpha, oxygen-III and sulfur-II. These emission lines are captured by using narrowband filters which only let through the light at very specific wavelengths, typically with a bandwidth of 12µm or less.

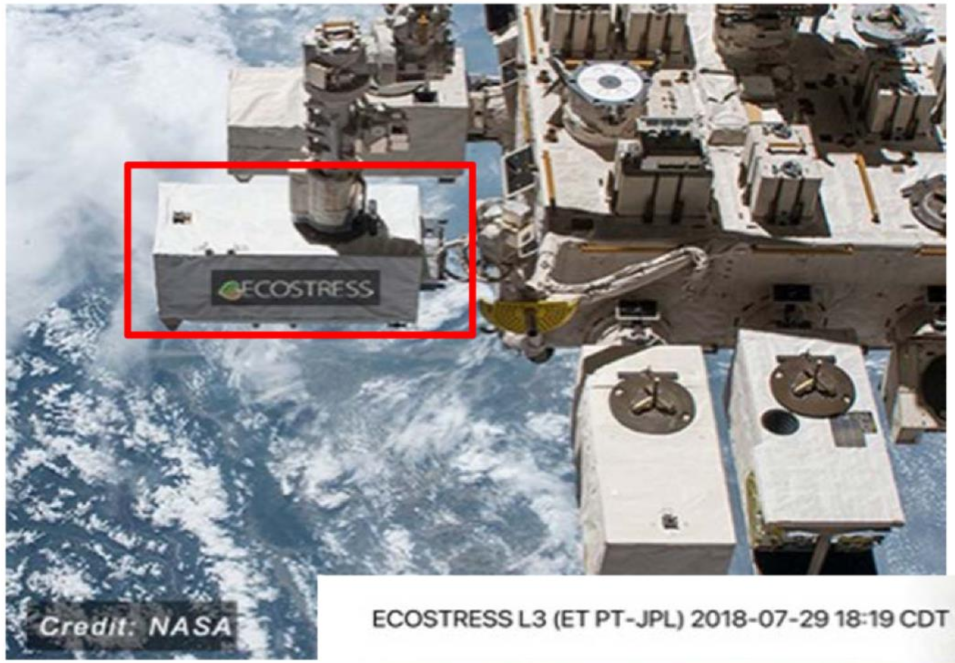
Element	Emission line	Wavelength	Colour
Hydrogen	Hα	656.3 nm	Red
Oxygen	O-III	500.7 nm	Green
Sulfur	S-II	672.4 nm	Red

Mapping Hα, O-III and S-II to red, green and blue is problematic when two of them are red, one is green and none is blue. Astronomers deal with this by using false colour — one or more of these elements is going to have to take a hit for the team and take on an unnatural hue. The Hubble palette assigns red to S-II, green to Hα, and blue to O-III: red is accurate, green and blue are false.

https://delsaert.com/deep-sky/nebulae/h-alpha/

ECOSTRESS

- Onboard the ISS
- Date Range: Aug 2018-Present
- Spatial Resolution: 70 m
- Spectral Resolution: 6 bands (160-1200 nm)
- Range: 53.6° N latitude to 53.6° S latitude
- Find Data At: Data Pool, NASA Earthdata Search, AppEEARS, and USGS EarthExplorer



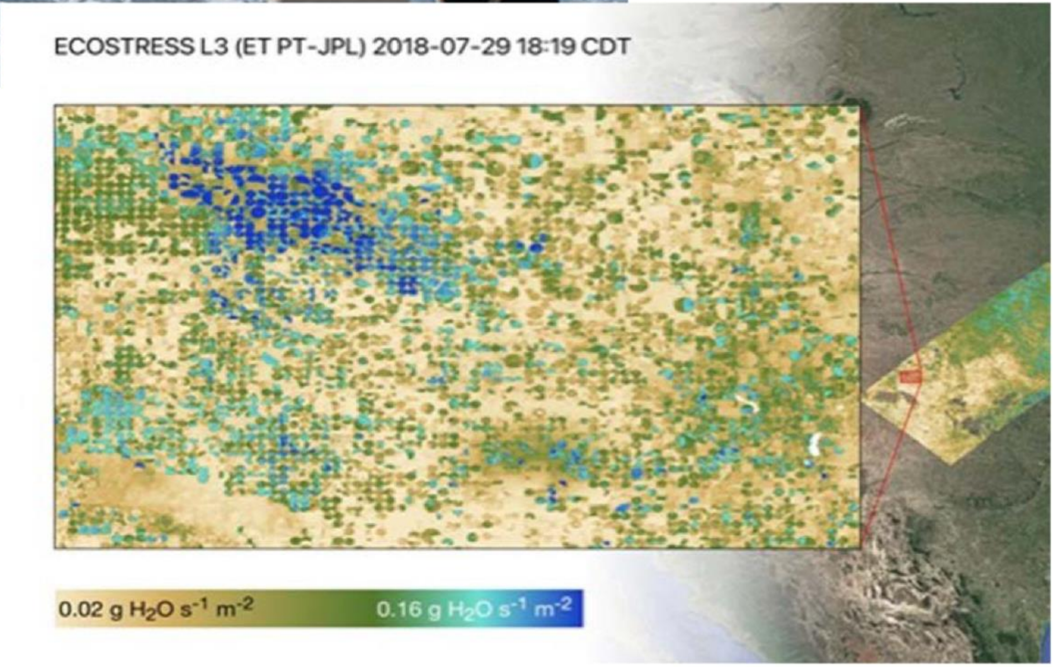
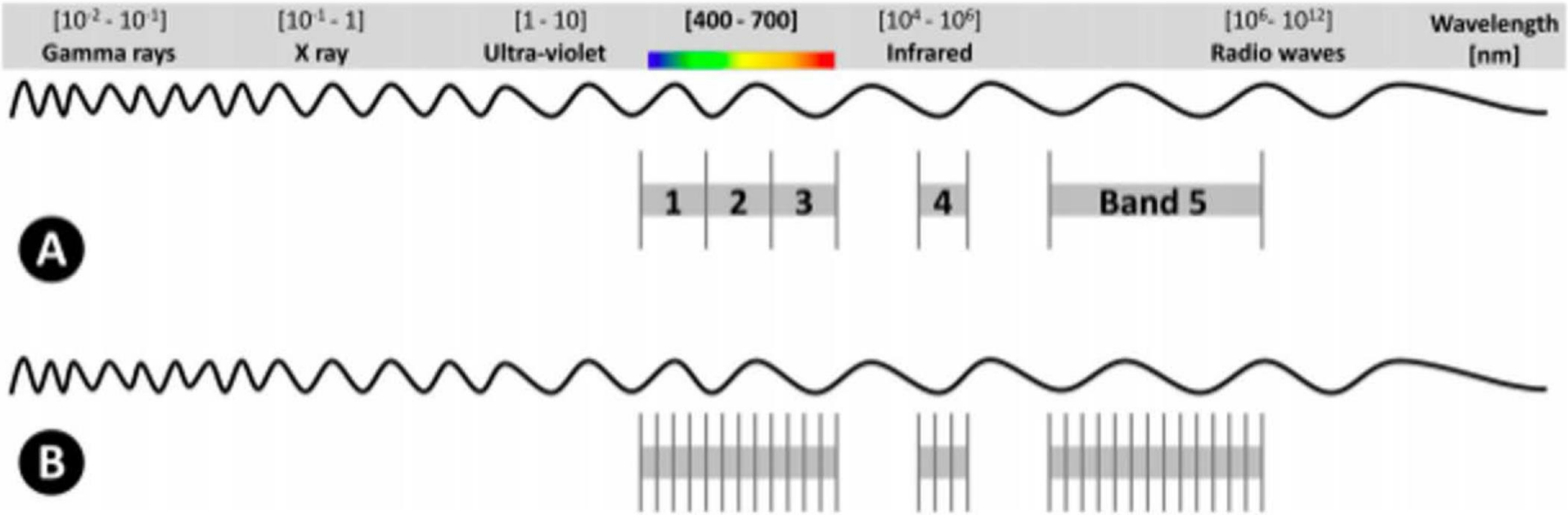


