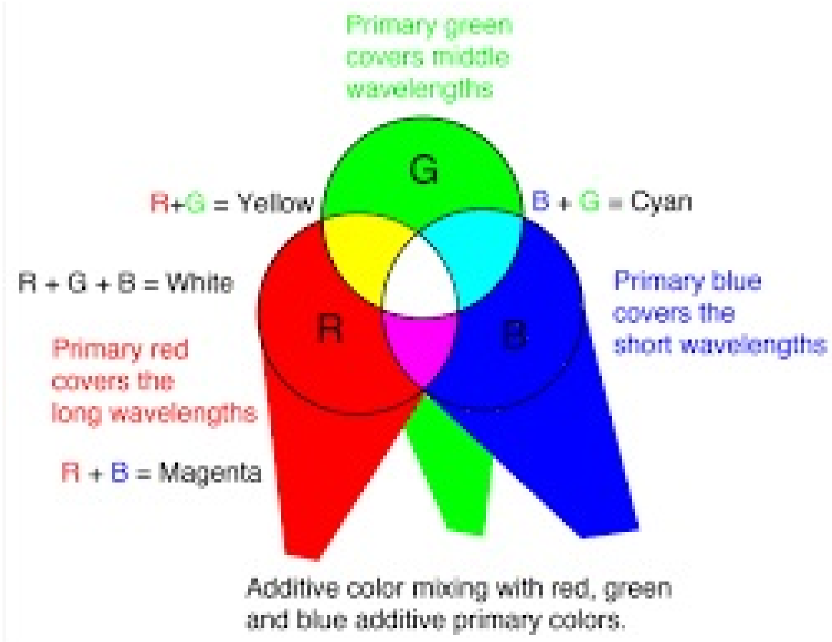
Colour Space Representation

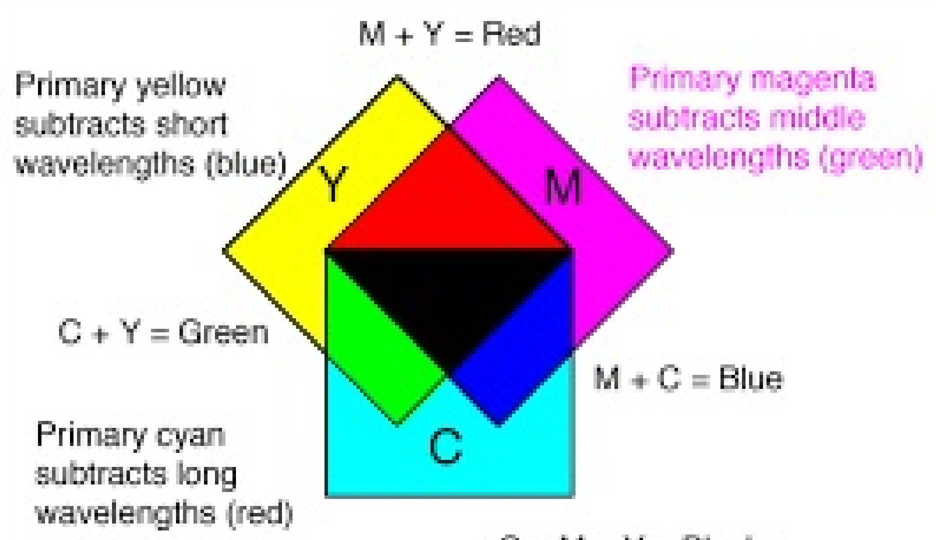
* popular color spaces are RGB, CMYK
* Additive colors are R, G, B
* Subtractive colors are C, M, Y
* Other color spaces are XYZ, YUV, Lab, YCbCr, HSV

* Standards are established by Comission Internationale d'Eclårage (CIE)
* Understanding of color spaces important in color industry.

so additive colours are RGB, R, G and B; subtractive colours are C, M and Y particular application where CMYK is used in practice is in printers. So, it happens that it is a lot easier to control colours using CMYK in printers, you can read more about these on these links provided below.

The CIE which establishes standards for colour spaces because this is an important, this is actually important for the printing and scanning industry

