Introduction to computer vision
Digital image, light and color

Vlad Shakhuro

18 September 2025

Outline

1. What is computer vision?

2. Digital image

3. Light and color

4. Color models

5. 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 Artifial Intelligence (AI)

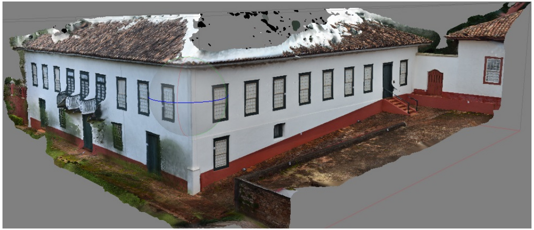
Turing test for CV: answer any question about the image that can be answered by a human

Object detection task

Image and object attributes recognition

Photogrammetry

→



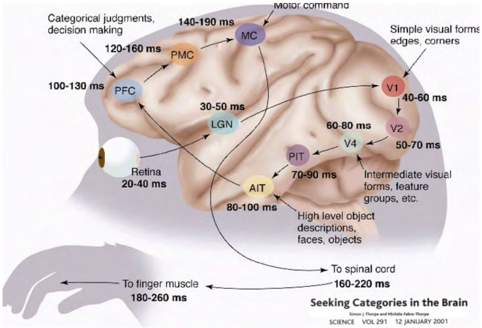
3D reconstruction and measurements from photos

Human vision

We match observations with our knowledge

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

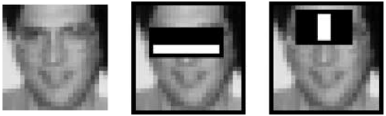
Works of David Marr (1970s)

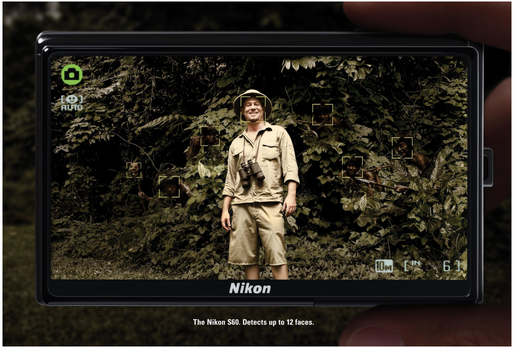


According to Marr, human vision is done in several steps:

- 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

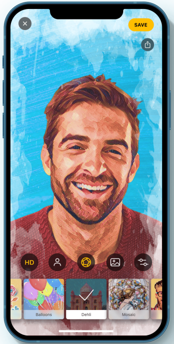
Viola-Jones face detector (2001)

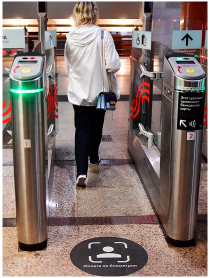




One of the first real-life applications of computer vision: face detection model that can be computed on a mobile CPU

Deep learning (2012 –)



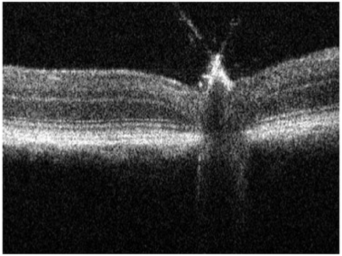


Starting from AlexNet in 2012 neural networks became mainstream in CV. A large number of models are deployed for video analytics, generative AI, self-driving, medical image analysis and other tasks

Outline

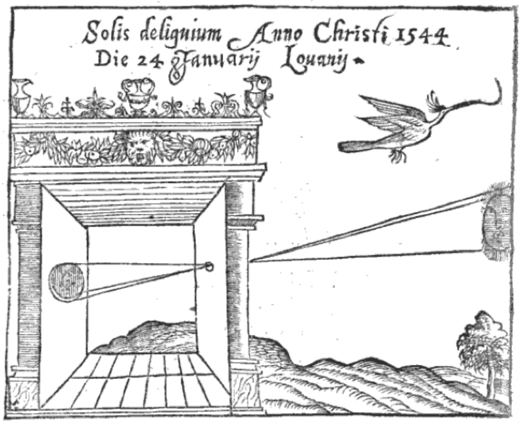
1. What is computer vision?
2. Digital image
3. Light and color
4. Color models
5. Features of human vision

