Introduction to computer vision
Digital image, light and color

Vlad Shakhuro










17 September 2026

Outline

1. A short intro about your career and AI
2. What is computer vision?
3. Digital image
4. Light and color
5. Color models
6. Features of human vision

This year course is AI-restricted. Why?

lecturer view:

- all online courses are dead
- but we want to teach those who want to learn

student view:

- AI can do a lot of work **for** you
- learn hard skills s.t. AI won't work **instead** of you

Outline

1. A short intro about your career and AI
2. What is computer vision?
3. Digital image
4. Light and color
5. Color models
6. Features of human vision

What is computer vision?



Vision task: understand, what is depicted on an image

Computer vision: developing a computer model for vision. CV is a part of Artificial Intelligence (AI)

Turing test for CV: answer any question about the image that can be answered by a human

Photogrammetry



→



3D reconstruction and measurements from photos

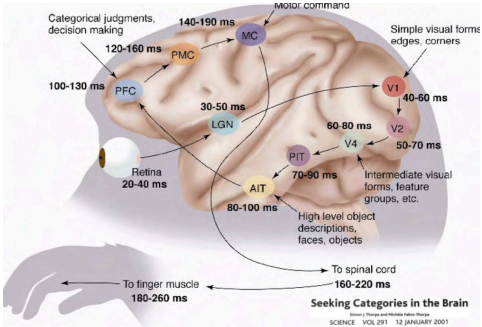
Human vision



We match observations with our knowledge

It is estimated that 25% of brain is used for solving vision task

Works of David Marr (1970s)

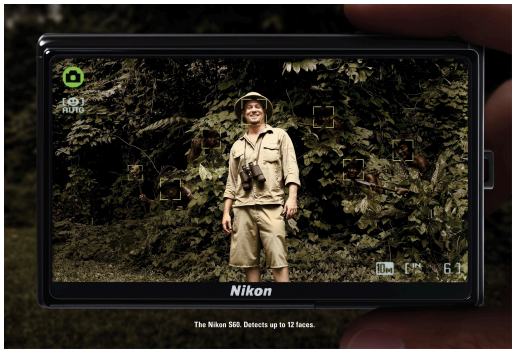


According to Marr, human vision is done in several steps:

- Primal sketch. Low-level features of an image: edges, regions
- 2.5D sketch. Ordering by depth (stereo vision), textures
- 3D model. Human recognizes objects and infers a continuous 3-dimensional map

Viola-Jones face detector (2001)



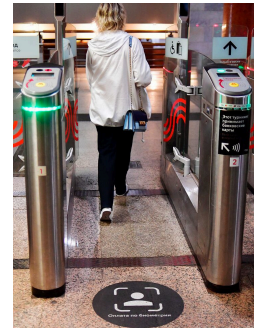


One of the first real-life applications of computer vision: face detection model that can be computed on a mobile CPU

Deep learning (2012 –)







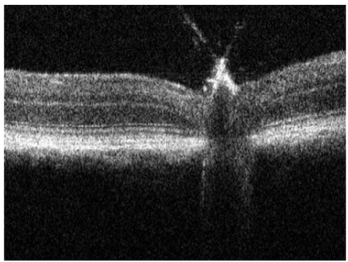
Starting from AlexNet in 2012 neural networks became mainstream in CV. A large number of models are deployed for video analytics, generative AI, self-driving, medical image analysis and other tasks

Outline

1. A short intro about your career and AI
2. What is computer vision?
3. Digital image
4. Light and color
5. Color models
6. Features of human vision

Image





Optical image is a visual representation of an object or scene formed through the use of optics, such as lenses or mirrors

Camera obscura

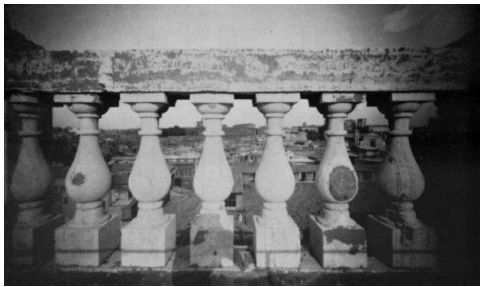


Simplest device for obtaining images. Basic principle known to Mozi (470 – 391 BCE) and Aristotle (384 – 322 BCE)

Was used as drawing aid for artists, described by Da Vinci (1452 – 1519)

Math model for camera obscura is perspective projection: beam of light passes through a focal point and forms an image on image plane

Perspective projection distortions



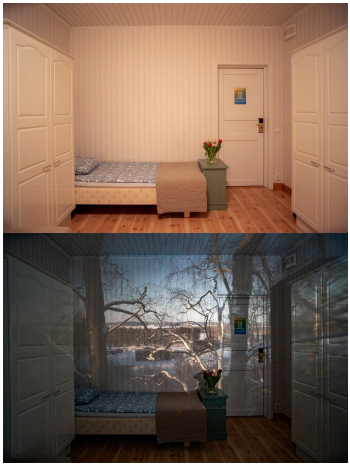


Perspective projection distorts object properties:

- columns near edges appear thicker,
- parallel lines may intersect on an image.

