UNIVERSITÄT
ZU KÖLN

BOZEK LAB
DATA SCIENCE OF BIOIMAGES

Computer Vision

Prof. Dr. rer. nat. Katarzyna Bozek

UNIVERSITÄT ZU KÖLN

BOZEK LAB
DATA SCIENCE OF BIOIMAGES

Course Information

Course Information

- What the course is about
  - Computer Vision from Algorithms to Deep Learning
  - Starting with history and background information
  - Practical oriented course
- Timeplan
  - 15 Weeks, but there are also some holidays
  - Exception: lecture on 17.04
  - In week 3 (23.04. – 25.04), there will be no lecture (Lab retreat)

Course Information

- Homework
    - Four assignments in total
    - You can start them in the practical blocks on thursdays and ask questions
    - Each assignment takes 2–3 weeks
- Journal Club
    - Discussions of given papers (in groups)
- Exam
    - Final exam will be a group project with presentation and documentation
- Grading
    - Project presentation and documentation
    - Participation in journal clubs

Used Books

In the beginning of some chapters you will find a list of book sections relevant to the unit.
All abbreviations are explained here:

Main resources

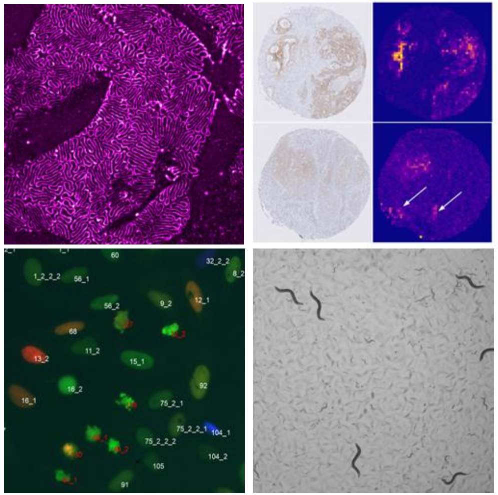
FP: Forsyth and Ponce. Computer vision: a modern approach. 2nd edition.
SR: Richard Szeliski. Computer Vision: Algorithms and Applications. 2nd edition
GW: Gonzalez-Woods. Digital image processing. 4th edition

Supplemental material

DL: https://www.deeplearningbook.org/
SL: A Cookbook of Self-Supervised Learning
GP: Bichot, Siarry. Graph Partitioning
DIP: Burger and Burge. Digital Image Processing
DV: David Vernon: Machine Vision. Automated Visual Inspection and Robot Vision

Who are we...

- me: Kasia Bozek
- Professor for Data-Analytics in Bioinformatics
- Institute of Biomedical Informatics, Faculty of Medicine



- Bozek Lab
  - Develop and apply deep learning (DL) methods
  - Biomedical image data
    - cancer histopathology, microscopy images of podocytes, c. elegans behavior analysis, cell death prediction, …
  - https://bozeklab.com/

Teaching Assistants (TAs)

- PhD students from Bozek Lab
- Will do the practical sessions, correct the homework, …
- If you have questions regarding the course, homework, etc. contact: maurice.deserno@uni-koeln.de
- Put CV-Course at the beginning of the email subject

Florian

Juan

Maurice

Introduction

Material:

- SR 1.1 – 1.2, 2.3.2
- GW 1.3, 2.1 – 2.2
- FP 1.1 – 2, 3.1–4

The power of human vision 🌈

- What do you see?
  - A bent plant and some pots
  - A fluffy chair with a pillow
  - Letters, even a word you can read
  - Carpets on a wood(-like) floor
  - Located in a 3D room
  - …
- All this you can extract from a 2D-RGB image 🤯
- How do you know all this?
- How to program a computer to "see" all this?