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

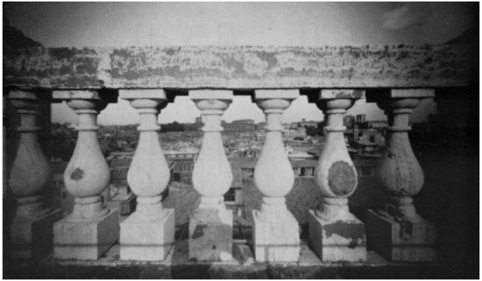


Simplest device for obtaining images. Basic principle known to Mozi (470 – 391 BCE) and Aristotle (384 – 322 BCE)

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

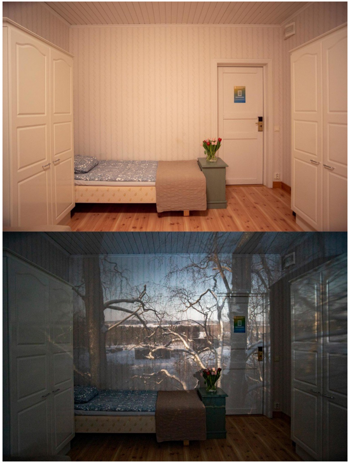
Camera obscura as dimension reduction

What information is lost by taking photo?

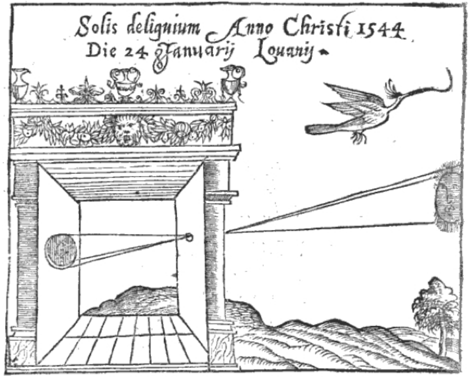
Camera obscura as dimension reduction

What information is lost by taking photo?
And what information is preserved?

Camera obscura in real life

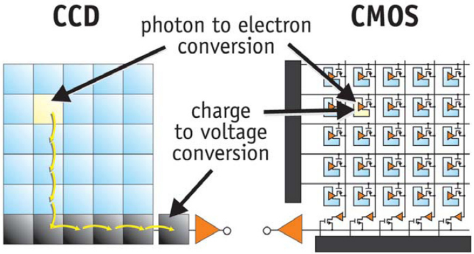


Digital camera

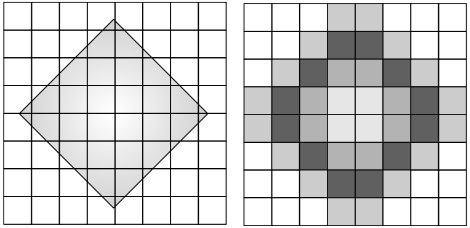


Camera is also described with perspective projection, but has lens, digital image sensor and other stuff

Image sensor



CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node. CMOS imagers convert charge to voltage inside each pixel.



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, …, 255}) or with higher precision (10–16 bits)

Outline

1. What is computer vision?

2. Digital image

3. Light and color

4. Color models

5. Features of human vision

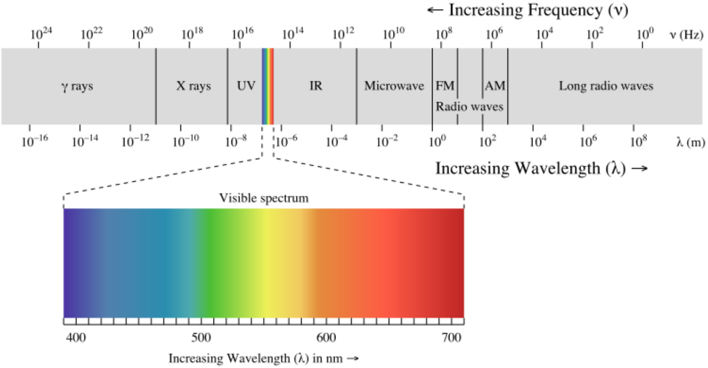
What is color?

What is color?