Image Credit: NASA

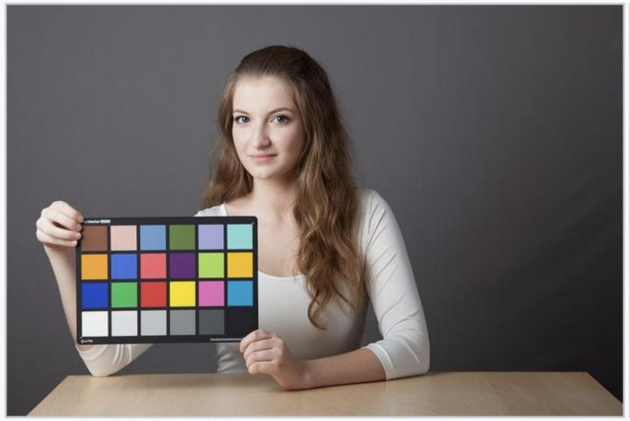
Multispectral vs. Hyperspectral Data



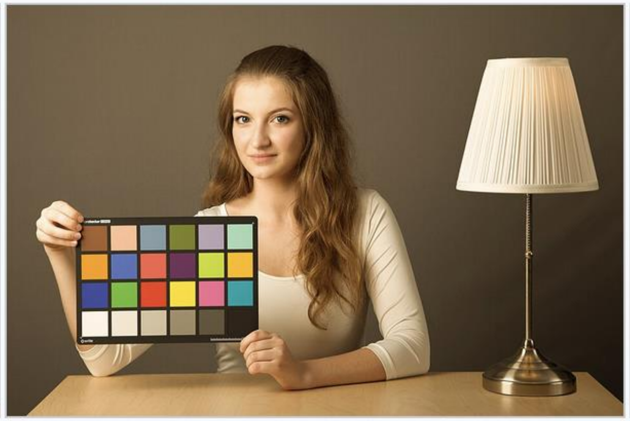
Spectrum representation including: (A) Multispectral example, with 5 wide bands; and (B) Hyperspectral example consisting of several narrow bands. Some hyperspectral sensors have several hundred bands. Image Credit: Adao, et al., 2017

Cameras don’t faithfully capture color.
RGB don’t correspond to particular wavelengths.

Examples [edit]



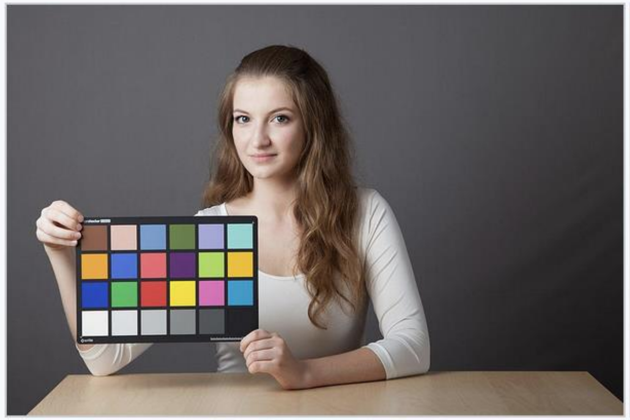
Neutral light



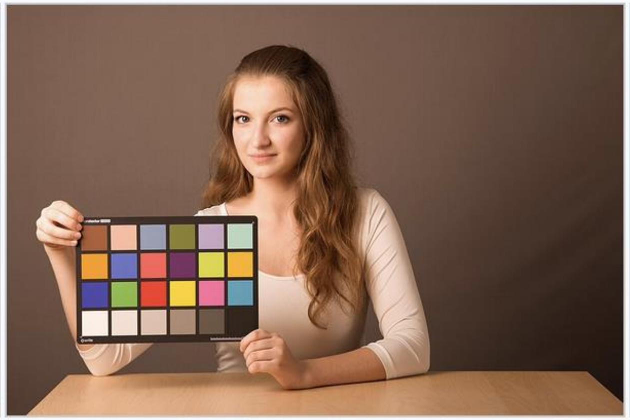
Warm light

Cold light

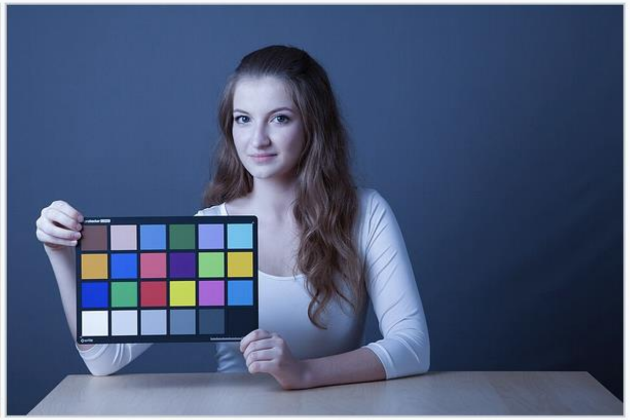
Comparison of resulted colors as shot by the digital camera for different light qualities (color temperature): Neutral, Warm and Cold.[16]



Setting: As shot



Setting: Cloudy



Setting: Tungsten

Example of different white balance settings on digital camera for neutral light[16]

Color can be ambiguous

Color can be ambiguous

The dress

From Wikipedia, the free encyclopedia

For other uses, see The Dress.

The dress is a photograph that became a viral internet sensation on 26 February 2015, when viewers disagreed over whether the dress pictured was coloured black and blue, or white and gold. The phenomenon revealed differences in human colour perception, which have been the subject of ongoing scientific investigations into neuroscience and vision science, producing a number of papers published in peer-reviewed science journals.

The photo originated from a washed-out colour photograph of a dress posted on the social networking service Facebook. Within the first week after the surfacing of the image, more than 10 million tweets mentioned the dress, using hashtags such as #thedress, #whiteandgold, and #blackandblue. Although the colour of the dress was eventually confirmed as black and blue,[3][4] the image prompted many discussions, with users discussing their different perceptions of the dress's colour. Members of the scientific community began to investigate the photo for fresh insights into human colour vision.

The dress itself, which was identified as a product of the retailer Roman Originals, experienced a major surge in sales as a result of the incident. The retailer also produced a one-off version of the dress in white and gold as a charity campaign.