The power of human vision 🌈

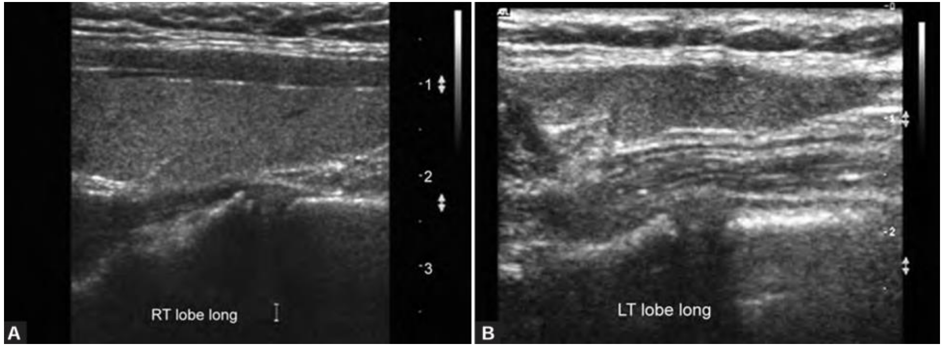


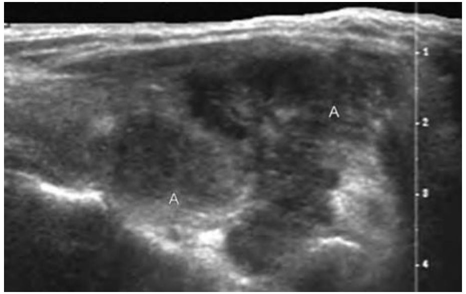


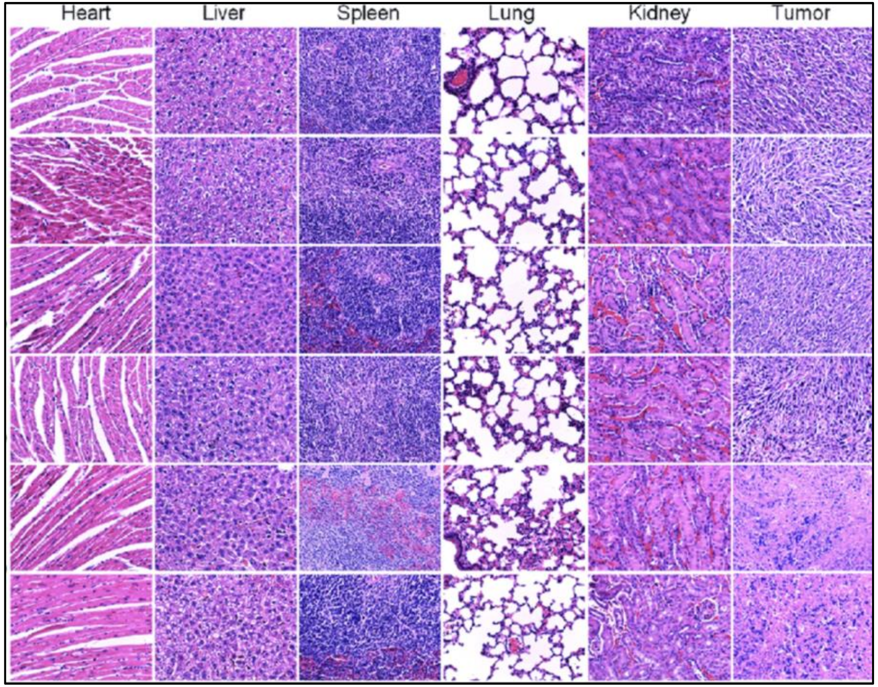





The power of human vision 🌈







https://www.aijoc.com/doi/AIJOC/pdf/10.5005/jp-journals-10003-1148

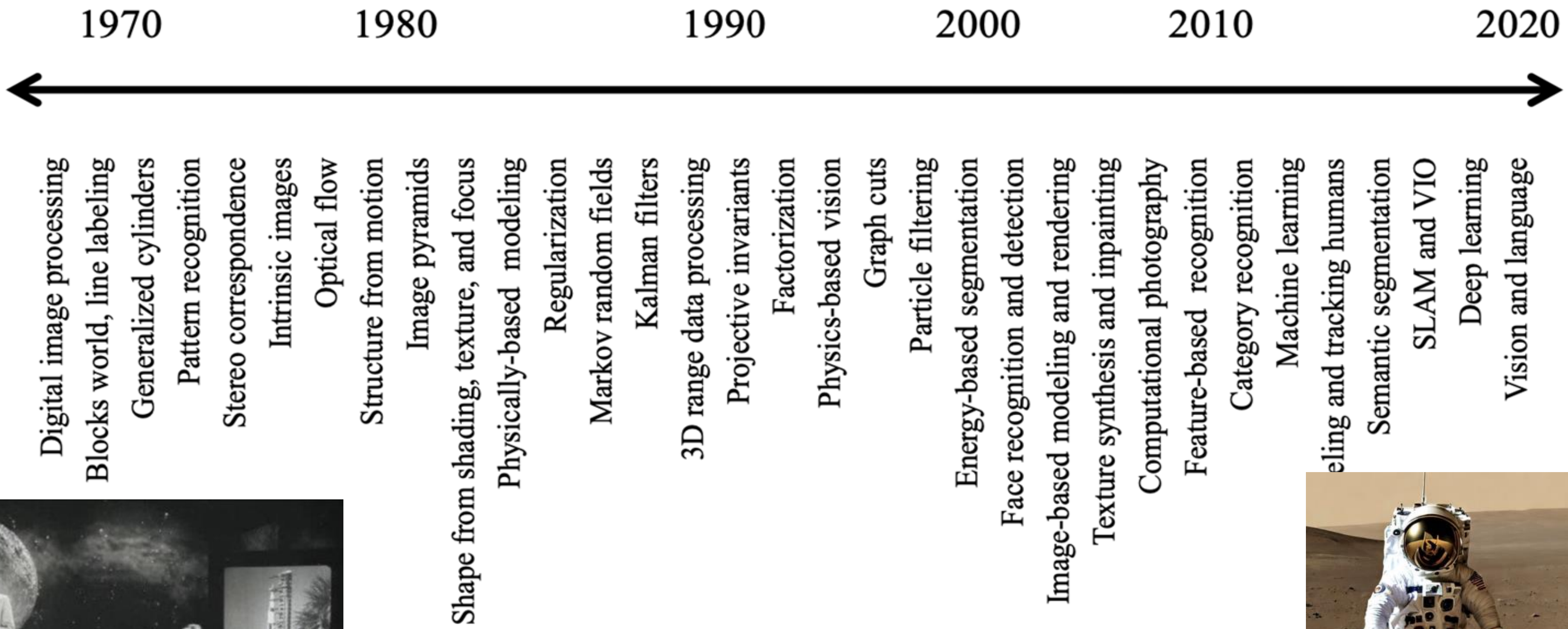
https://pubs.acs.org/doi/10.1021/acs.nanolett.6b02786

What is Computer Vision?

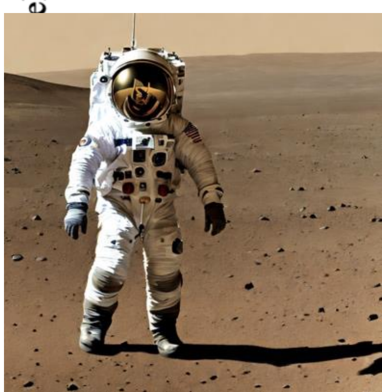
- Meaning changed over time
- Mimicking human vision in the 1960s – Today the field is quite broad