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

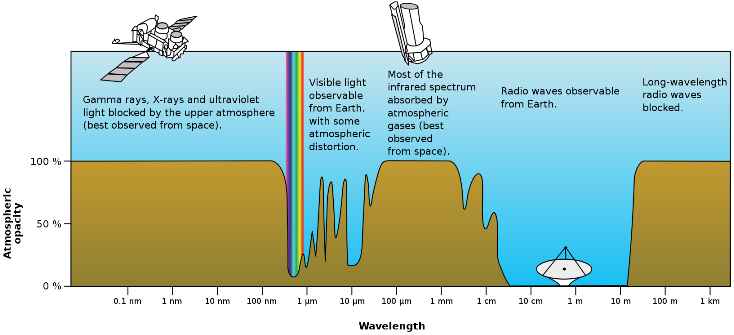
Light perception is studied by photometry and colorimetry

Visible light



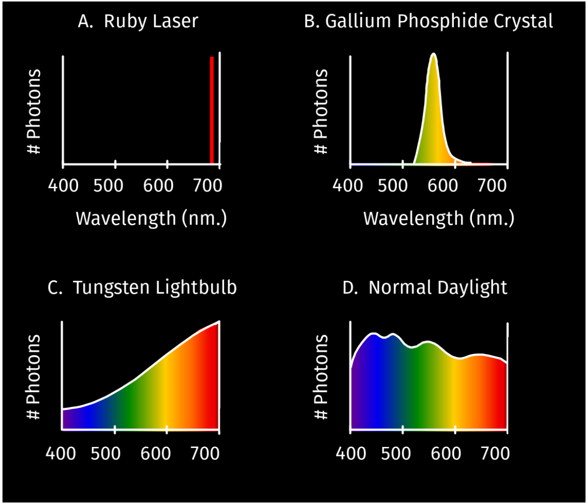
Visible light is electromagnetic radiation in range [380, 780] nm.

Visible light



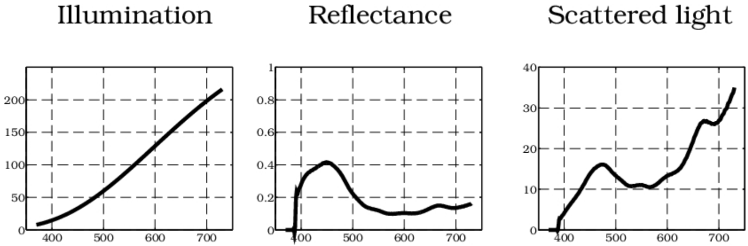
Visible light is electromagnetic radiation in range [380, 780] nm. It isn't blocked by earth's atmosphere

Light source spectrum

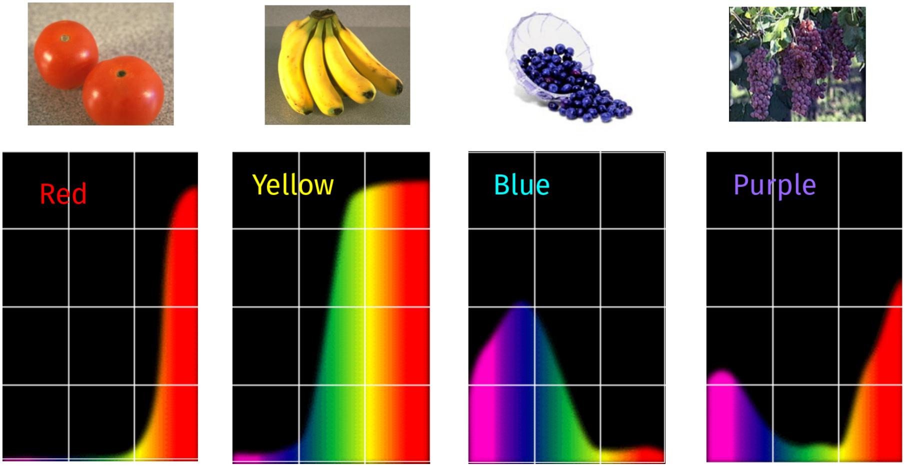


Any light source may be described by spectrum: the amount of energy emitted per unit of time for each wavelength