These distortions aren't due to optical system inaccuracies

Camera obscura in real life





Digital camera

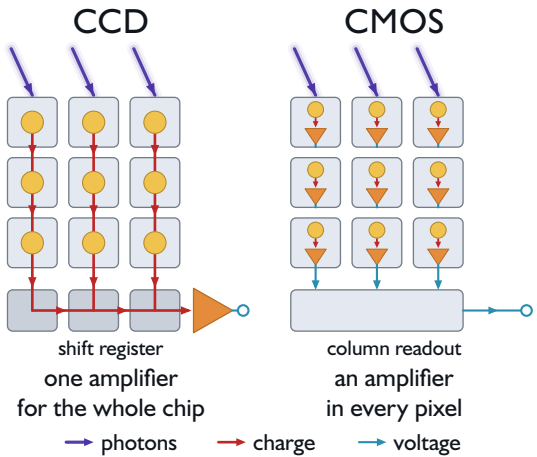


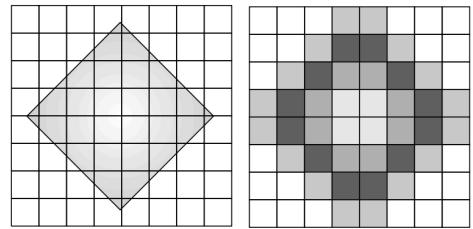




Camera is also modelled with perspective projection, but has lens, digital image sensor and other stuff

Image sensor





spatial and value discretization

Digital image



first digital image,
176 × 176 pixels

Digital grayscale image is a matrix with values in $[0, 1]$

0 means black (no light),
1 means white (maximum luminance)

Values are usually encoded with bytes (values $\{0, 1, \ldots, 255\}$) or with higher precision (10 – 16 bits)

Outline

1. A short intro about your career and AI
2. What is computer vision?
3. Digital image
4. Light and color
5. Color models
6. Features of human vision

What is color?



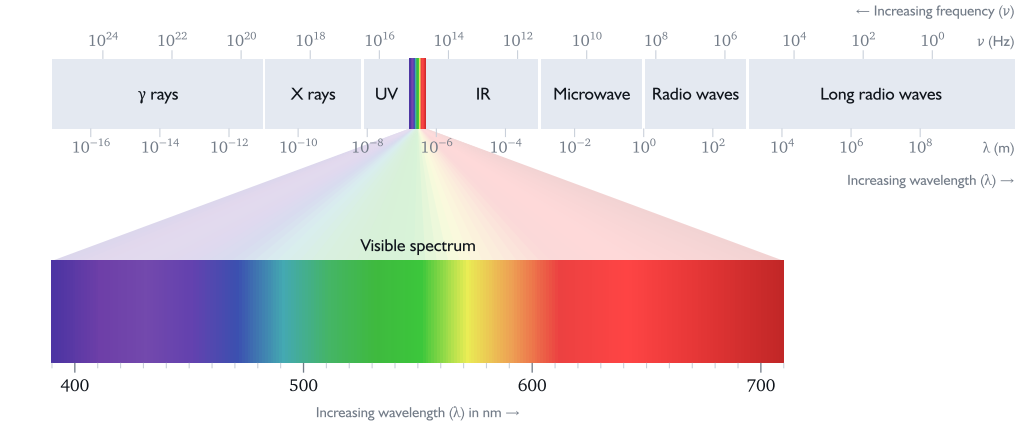
What is color?



Color is a psychophysiological property of our vision that occurs when observing objects and light, and not the physical properties of objects and light

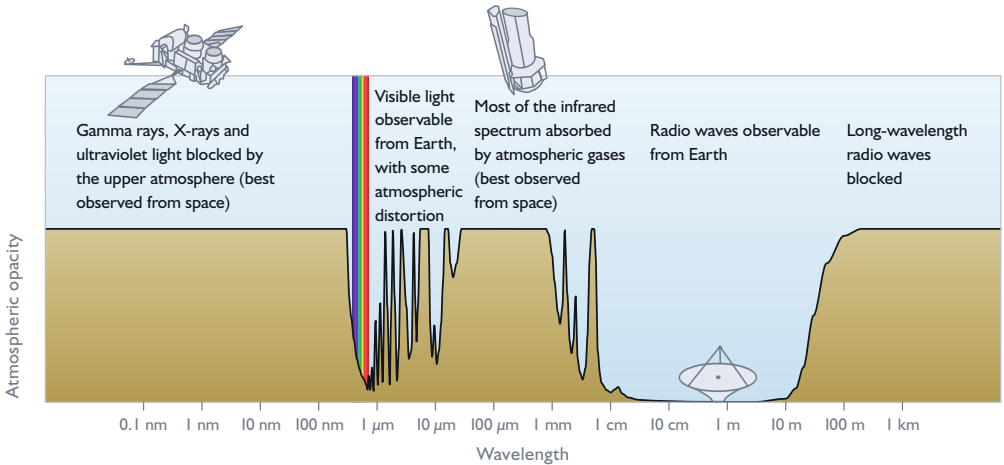
Light perception is studied by photometry and colorimetry

Visible light



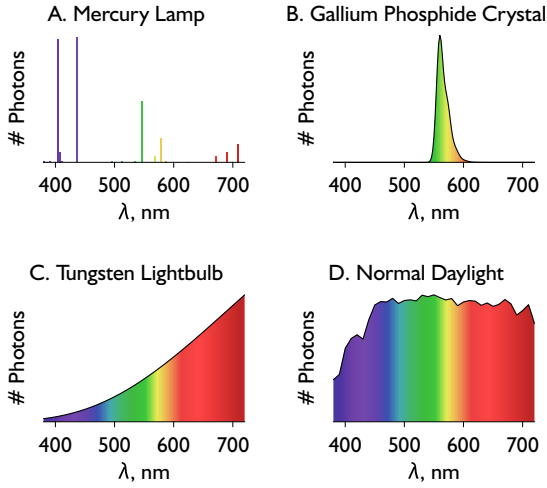
Visible light is electromagnetic radiation in range [380, 720] nm.