- Short: Enable a computer to see
- Include methods for:
    - acquiring
    - processing
    - analyzing and understanding digital images
    - extraction of high-dimensional data
    - produce numerical or symbolic information

What is Computer Vision?



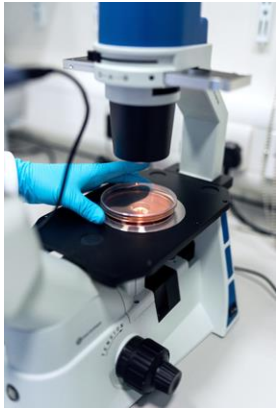
https://www.scienceandmediamuseum.org.uk/objects-and-stories/moon-to-living-room-apollo-11-broadcast



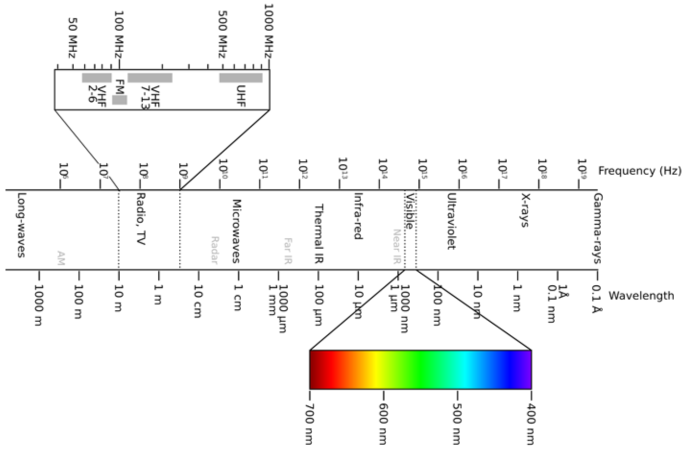
https://huggingface.co/spaces/stabilityai/stable-diffusion