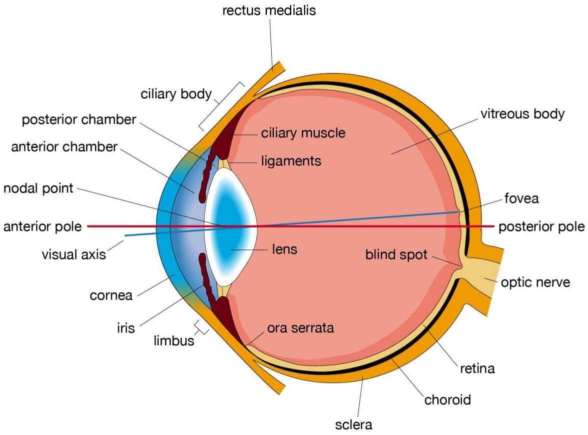
Interaction of light and objects



Example spectra of reflected light



Human eye as camera

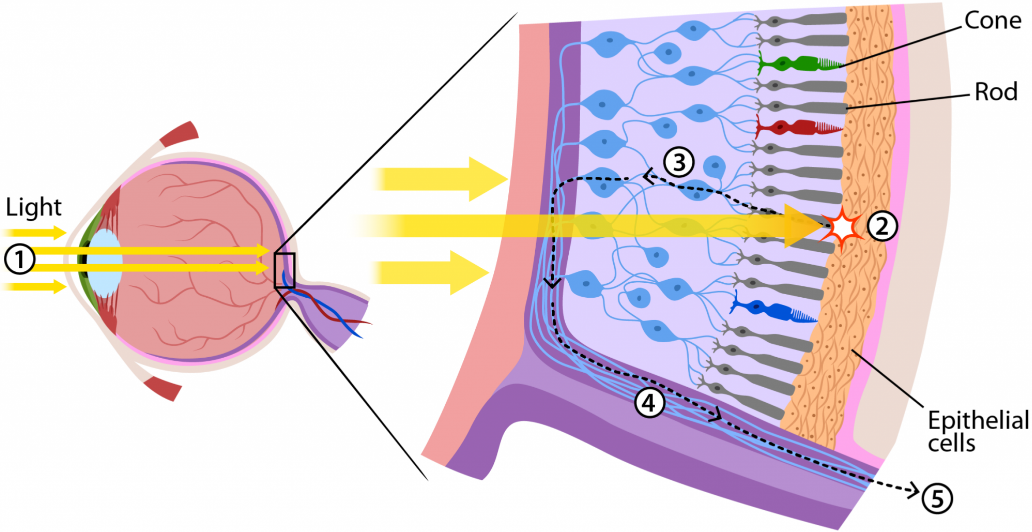


Iris — thin colored membrane with radial muscles, controls size of pupil and amount of incoming light

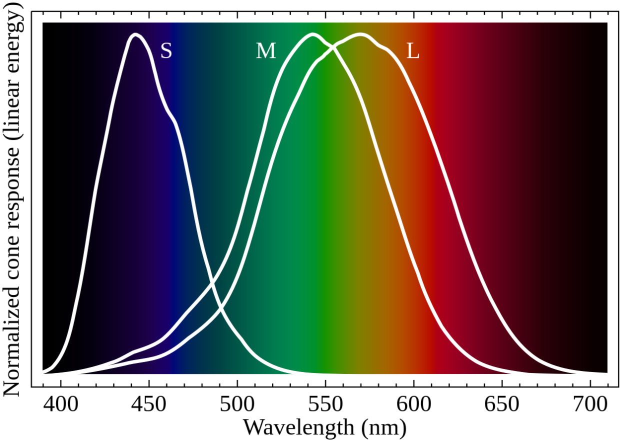
Lens size are changed using ciliary muscle (accomodation)

Light is perceived by cells on retina

Eye retina



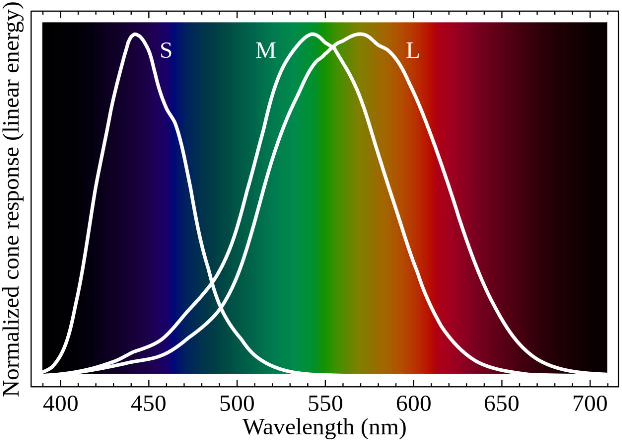
Color perception



Spectrum is multiplied by reponse 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 reponse 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