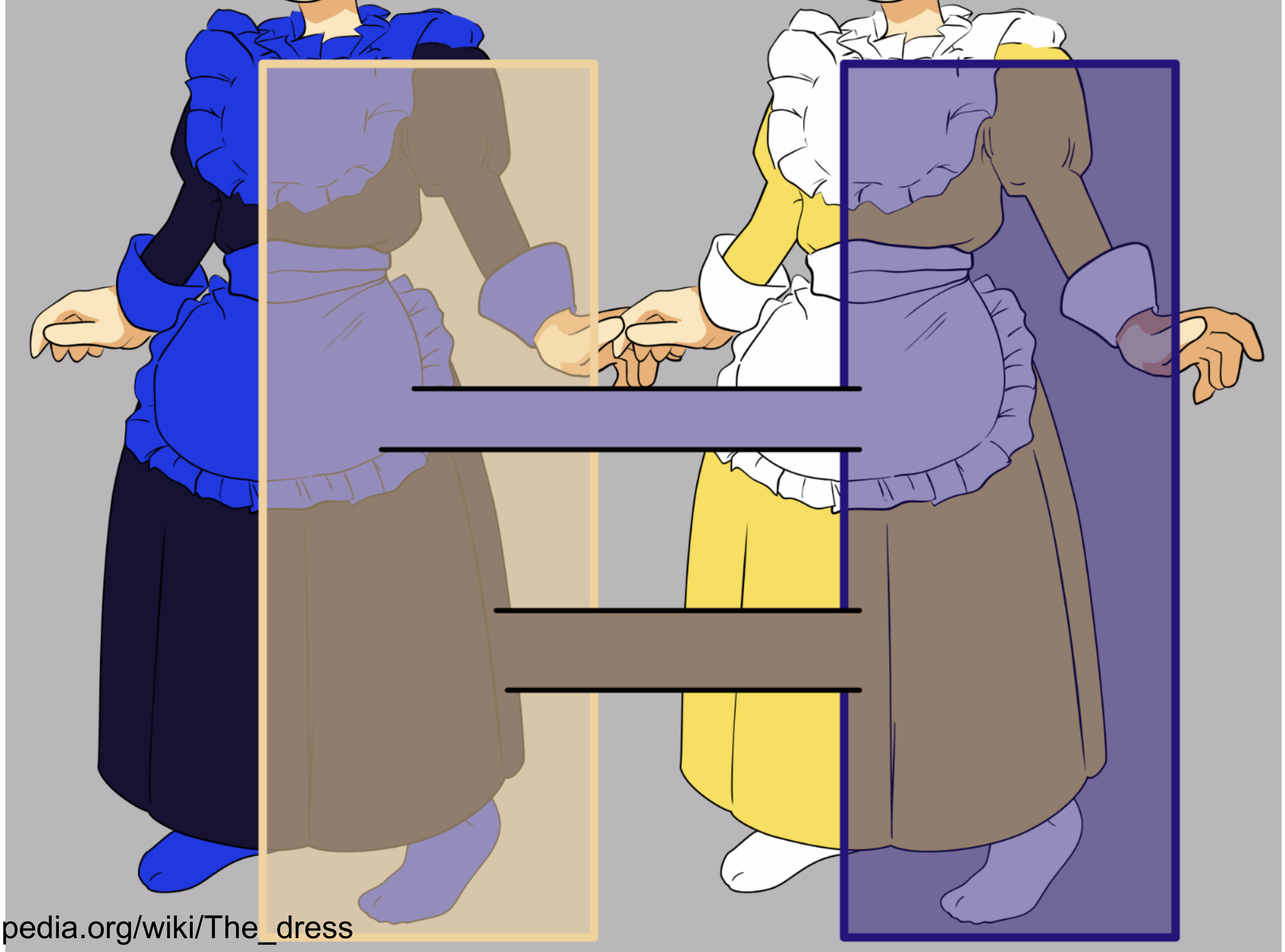
The dress

The original *The dress* picture

Designer Roman Originals[1]

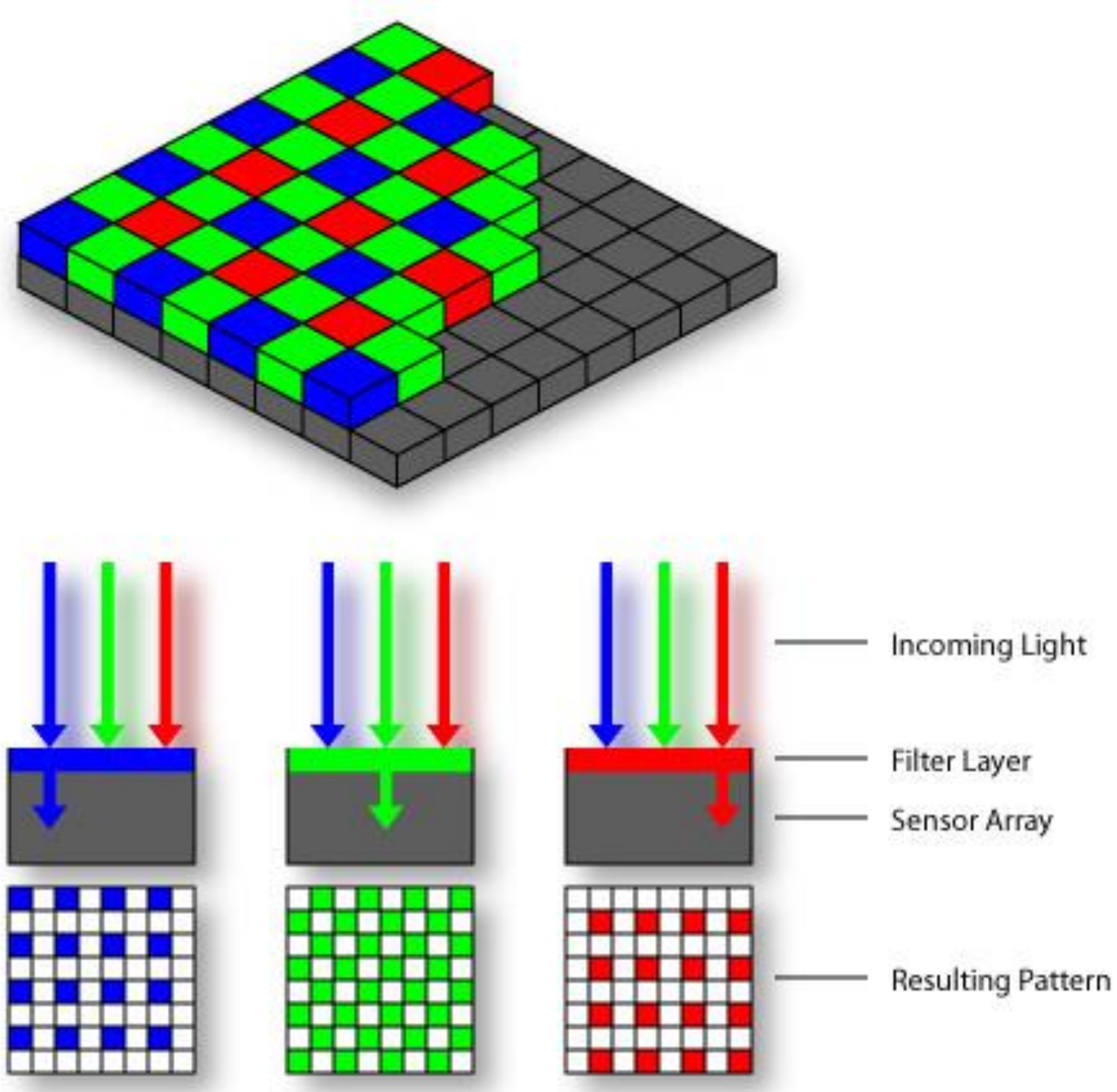
Contents [hide]