Applications

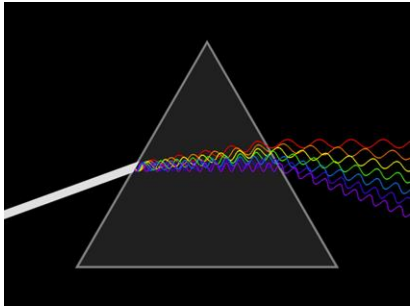
- Self-driving vehicles
- Robotics
- Bio-Medical Imaging
- Optical Character Recognition
- Retail/Warehouse-logistics
- Image/Video generation
- Surveillance
- … many more

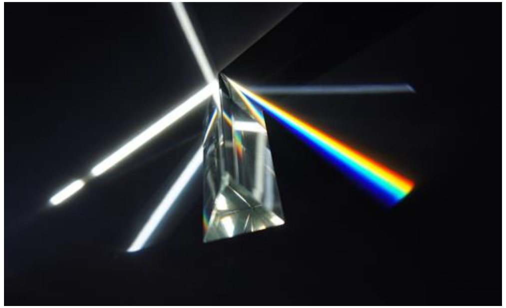


Electromagnetic Spectrum ⚡



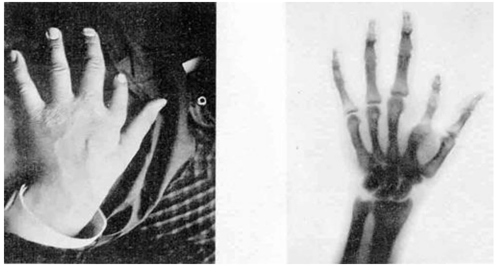
- Isaac Newton first used the term Spectrum
- Principal energy source for images (nowadays)
- Alternatives sources: sound (eg. ultrasonic), synthetic data, …
- Most known are images from the human visible range
- But other ranges are also used (see next slide)
- Imagine it as sinusoidal waves with a specific wavelength λ





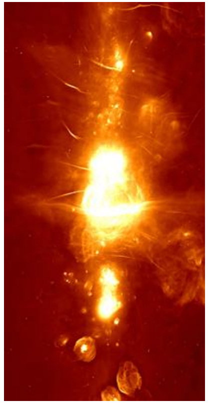
Dispersive Prizm
“splitting” white light