Outline

1. What is computer vision?

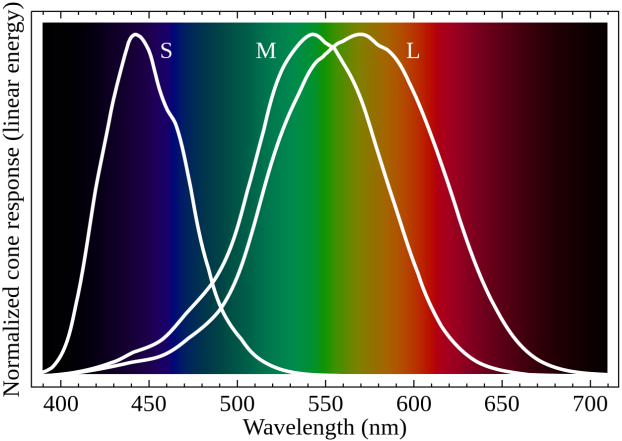
2. Digital image

3. Light and color

4. Color models

5. Features of human vision

LMS model



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

No:

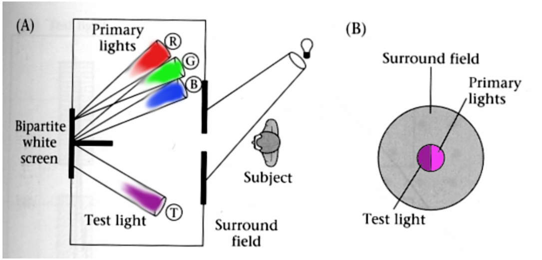
LMS is inconvenient model: colors with M > 0 and S = L = 0 don't exist

It's impossible to make a display that works in LMS space

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:

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

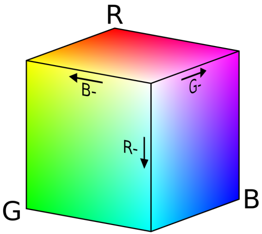
- Mixture of two light sources is percepted same as mixture of correspoding primary light sources:

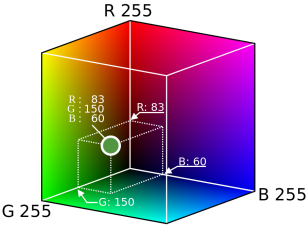
$$a = u_1 p_1 + u_2 p_2 + u_3 p_3, \quad b = v_1 p_1 + v_2 p_2 + v_3 p_3$$
$$a + b = (u_1 + v_1) p_1 + (u_2 + v_2) p_2 + (u_3 + v_3) p_3$$

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

$$a = u_1 p_1 + u_2 p_2 + u_3 p_3$$
$$ka = (ku_1) p_1 + (ku_2) p_2 + (ku_3) p_3$$

CIE RGB space (1931)





3 monochromatic primary colors (max response of cones):

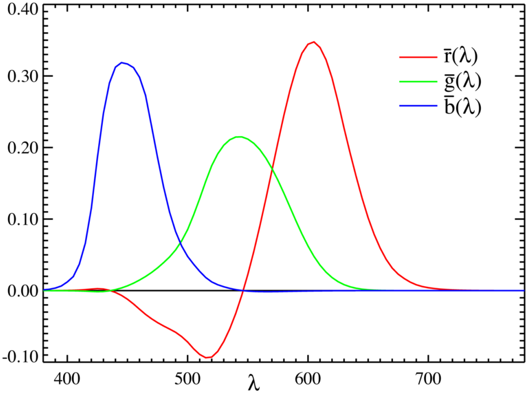
$p_1 - 645.2\ nm$

$p_2 - 525.3\ nm$

$p_3 - 444.4\ nm$

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

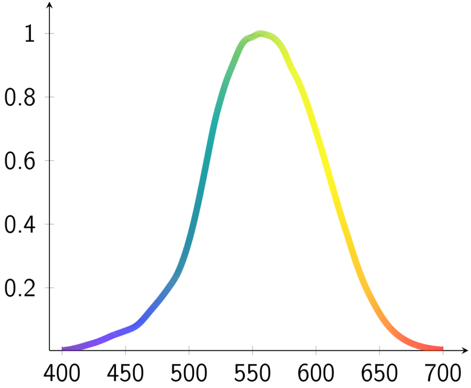
Color matching function



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

Luminance and chrominance

Intuitively human denotes two characteristics of light: luminance and chrominance (may be further divide into hue and saturation)