Visible light



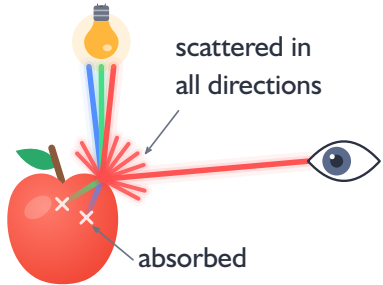
Visible light is electromagnetic radiation in range [380, 720] nm. It isn't blocked by earth's atmosphere

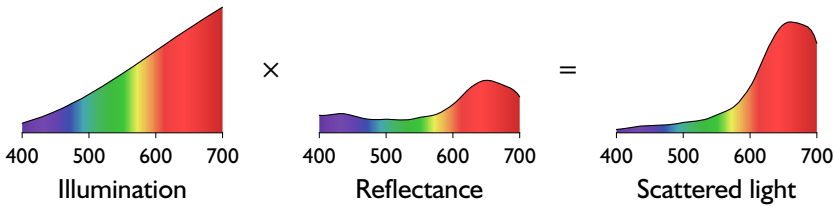
Light source spectrum



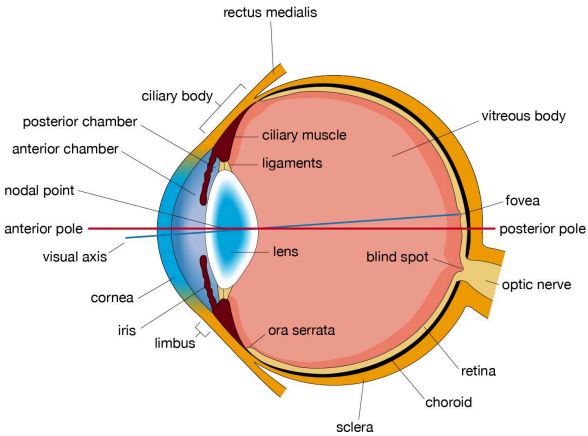
Any light source may be described by spectrum: the amount of energy emitted per unit of time for each wavelength

Interaction of light and objects





Human eye as camera

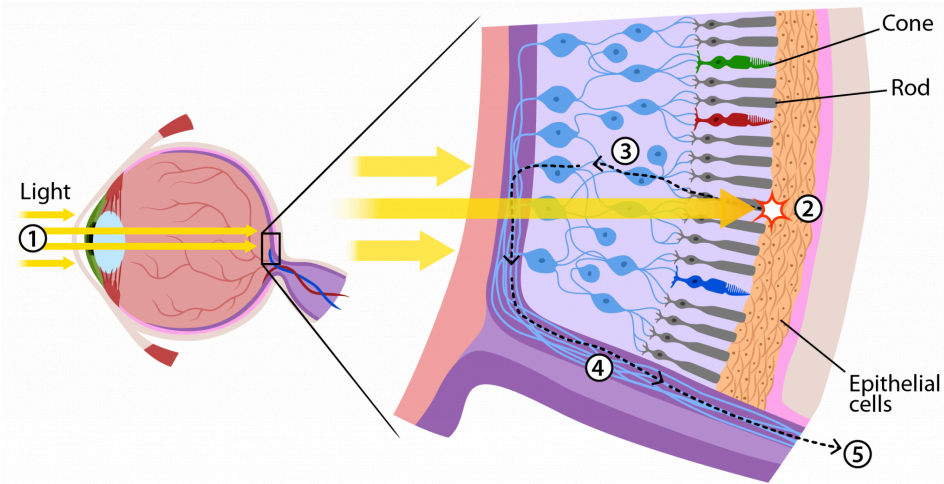


Iris — thin colored membrane with radial muscles, controls size of pupil and amount of incoming light

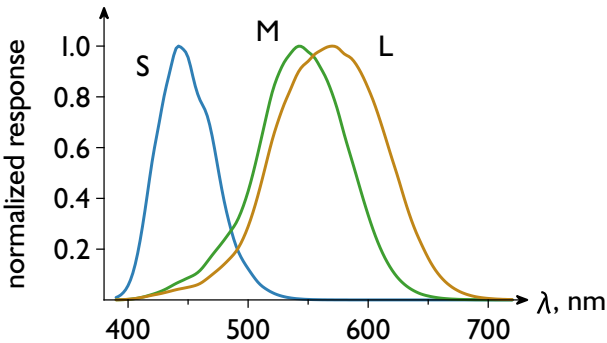
Lens size are changed using ciliary muscle (accommodation)

Light is perceived by cells on retina

Eye retina



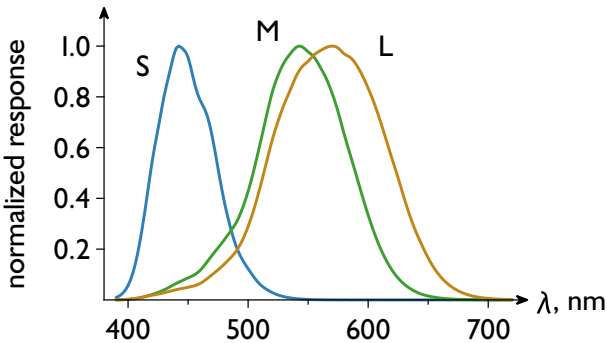
Color perception



Spectrum is multiplied by response function and integrated. Every type of cone gives 1 number

How can we describe whole spectrum using 3 numbers?