Some examples



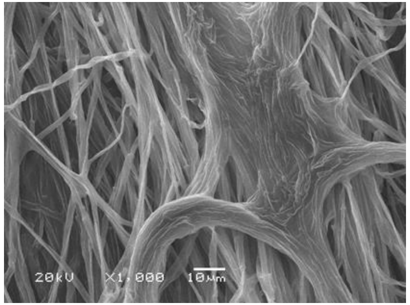
X-ray of a hand

Different infrared ranges

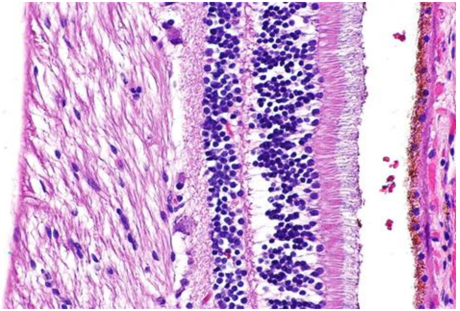


Radio image of the Milky Way

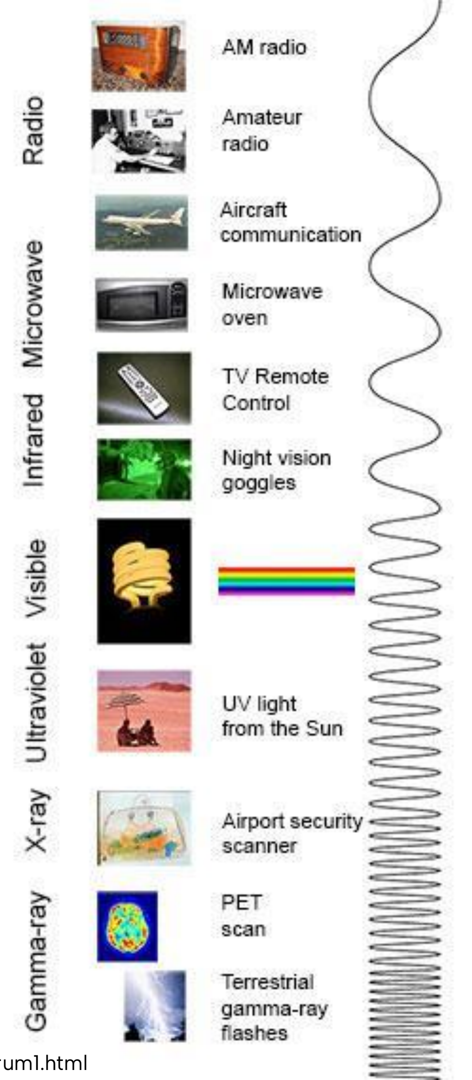
Ultrasonic image of a healthy 18 weeks old female



Mycelial film under a scanning electron microscope



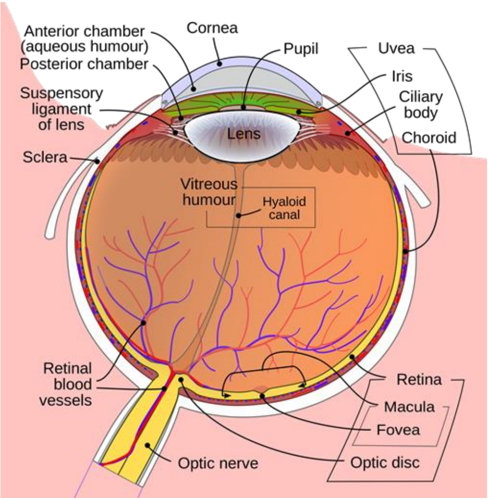
Part of the retina under a microscope with H&E staining

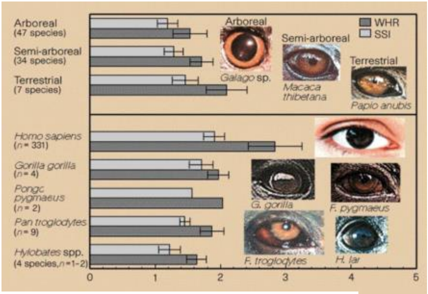


https://imagine.gsfc.nasa.gov/science/toolbox/emspectrum1.html

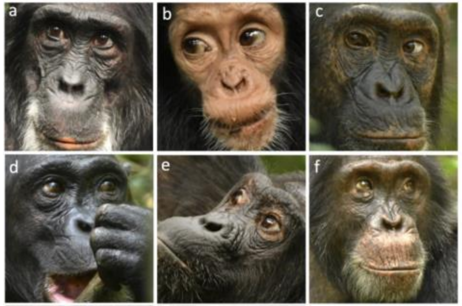
Human Eyes

- Powerful sensory organ
- Only covered roughly
- Important parts of the eye:
  - Pupil: Controlling the amount of light hitting the lens
  - Cornea + Lens: Refracting the light, forming an image on the retina
  - Retina: Contains the photoreceptors, reacting to the light
- Shape of the lens varies to achieve a sharp image (change of focal length)
- Fun fact for CS people:
  - Human retinas bandwidth approx. 8.75 Mbit/s
  - Guinea pig’s bandwidth approx. 875 Kbit/s