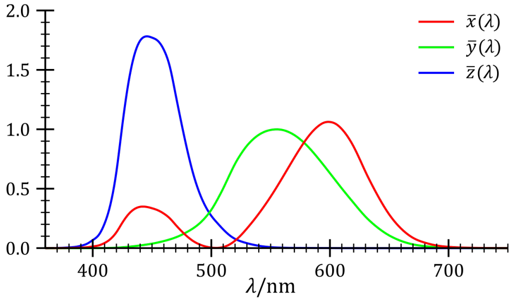
YIQ is a linear color model with separate component for luminance. Used in color television and JPEG image standard

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B$$
$$I = 0.596 \cdot R - 0.275 \cdot G - 0.321 \cdot B$$
$$Q = 0.212 \cdot R - 0.528 \cdot G + 0.311 \cdot B$$

CIE XYZ space (1931)

Linear additive model, a standard reference against which many other color spaces are defined

- covers all visible colors (matching functions are non-negative)
- Y is for perceived luminance
- XZ are for chrominance components
- X, Y, Z values are in $[0; +\infty]$
- points $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ are imaginary primary colors



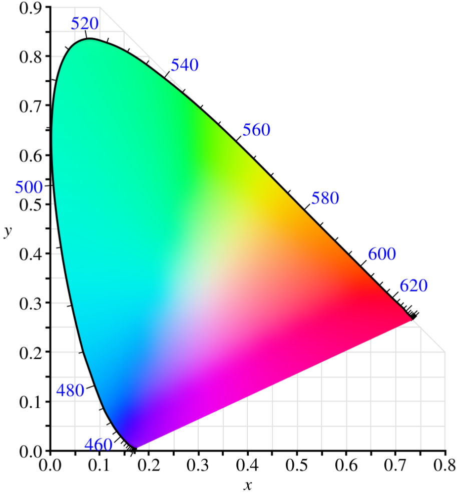
CIE xyY and color gamut

We'll describe chromaticity with two values: $x = \frac{X}{X+Y+Z}$, $y = \frac{Y}{X+Y+Z}$

Plot color diagram for $x, y \in [0,1]^2$

Observations:

- coherent light sources are placed on the curve
- lower line correspond to colors that can't be reproduced with coherent light sources
- visible lights can't be covered with any three real primary light sources



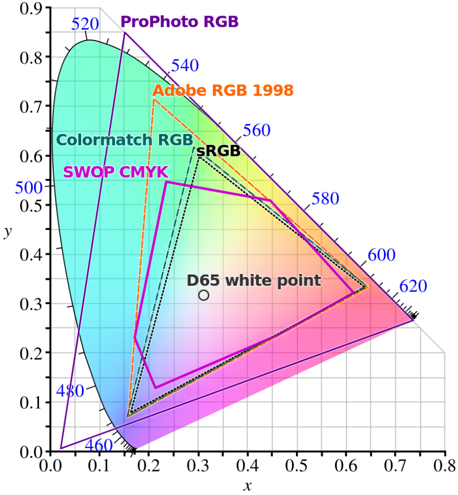
CIE xyY and color gamut

We'll describe chromaticity with two values: $x = \frac{X}{X+Y+Z}$, $y = \frac{Y}{X+Y+Z}$

Plot color diagram for $x, y \in [0, 1]^2$

Observations:

- coherent light sources are placed on the curve
- lower line correspond to colors that can't be reproduced with coherent light sources
- visible lights can't be covered with any three real primary light sources