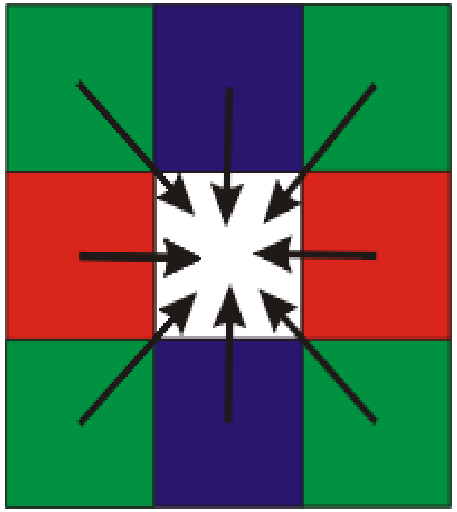
https://en.wikipedia.org/wiki/The_dress



https://en.wikipedia.org/wiki/The_dress

Practical Color Sensing: Bayer Grid





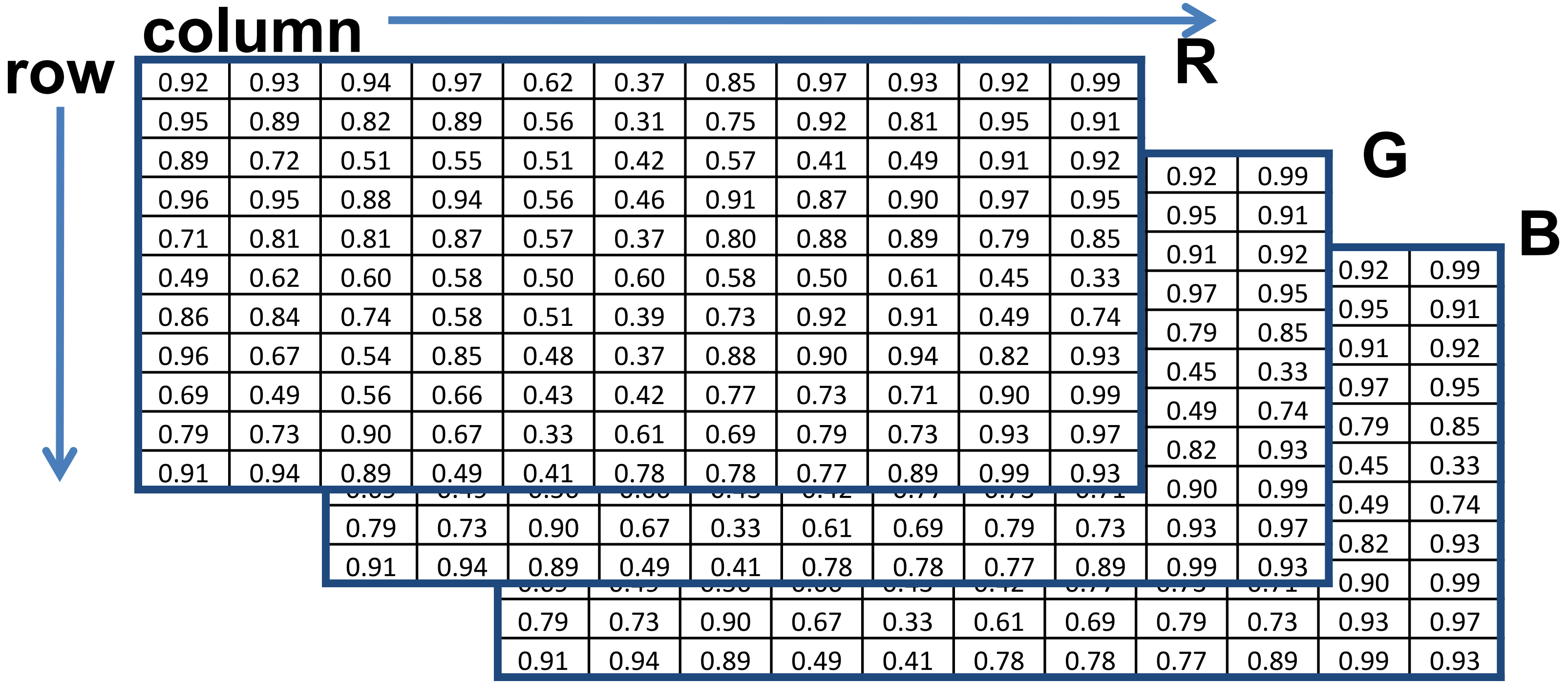
- Estimate RGB at ‘G’ cells from neighboring values

Slide by Steve Seitz

Color Image

Images in Python

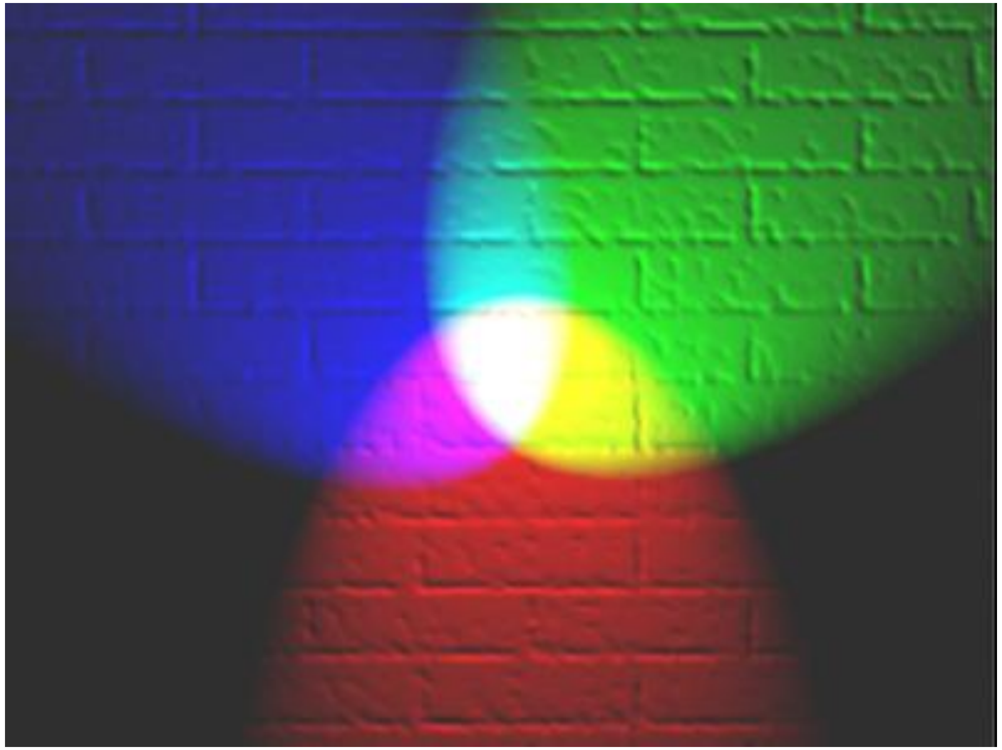
- Images represented as a matrix
- Suppose we have a NxM RGB image called “im”
  - im(0,0,0) = top-left pixel value in R-channel
  - im(y, x, b) = y pixels down, x pixels to right in the bth channel
  - im(N-1, M-1, 2) = bottom-right pixel in B-channel



0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

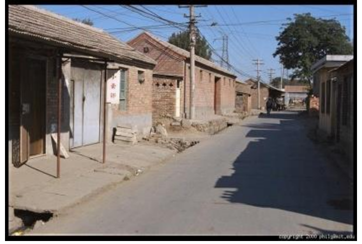
Color spaces

- Are there other ways to encode color beyond R, G, B?