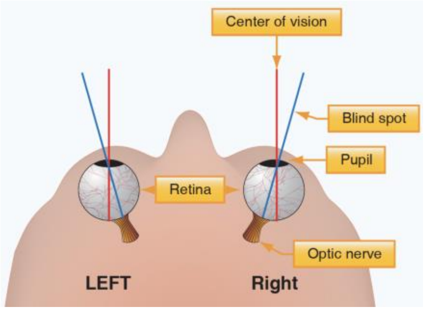




Kobayashi, H., Kohshima, S. 1997

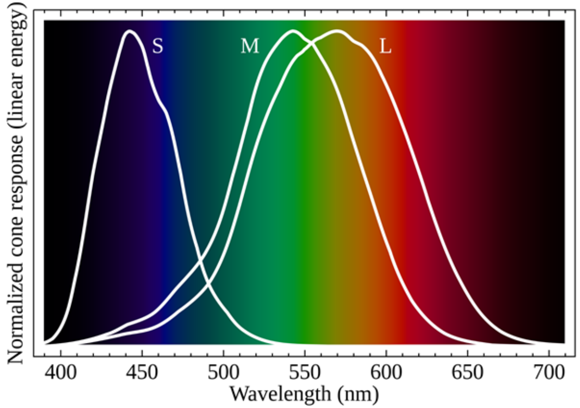


Clark et al. 2023



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

Light and Color in Human Vision

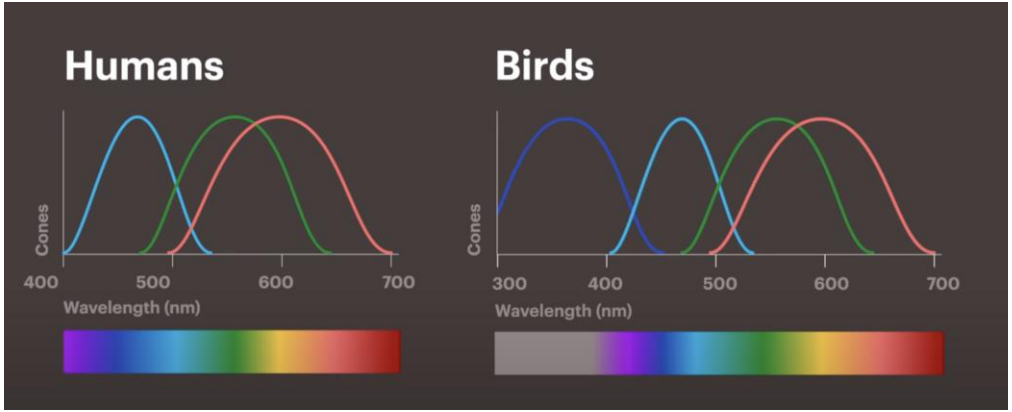


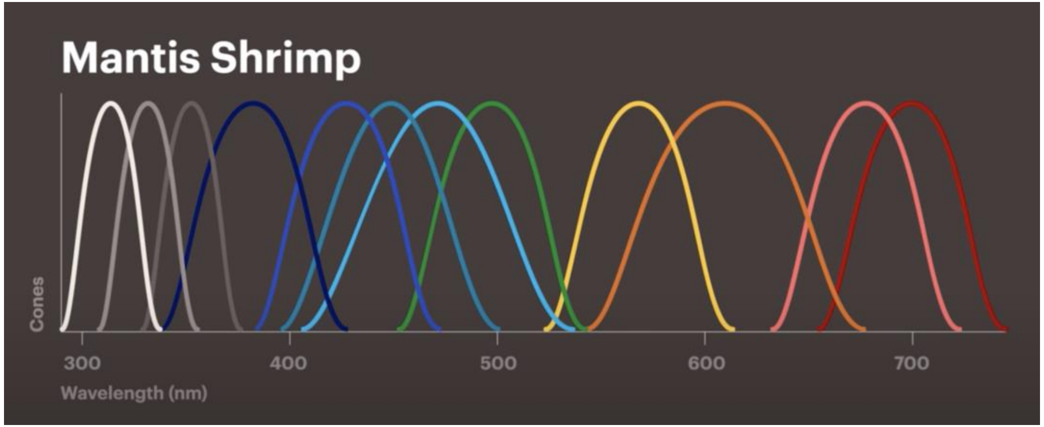
- Retina contains the photoreceptors
  - Rods and cones
  - Approx. 100 million rods, 5 million cones

- Rods are extremely sensitive to light and are important for night vision
- They play only a minor role in color vision (that's why you see less colors in darkness)
- Their density increases to the outer edges of the retina

- Cones are less light sensitive but enable color vision
- Three different types (S, M, L), categorized by their sensitivity to different wavelength (see image)