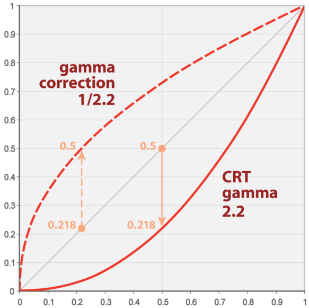


Gamma correction

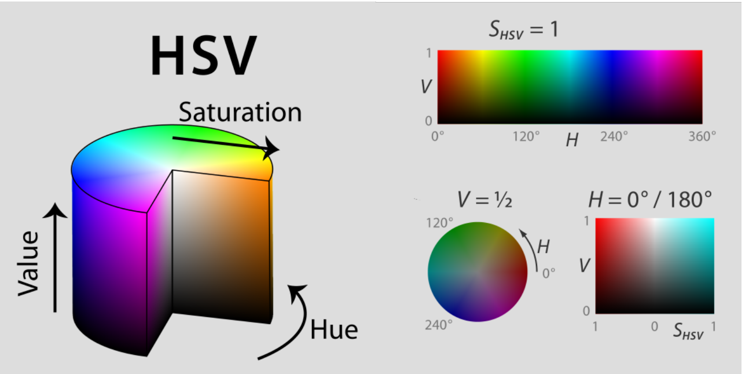
Physically uniform luminance

Subjectively uniform luminance

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

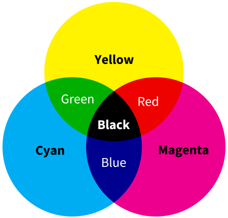


HSV color space



Compared to YIQ, this space has more intuitive chromaticity coordinates: Hue and Saturation

CMYK subtractive model

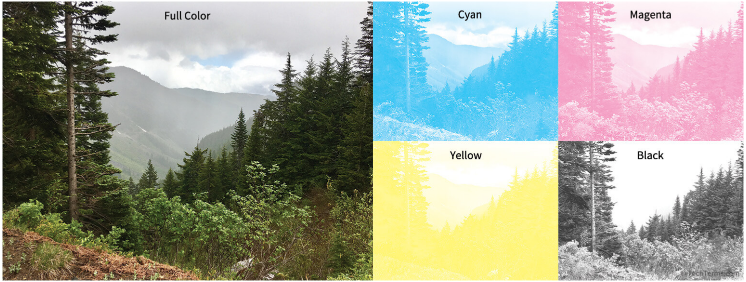


RGB is additive color model (light is emmited)
CMYK is subtractive model (light is reflected)

$$C = G + B$$

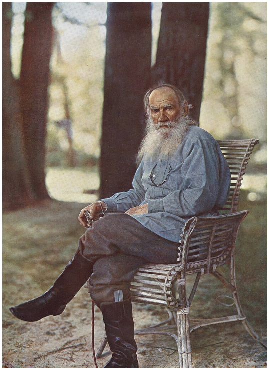
$$M = R + B$$

$$Y = R + G$$



First color photos

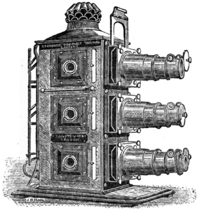
Sergey Prokudin-Gorsky



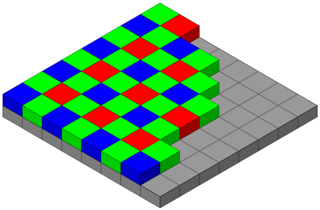
Leo Tolstoy

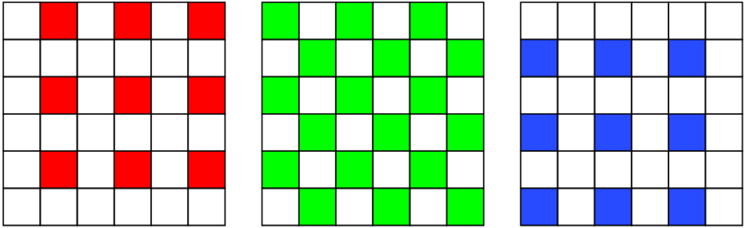
Film about Prokudin-Gorsky: Parfenov. The color of the nation