Color perception



Spectrum is multiplied by response function and integrated. Every type of cone gives 1 number

How can we describe whole spectrum using 3 numbers?

We can't. Two different indistinguishable spectra are called metamers

Color perception variants

1. Whales, dolphins, seals
2. Dogs, cats, horses, bulls
3. Bees (UV instead of red)
4. Birds and fish (+UV), rare women

16. Mantis shrimp



Color perception variants

1. Whales, dolphins, seals
2. Dogs, cats, horses, bulls
3. Bees (UV instead of red)
4. Birds and fish (+UV), rare women

16. Mantis shrimp



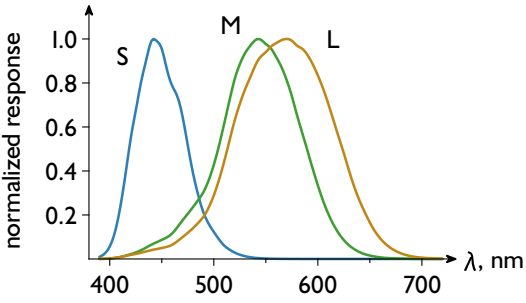
How to measure number of cones?

- histology
- microspectrophotometry
- electrophysiology
- genetics
- behavior experiments

Outline

1. A short intro about your career and AI
2. What is computer vision?
3. Digital image
4. Light and color
5. Color models
6. Features of human vision

LMS model



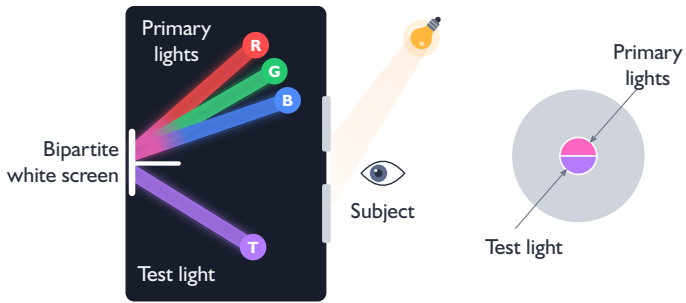
Can we use LMS responses for describing color in digital images?

- Fine for encoding and modelling human vision: chromatic adaptation, color blindness, etc.
- Not as display primaries: L and M are heavily correlated, there is no spectrum that excites only M and not L

Trichromatic color theory

Choose 3 primary colors p_1, p_2, p_3. Other colors are described as linear combinations of primary colors: $a = u_1 p_1 + u_2 p_2 + u_3 p_3$. 3 weights are 3 "coordinates" of color

Trichromatic theory states that all visible colors may be obtained as combination of 3 primary colors. It was studied in color matching experiments



Grassmann's laws (1853)

Empirical results about the perception of mixtures of colored lights:

- If two light sources have same weights for primary colors, they are perceived identically:

$$a = u_1 p_1 + u_2 p_2 + u_3 p_3, \quad b = u_1 p_1 + u_2 p_2 + u_3 p_3$$
$$a = b$$

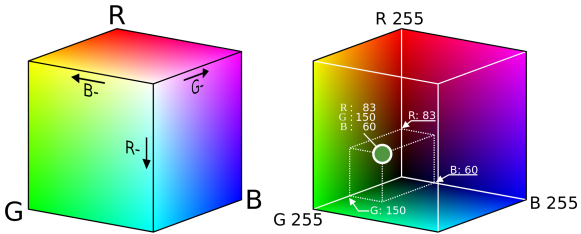
- Mixture of two light sources is perceived same as mixture of corresponding primary light sources:

$$a = u_1 p_1 + u_2 p_2 + u_3 p_3, \quad b = v_1 p_1 + v_2 p_2 + v_3 p_3$$
$$a + b = (u_1 + v_1) p_1 + (u_2 + v_2) p_2 + (u_3 + v_3) p_3$$

- Increasing light source intensity is the same as increasing intensity of primary light sources:

$$a = u_1 p_1 + u_2 p_2 + u_3 p_3$$
$$ka = (k u_1) p_1 + (k u_2) p_2 + (k u_3) p_3$$

CIE RGB space (1931)