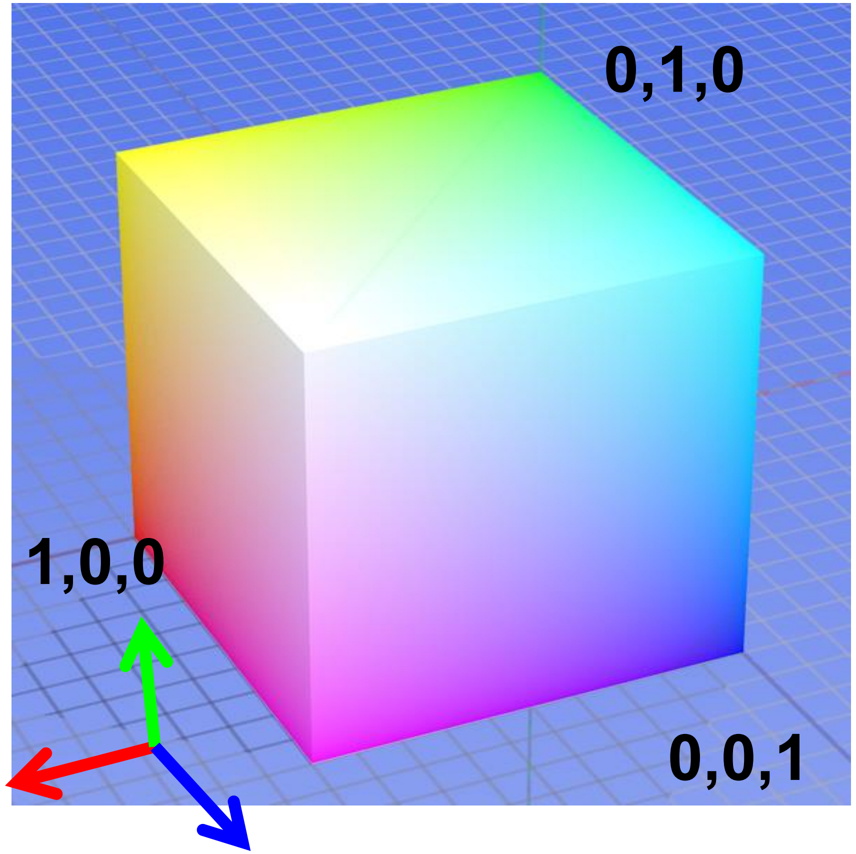


http://en.wikipedia.org/wiki/File:RGB_illumination.jpg

Color spaces: RGB



Default color space



R
(G=0,B=0)

G
(R=0,B=0)

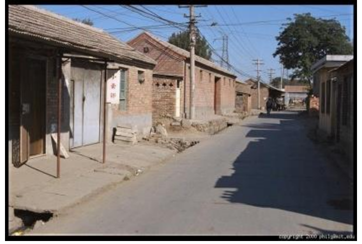
B
(R=0,G=0)

Some drawbacks

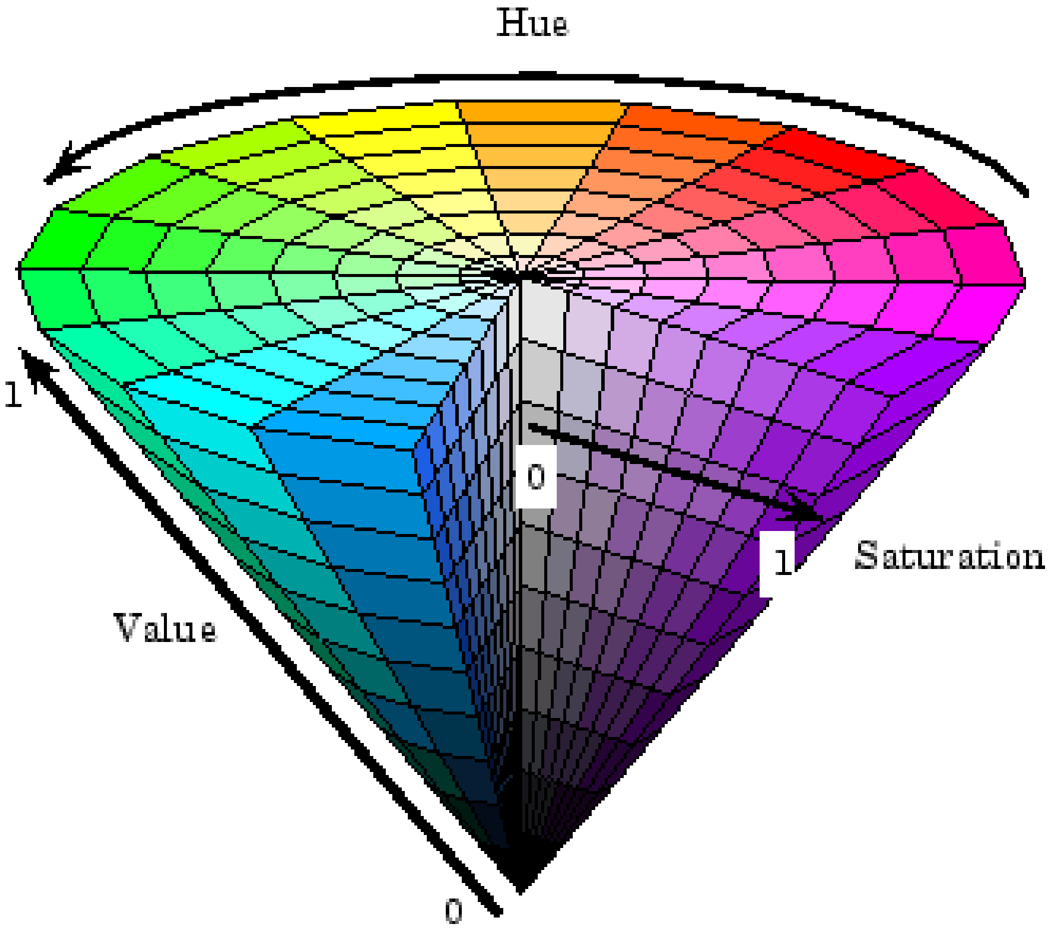
- Strongly correlated channels
- Non-perceptual

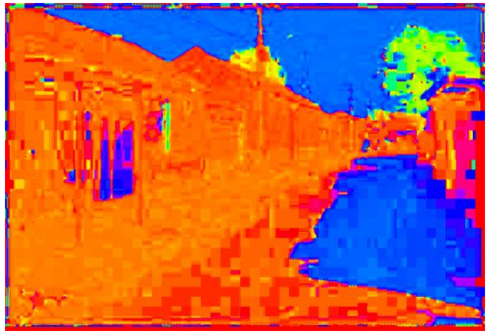
Image from: http://en.wikipedia.org/wiki/File:RGB_color_solid_cube.png

Color spaces: HSV



Intuitive color space



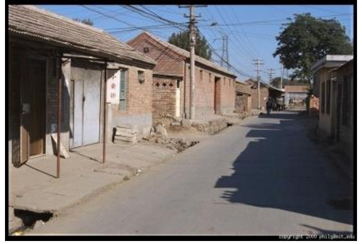


H
(S=1,V=1)

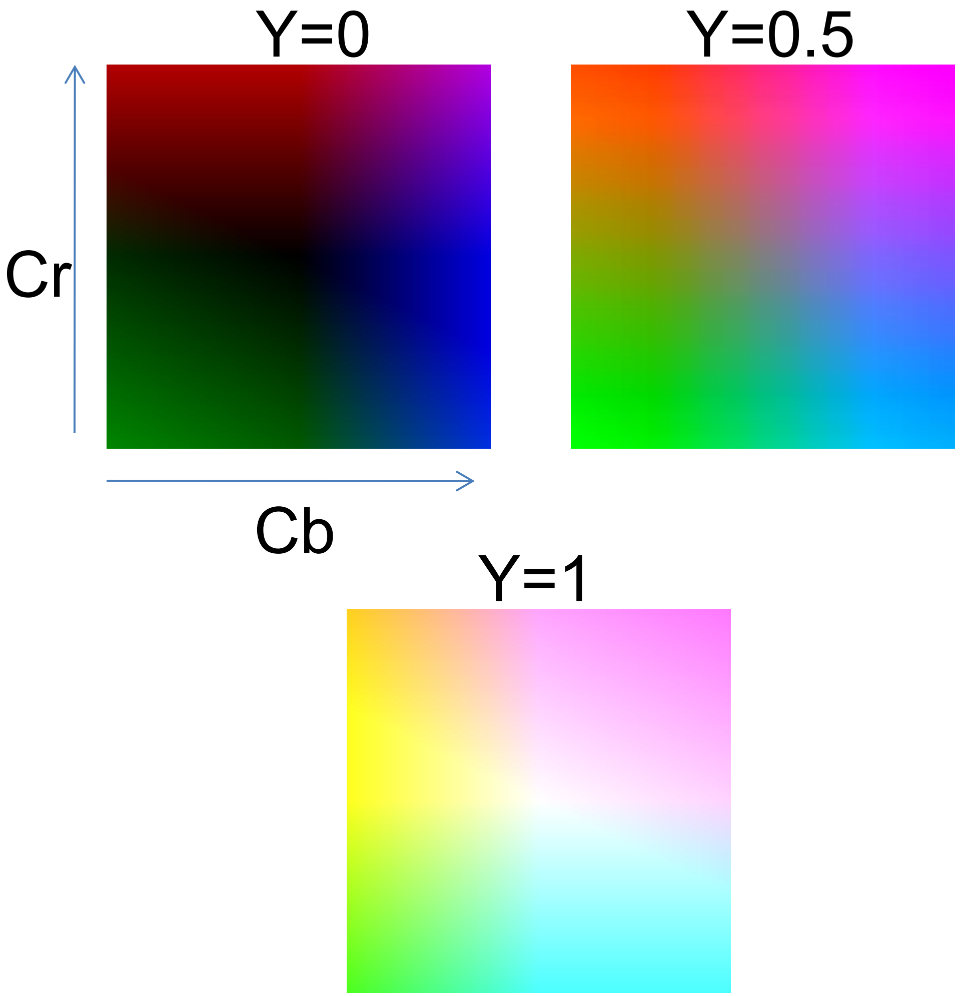
S
(H=1,V=1)