Light and Color in Human Vision

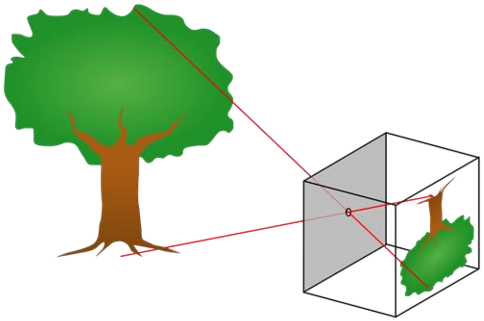




https://www.scientificamerican.com/article/camera-mimics-mantis-shrimps-astounding-vision/

How to mimic an eye? – The Pinhole Camera

- Very basic camera – Uses no lenses
- Usually a light proof box with a tiny hole at one side
- “Pinhole” because the hole is commonly punched with a pin
- The small hole blocks rays which do not reflect directly from the object to the camera
    - Otherwise we would not achieve a clear image of the subject
- The smaller the hole, the sharper the image (but also darker)
- Image appears upside down (like inside the human eye)

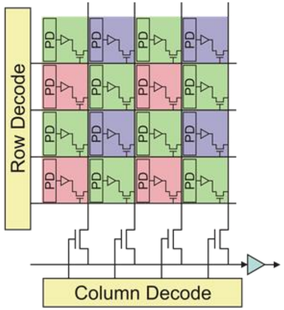


- One thing is missing to capture the actual image… The retina!
- The image can be captured analog by a film… or by a digital sensor!

From Human Vision to Computer Vision 👩‍💻

- Most practical method: Digital photography, recording a color image
- Instead of your rods and cones an image sensor uses photodiodes (PD)
- Photodiodes produce electricity when hit by photons (e.g. from visible light)
- To detect colors (specific wavelength) a color filter array is placed over the sensor
- Common filter is the Bayer Filter consisting of ½ green, ¼ red and ¼ blue filters

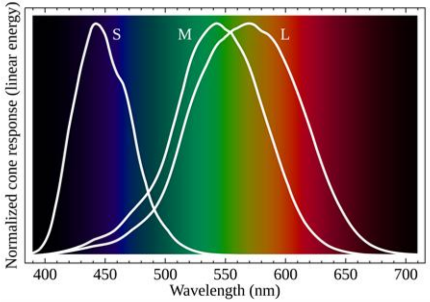
CMOS camera sensor

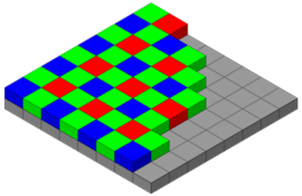


Photodiodes (PD) of a CMOS Image Sensor with Bayer Filter

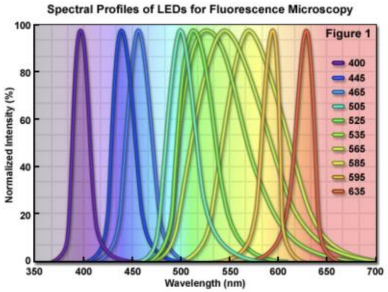
Enabling digital color vision – Bayer Filter

- The bayer filter is used so that a photodiode “reacts” only to a specific color
- Otherwise we would not know which color (wavelength) the photon had
- But why does the bayer filter have ½ green pixels?
- To mimic the human perception of color!
  - Our cones are most sensitive to green light (see graph)
- To get a red, green and blue (RGB) value for each square (pixel) demosaicing (interpolation) is applied