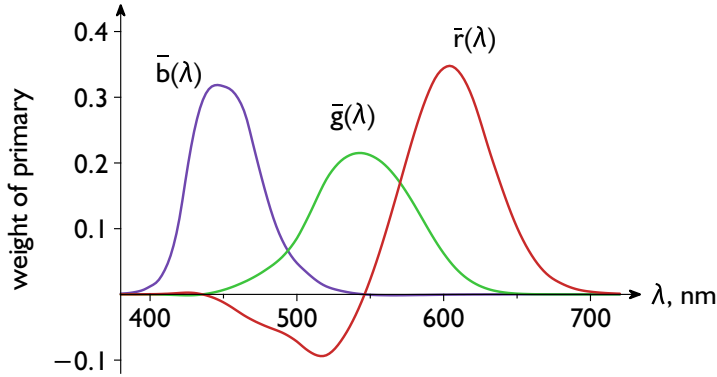
3 monochromatic primary colors, picked to be reproducible:

p_1 – 700 nm
p_2 – 546.1 nm
p_3 – 435.8 nm

p_2, p_3 are mercury lamp lines; at 700 nm hue barely changes with wavelength

Is RGB model able to reproduce all visible colors? Let's make color matching experiments for monochromatic light sources

Color matching functions

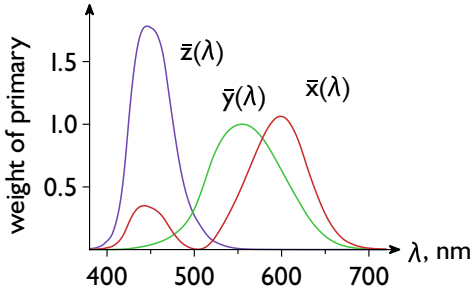


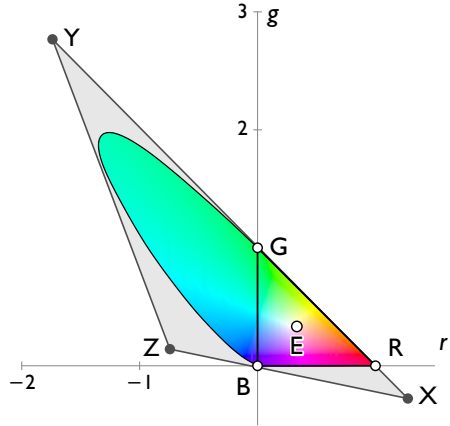
There exists colors that may not be reproduced with RGB primaries. We need to “subtract”, i.e. add primary light to the light that is reproduced

CIE XYZ color space (1931)

Drop intensity, keep chromaticity:
$r = R/(R+G+B)$, $g = G/(R+G+B)$.
Choose new linear basis XYZ:

- matching functions are non-negative
- Y is luminance
- $S(\lambda) = \text{const} \rightarrow X = Y = Z$

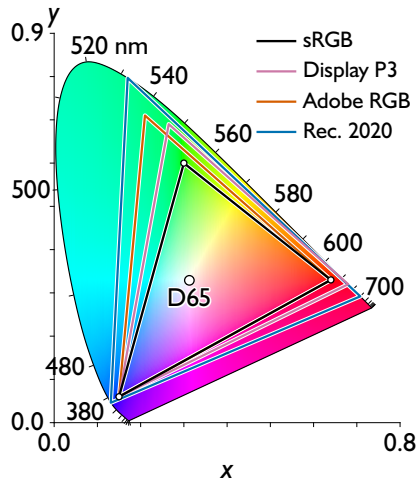




Color gamut

Plot xy-chromaticity diagram
$x = \frac{X}{X+Y+Z}$, $y = \frac{Y}{X+Y+Z}$:

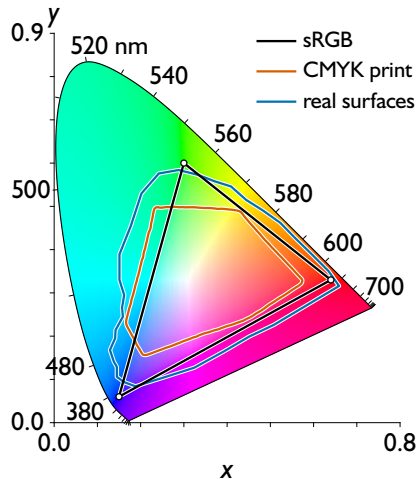
- monochromatic light sources are placed on the curve
- lower line: colors that can't be reproduced with monochromatic light sources
- no three real primaries cover all visible colors



Color gamut

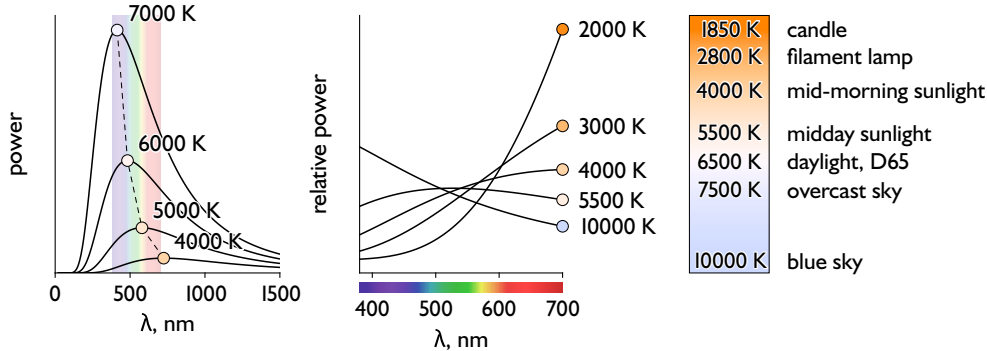
Plot xy-chromaticity diagram
$x = \frac{X}{X+Y+Z}$, $y = \frac{Y}{X+Y+Z}$:

- monochromatic light sources are placed on the curve
- lower line: colors that can't be reproduced with monochromatic light sources
- no three real primaries cover all visible colors



Black body radiation

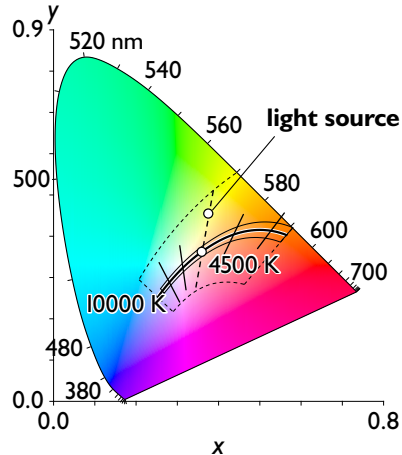
$$B(\lambda, T) = \frac{2hc^2}{\lambda^5} \cdot \frac{1}{e^{hc/(\lambda k T)} - 1}$$



Only the candle and the filament are hot bodies, and even the candle glows at about 1500 K: soot overemits the blue, so its light CCT is 1850 K

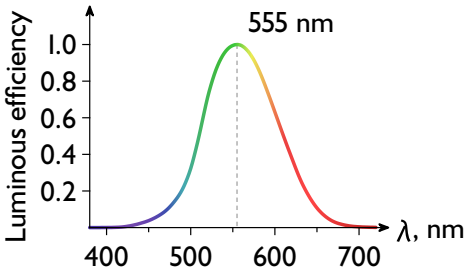
Correlated Color Temperature (CCT)

- black bodies lie on a curve in the xy plane, the Planckian locus
- CCT is the temperature of the nearest point of the locus in a wide band (dashed on the right)
- standard light sources lie in a narrow band (solid on the right)
- CCT is about light source on its own, color faithfullness of light on objects is measured with color rendering index (CRI)



Luminance and chrominance

Humans perceive light along two axes: luminance and chrominance (the latter splits into hue and saturation)

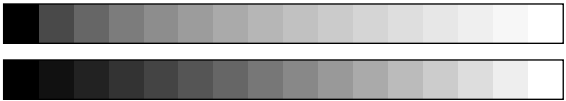


YC_bC_r is a linear color model with separate component for luminance and two color differences. Used in JPEG and digital video

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$$
$$C_b = 0.564 \cdot (B - Y)$$
$$C_r = 0.713 \cdot (R - Y)$$

Analog TV used relatives with the same Y: YUV (PAL) scales $B - Y$ and $R - Y$ differently, YIQ (NTSC) is UV rotated by 33°. "YUV" in software usually means YC_bC_r

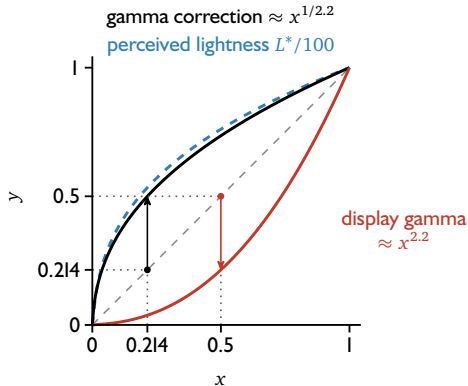
Gamma correction



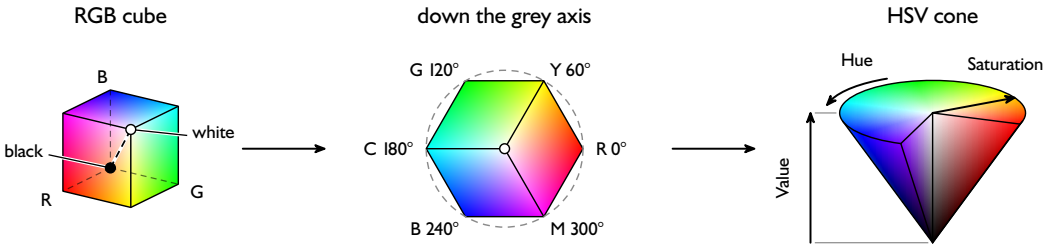
Physically uniform luminance

Subjectively uniform luminance

Luminance is perceived nonlinearly (dark regions are discriminated better than light regions), storing linear values is inefficient. To compensate that apply gamma correction $y = c \cdot x^{\gamma}$

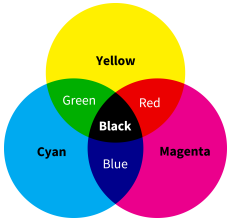


HSV



- $V = \max(R, G, B)$ is the position along the grey axis. V is **not** lightness
- H the angle around it and S is the distance out. Compared to YC_bC_r, the chromaticity coordinates are intuitive