V
(H=1,S=0)

Color spaces: YCbCr



Fast to compute, good for compression, used by TV

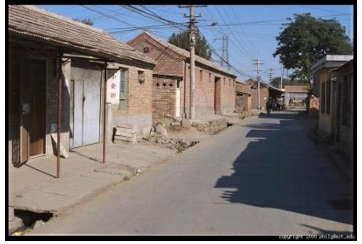


Y
(Cb=0.5,Cr=0.5)

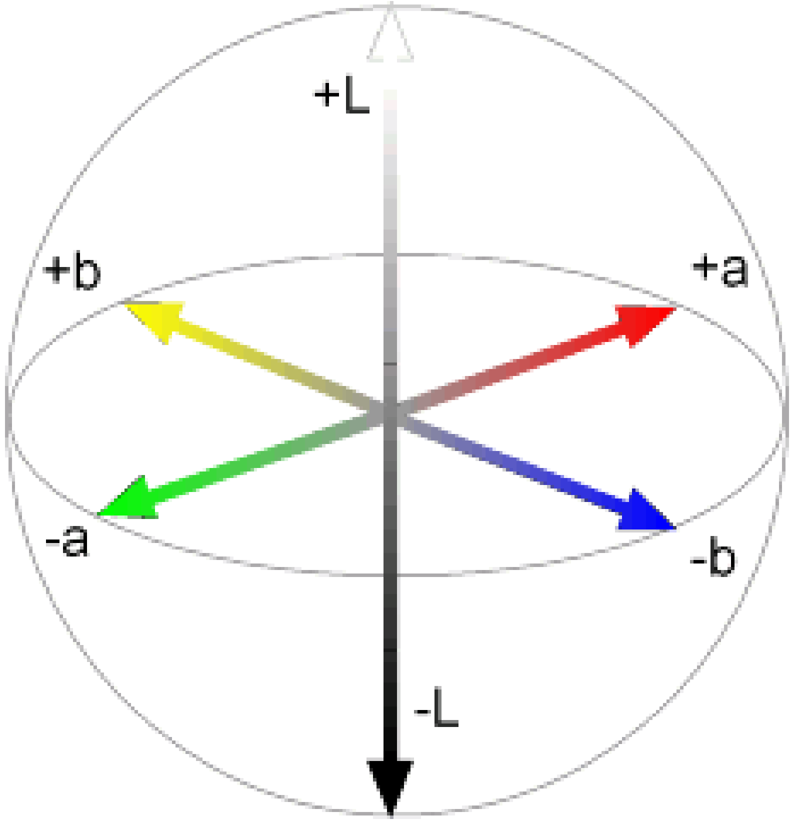
Cb
(Y=0.5,Cr=0.5)

Cr
(Y=0.5,Cb=05)

Color spaces: L*a*b*



"Perceptually uniform"* color space



L
(a=0,b=0)

a
(L=65,b=0)

b
(L=65,a=0)

If you had to choose, would you rather go without luminance or chrominance?

If you had to choose, would you rather go without luminance or chrominance?