First color photos



Digital color image

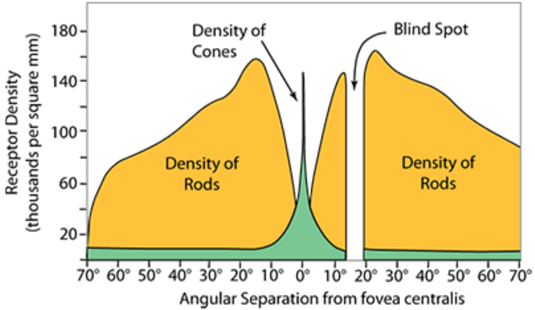




Outline

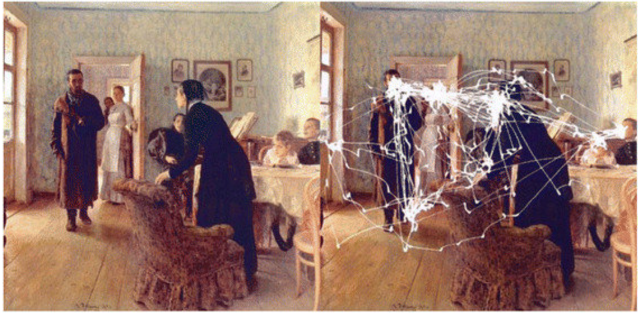
1. What is computer vision?
2. Digital image
3. Light and color
4. Color models
5. Features of human vision

Density of rods and cones

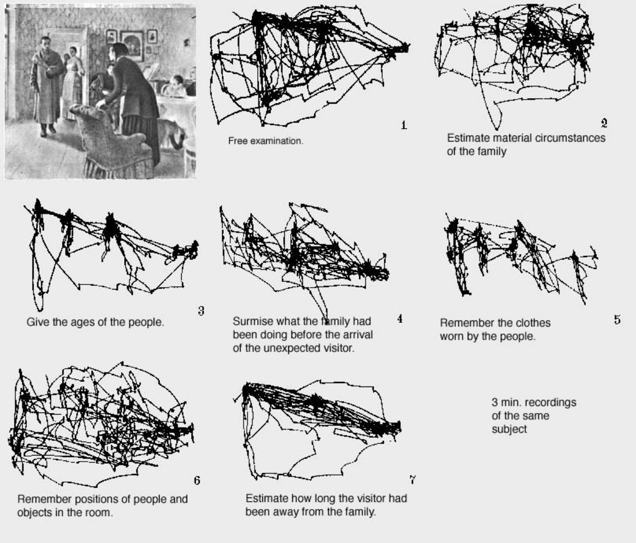


Rods and cones are distributed inequally on retina
Most cones are in fovea — small area (1 or 2 degree of visual angle)

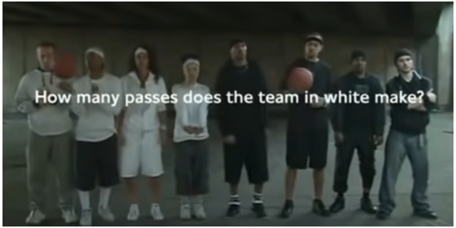
Eye tracking



Viewing scenarios



Awareness videos



Basketball players

Whodunnit?

Color and luminance constancy

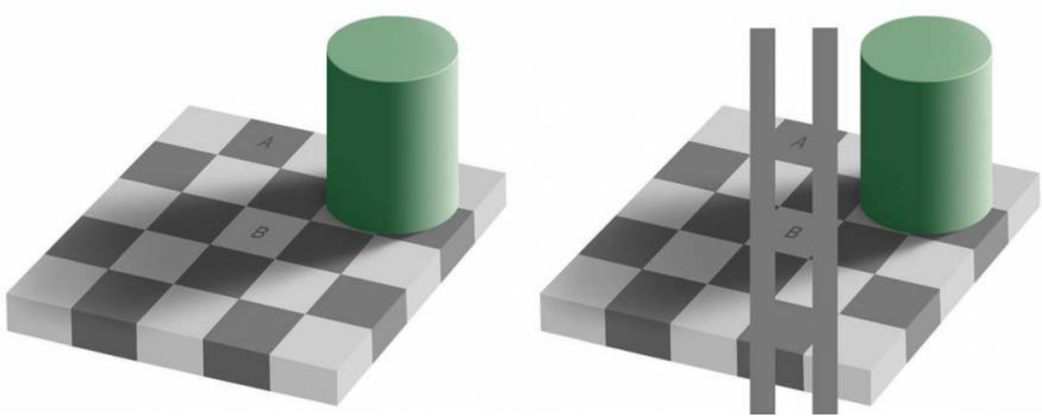
J. Sargent. The Daughters of Edward Darley Boit. 1882

Human vision system is able to adapt to different lighting conditions

Lumninance:

- 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, YIQ, HSV
- features of human vision