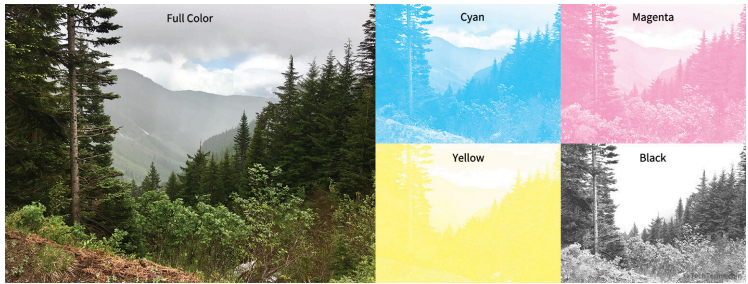
CMYK



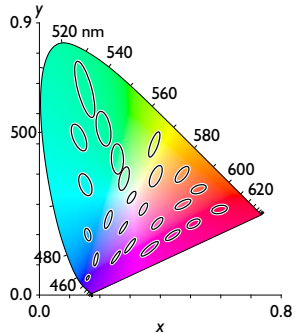
RGB is additive: the screen starts black and each primary adds its own light

CMY is subtractive: the paper starts white and each ink absorbs its complementary primary

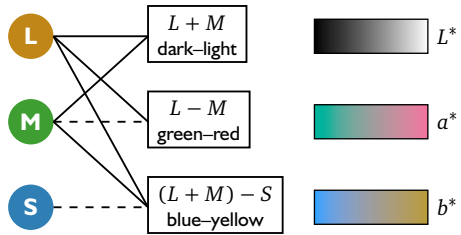
$$C = 1 - R, \quad M = 1 - G, \quad Y = 1 - B$$



Perceptually uniform color spaces



MacAdam ellipses: colors indistinguishable from center. Ellipses are small in blue, large in green: xy distance is not perceived difference



Opponent process: hues come in opponent pairs green–red, blue–yellow. Retina passes to brain one luminance and two cone-opponent signals

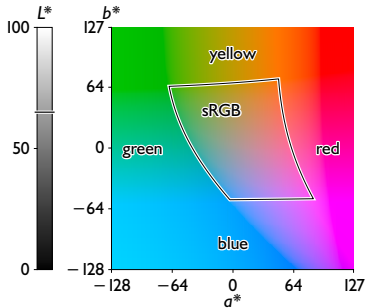
CIE Lab color space (1976)

$$L^* = 116 \cdot f(Y/Y_n) - 16$$
$$a^* = 500 \cdot (f(X/X_n) - f(Y/Y_n))$$
$$b^* = 200 \cdot (f(Y/Y_n) - f(Z/Z_n))$$
$$f(t) \approx \sqrt[3]{t}$$

- X_n, Y_n, Z_n – white point, e.g. D65
- Euclidean distance measures perceptual color distance better:



$\Delta E = 1$ only at the seam | $\Delta E = 2$ print tolerance | $\Delta E = 5$ | $\Delta E = 10$ different colors

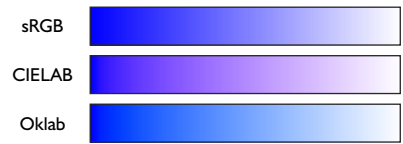


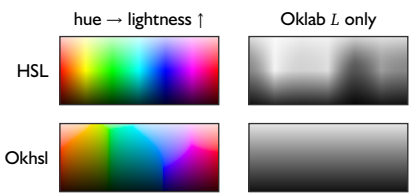
Oklab

Oklab compresses cone-like lms, not XYZ:

$$\begin{pmatrix} l \\ m \\ s \end{pmatrix} = M_1 \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \quad \begin{pmatrix} L \\ a \\ b \end{pmatrix} = M_2 \begin{pmatrix} \sqrt[3]{l} \\ \sqrt[3]{m} \\ \sqrt[3]{s} \end{pmatrix}$$

- $M_1 - XYZ$ (D65) to cone-like LMS
- $M_2 - lms$ to lightness $L \in [0, 1]$ and opponent axes a, b
- M_1, M_2 are fitted to perception data: lightness, chroma and hue change independently