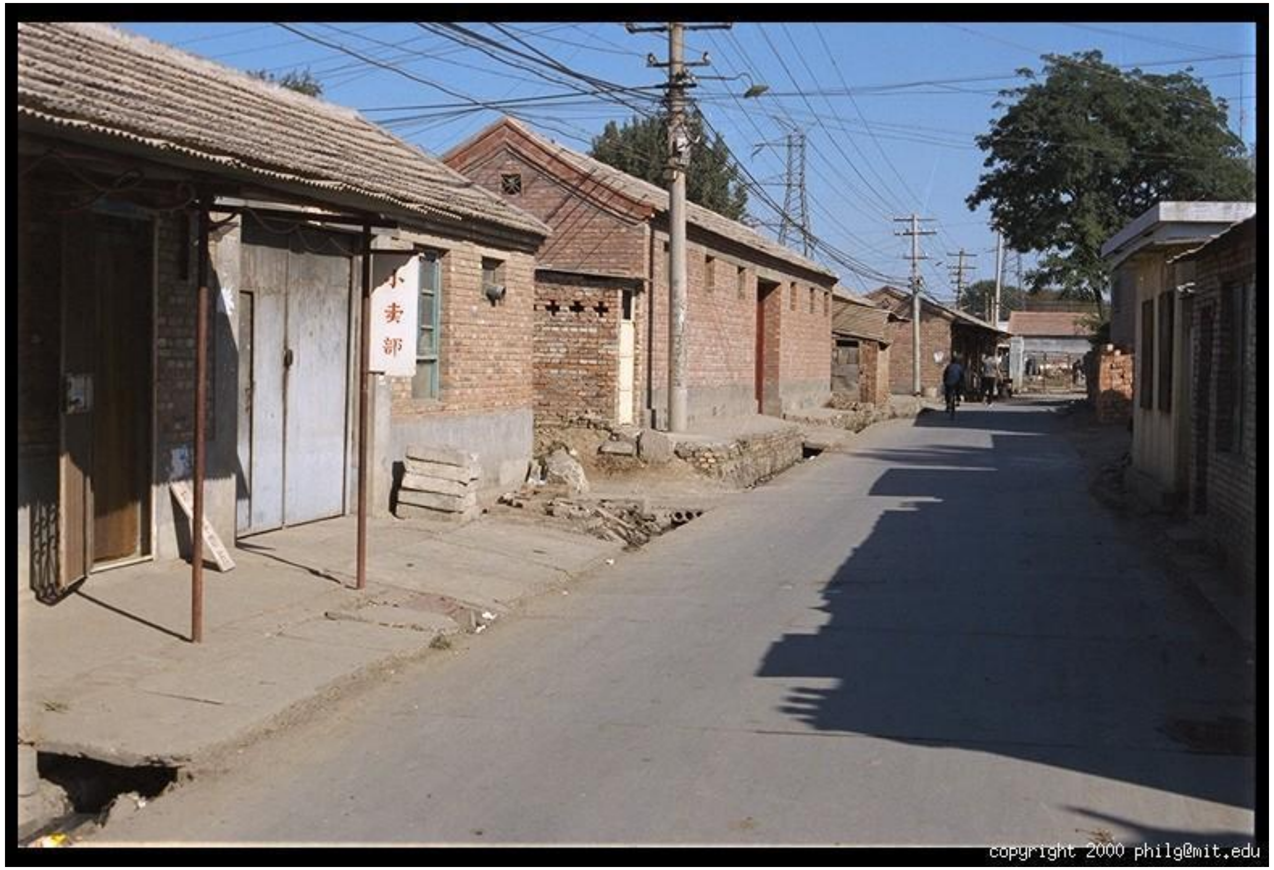
Most information in intensity

Only color shown – constant intensity

Most information in intensity

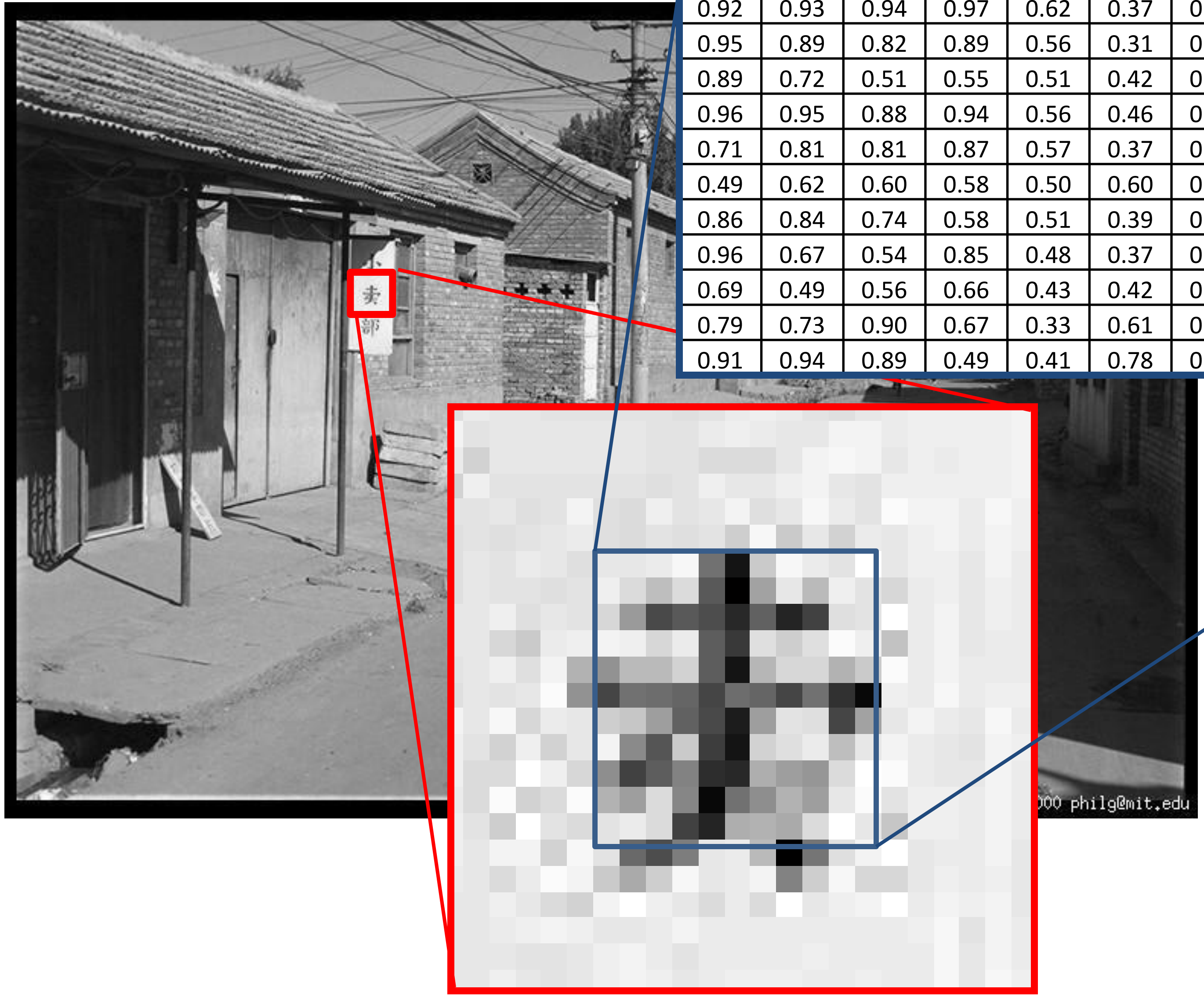
Only intensity shown – constant color

Most information in intensity



Original image

Back to grayscale intensity



0.92	0.93	0.94	0.97	0.62	0.37	0.85	0.97	0.93	0.92	0.99
0.95	0.89	0.82	0.89	0.56	0.31	0.75	0.92	0.81	0.95	0.91
0.89	0.72	0.51	0.55	0.51	0.42	0.57	0.41	0.49	0.91	0.92
0.96	0.95	0.88	0.94	0.56	0.46	0.91	0.87	0.90	0.97	0.95
0.71	0.81	0.81	0.87	0.57	0.37	0.80	0.88	0.89	0.79	0.85
0.49	0.62	0.60	0.58	0.50	0.60	0.58	0.50	0.61	0.45	0.33
0.86	0.84	0.74	0.58	0.51	0.39	0.73	0.92	0.91	0.49	0.74
0.96	0.67	0.54	0.85	0.48	0.37	0.88	0.90	0.94	0.82	0.93
0.69	0.49	0.56	0.66	0.43	0.42	0.77	0.73	0.71	0.90	0.99
0.79	0.73	0.90	0.67	0.33	0.61	0.69	0.79	0.73	0.93	0.97
0.91	0.94	0.89	0.49	0.41	0.78	0.78	0.77	0.89	0.99	0.93

Wrap up: Why do we care about cameras vs eyes?

Inside Tesla as Elon Musk Pushed an Unflinching Vision for Self-Driving Cars

The automaker may have undermined safety in designing its Autopilot driver-assistance system to fit its chief executive's vision, former employees say.

Hardware choices have also raised safety questions. Within Tesla, some argued for pairing cameras with radar and other sensors that worked better in heavy rain and snow, bright sunshine and other difficult conditions. For several years, Autopilot incorporated radar, and for a time Tesla worked on developing its own radar technology. But three people who worked on the project said Mr. Musk had repeatedly told members of the Autopilot team that humans could drive with only two eyes and that this meant cars should be able to drive with cameras alone.

Schuyler Cullen, who oversaw a team that explored autonomous-driving possibilities at the South Korean tech giant Samsung, said in an interview that Mr. Musk's cameras-only approach was fundamentally flawed. "Cameras are not eyes! Pixels are not retinal ganglia! The F.S.D. computer is nothing like the visual cortex!" said Mr. Cullen, a computer vision specialist who now runs a start-up that is building a new kind of camera-based sensor.

Amnon Shashua, chief executive of Mobileye, a former Tesla supplier that has been testing technology that is similar to the electric-car maker's, said Mr. Musk's idea of using only cameras in a self-driving system could ultimately work, though other sensors may be needed in the short term. He added that Mr. Musk might exaggerate the capabilities of the company's technology, but that those statements shouldn't be taken too seriously.