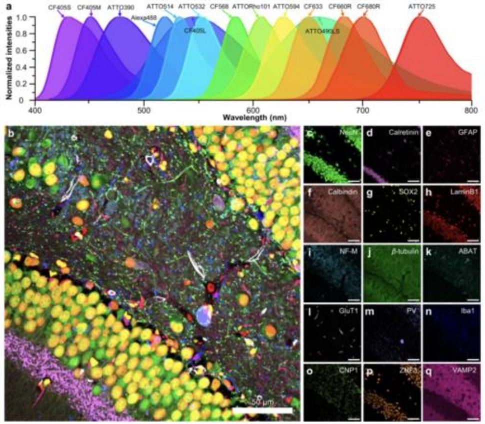




Bayer Filter Array

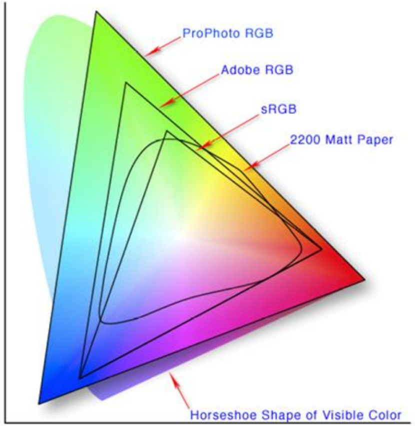


https://zeiss-campus.magnet.fsu.edu/articles/lightsources/leds.html



Seo, et al. 2022

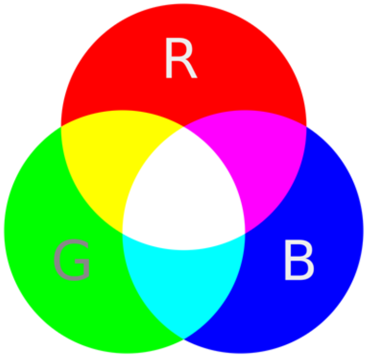
Color Spaces



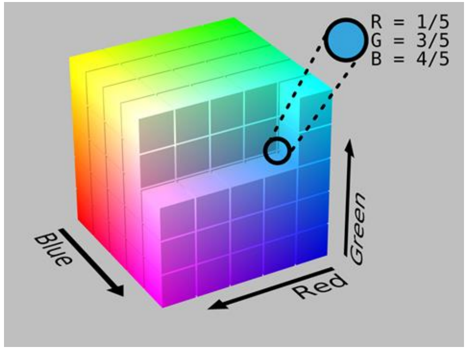
- We already talked about RGB, let us define this a bit more
- A color space is a combination of a color model and a mapping function
- A color model describes how colors can be represented as numbers

- Examples of color models:
  - RGB: Additive color mixing; individual values for the three channels red, green and blue
  - CMYK: Subtractive color mixing; individual values for the four channels cyan, magenta, yellow and black
  - HSV: encodes color with individual values for hue, saturation and value
- Common color spaces are: sRGB, Adobe RGB or DCI-P3; all using RGB color model but covering different colors (see image)

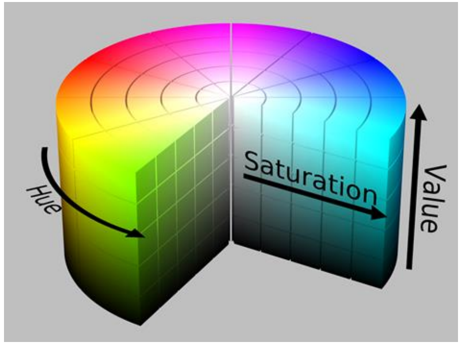
RGB – Color Model



- Probably the most common one
- Describes which light needs to be emitted for a specific color
  - Additive Model → Combine red, green and blue to get white color
  - No color (0, 0, 0) → Black image
- RGB color spaces can be mapped to a cube
- Commonly used for images in computer vision
- Typically colors of a pixel are identified by a triplet (R, G, B)
  - Floats between 0 and 1 → (0.96, 0.62, 0.11)
  - Integer between 0 and 255 (8 bit) → (245, 157, 29)
  - Hexcode → #F59D1D
  - If all three values are the same, we get a grayscale color



HSV – Color Model



- Cylindrical coordinate representation in a RGB color model
- Defines a color with three values:
  - Hue: angular axis, encodes the base color
    - 0° red, 120° green, 240° blue
  - Saturation: defines the intensity of the color
  - Value (also called brightness): encodes the perceived brightness
- Often used in CV
  - Many CV algorithms for color images are extensions of the grayscale version
  - As a result, sometimes each color component is passed individually
  - With RGB all channels are correlated (all encode the amount of light hitting the object)
  - With HSV they are easier to separate

Looking at bits

- Until now we talked about the hardware to capture an image and how color is represented
- But how does an image look like “under the hood”?
- A simple and common way is to store an image as a matrix of pixel values
- The image coordinate systems origin (0, 0) lies in the top left corner
- For grayscale images a pixel is represented by one value, resulting in a 2D array
- For RGB images a pixel is represented by three values, resulting in a 3D array
- In practice images often come with a lot of (meta-)information
    - In photography: GPS coordinates, camera model, aperture f-number, …
    - In bioimages (e.g. microscopy): magnification, resolution, data to identify the sample/experiment, …
- Images can also be 2-byte, 12-bit and contain many channels!