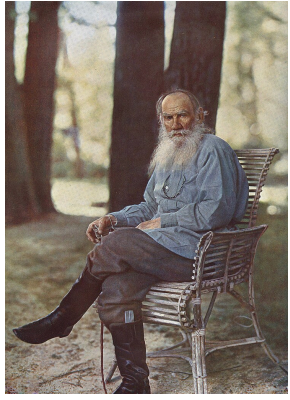


B. Ottosson, 2020. bottosson.github.io/posts/oklab

First color photos



Sergey Prokudin-Gorsky

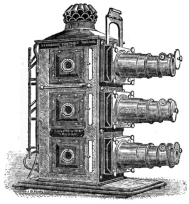


Leo Tolstoy

Film about Prokudin-Gorsky: Parfenov. The color of the nation

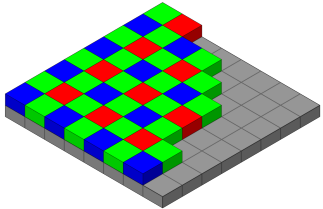
First color photos

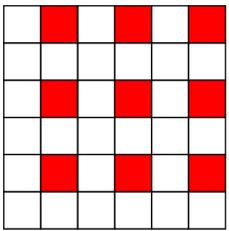


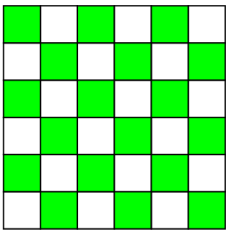


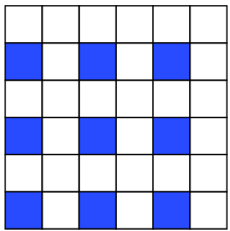


Digital color image





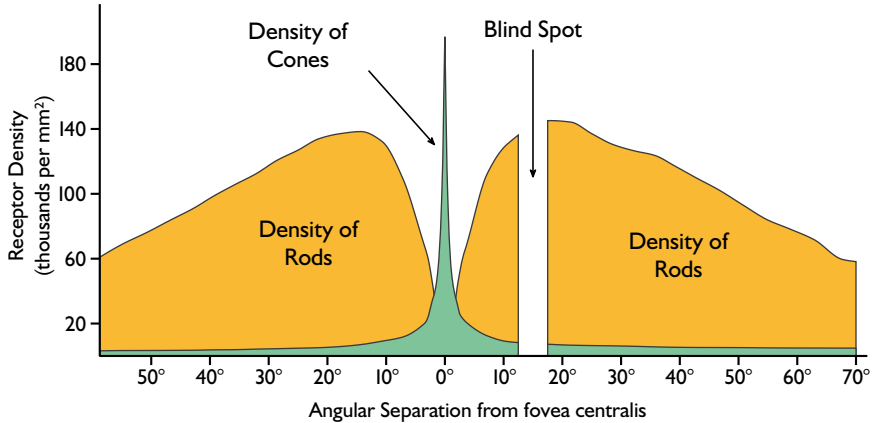




Outline

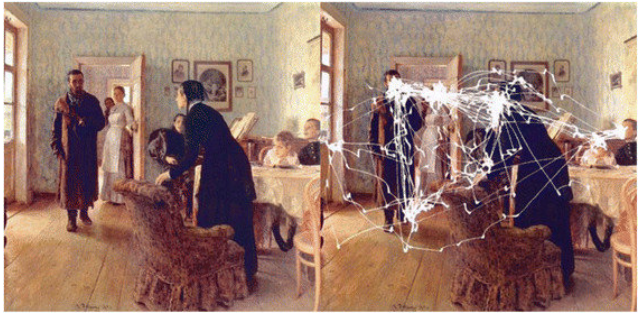
1. A short intro about your career and AI
2. What is computer vision?
3. Digital image
4. Light and color
5. Color models
6. Features of human vision

Density of rods and cones



Rods and cones are distributed unequally on retina. Most cones are in fovea — small area (1 or 2 degree of visual angle)

Eye tracking



Viewing scenarios





1

Free examination.



2

Estimate material circumstances of the family



3

Give the ages of the people.



4

Surmise what the family had been doing before the arrival of the unexpected visitor.



5

Remember the clothes worn by the people.



6

Remember positions of people and objects in the room.

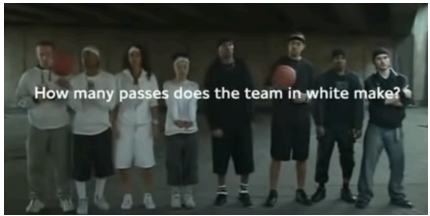


7

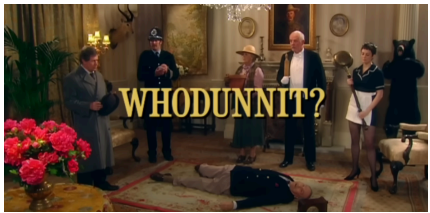
Estimate how long the visitor had been away from the family.

3 min. recordings of the same subject

Awareness videos



Basketball players



Whodunnit?

Color and luminance constancy



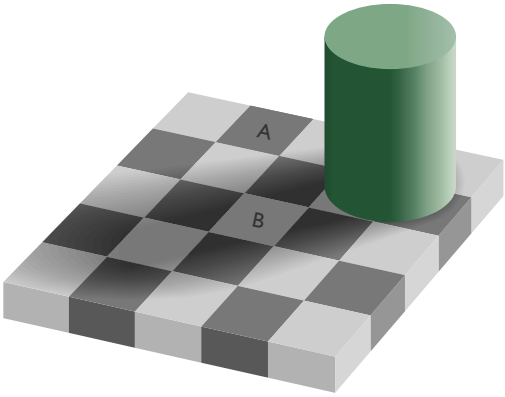
J. Sargent. The Daughters of Edward Darley Boit. 1882

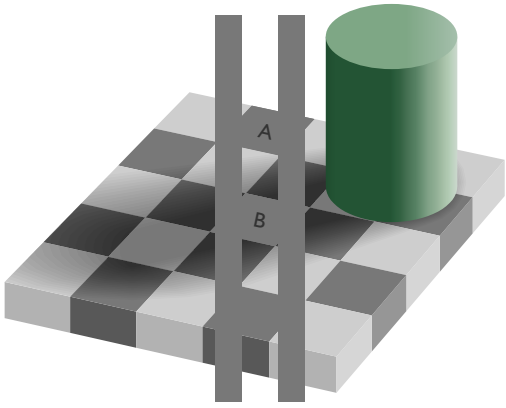
Human vision system is able to adapt to different lighting conditions

Luminance:

- full moon – 0.27 lux
- cloudy day – 100 – 1000 lux
- bright day (in shadow) – 20000 lux
- under the sun – 100000 lux

Checkerboard illusion





Conclusion

We reviewed following topics:

- definition of computer vision
- how human eye and photo camera work
- trichromatic theory of color
- color models LMS, RGB, XYZ, CMYK, YCbCr, HSV, Lab
- features of human vision