Another driving system with human level eyesight

https://www.youtube.com/watch?v=5fvsItXYgzk

Toonces the Driving Cat: Driver's Test - SNL

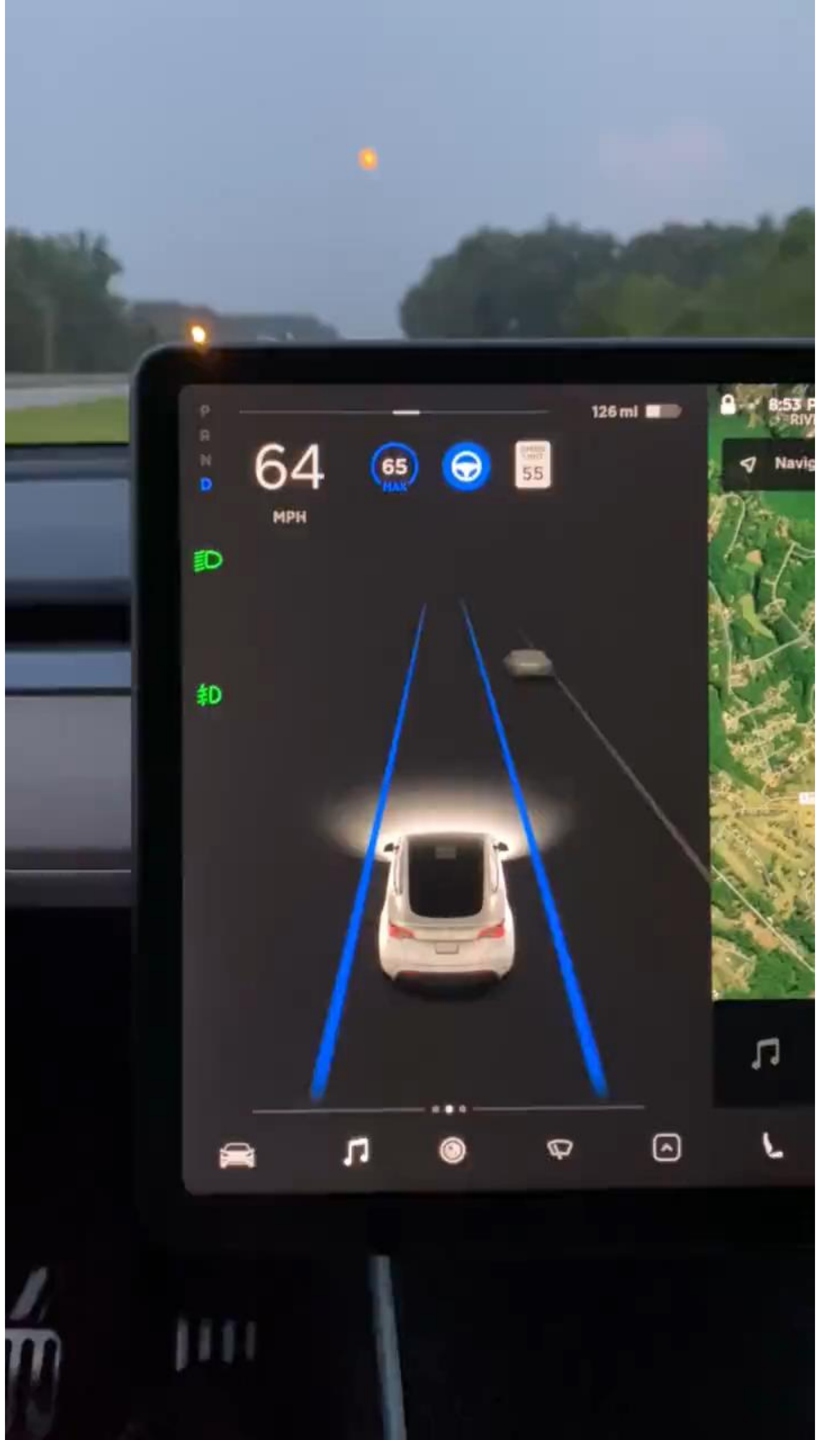
Saturday Night Live
16.1M subscribers

Subscribe

26K

Share

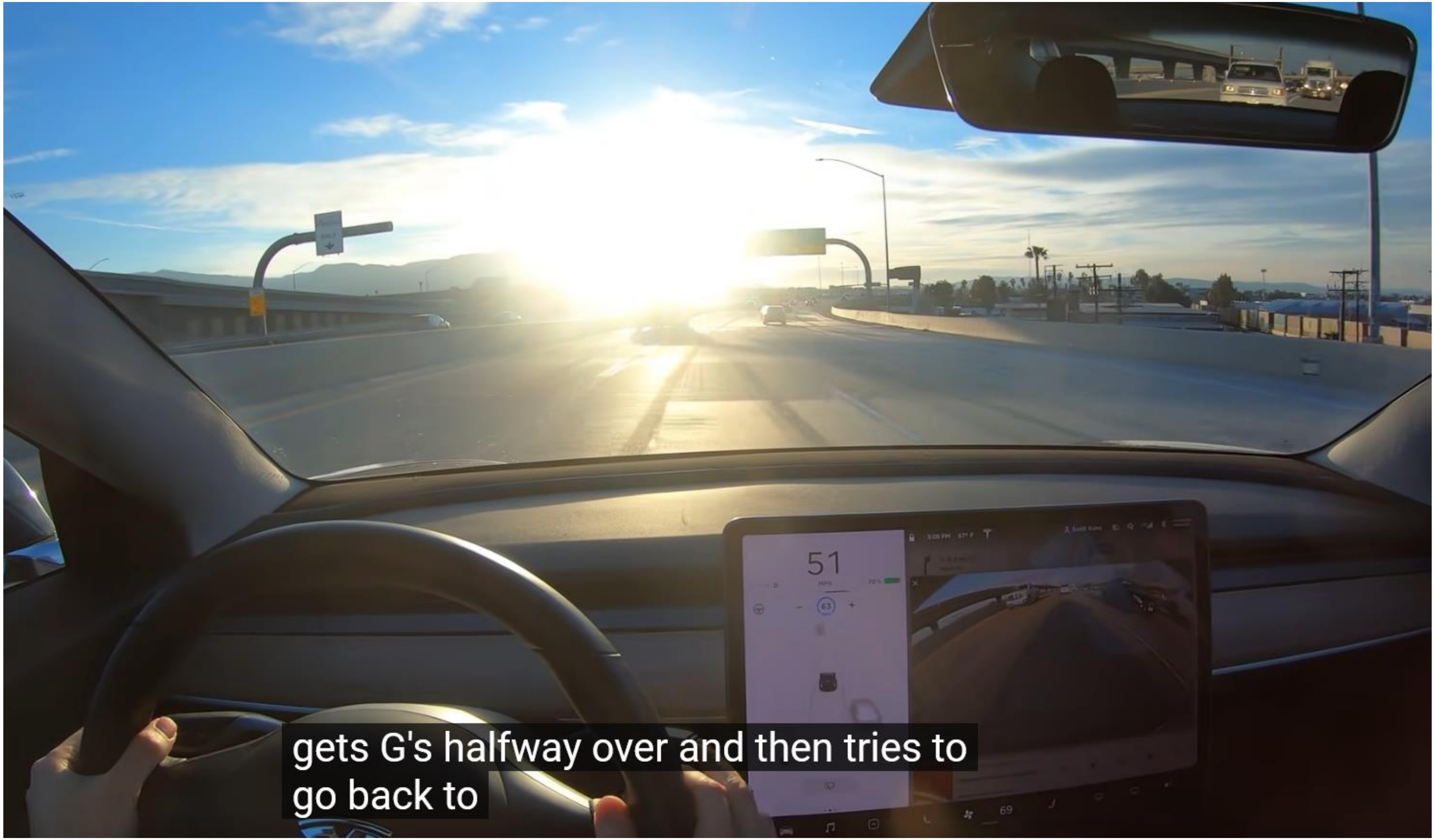
Download



https://twitter.com/JordanTeslaTech/status/1418413307862585344

https://www.reddit.com/r/teslamotors/comments/nrs8kf/you_think_ice_cream_truck_stop_signs_are_a_problem/

- On the whole, cameras *are* a reasonable analogy for eyes. They do capture sufficient information for safe driving 99.9% of the time.



https://www.youtube.com/watch?v=s1eTBMYEs-Y

- On the whole, cameras *are* a reasonable analogy for eyes. They do capture sufficient information for safe driving 99.9% of the time.
  - Imagine remote controlling a vehicle based on a camera feed.
- *But* the computer vision and machine learning methods that interpret the camera images *are not* yet a reasonable analogy for the human brain.
- *And* there are outlier cases where humans must do more than just passively perceive from a single fixed viewpoint.

Next: Interest points and corners