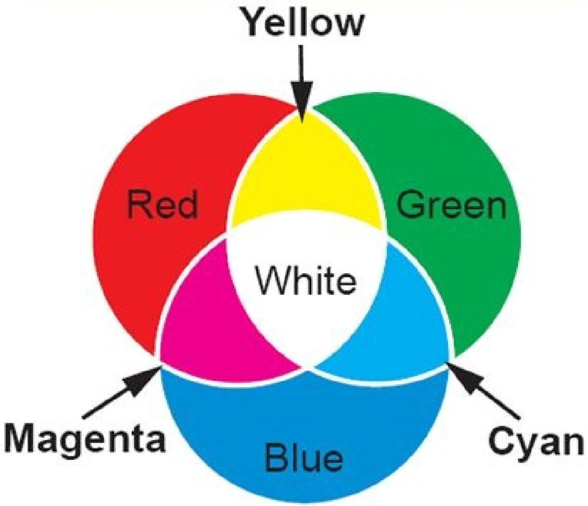




Color subtraction with idealized primary filters

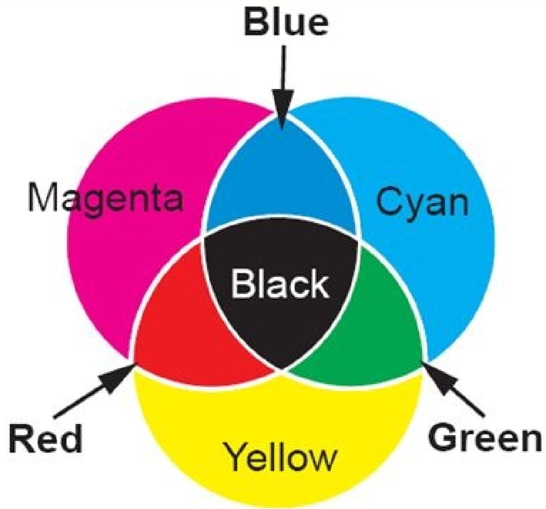
Additive and Subtractive Primary Colors

The additive primary colors



White = red + green + blue
Yellow = red + green
Magenta = red + blue
Cyan = blue + green

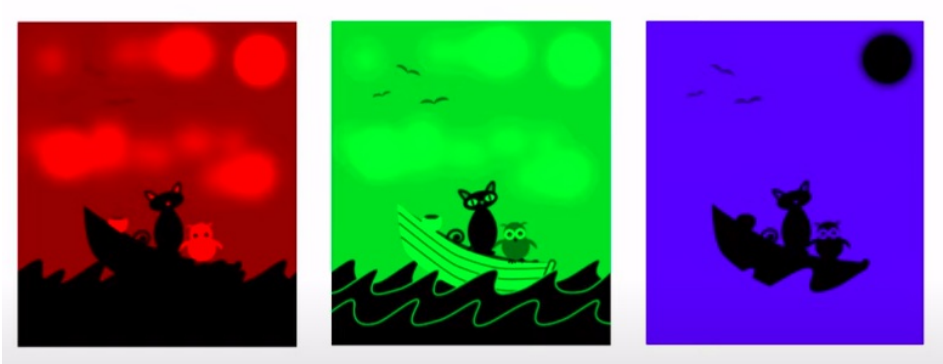
The subtractive primary colors

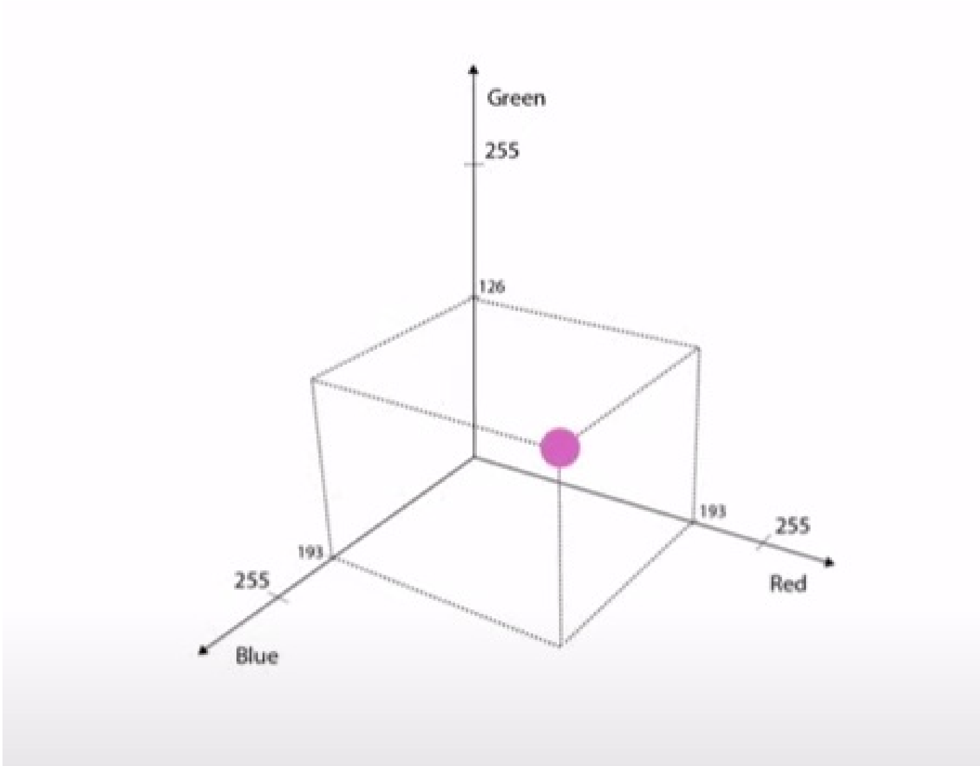


Black = magenta + yellow + cyan
Red = magenta + yellow
Green = cyan + yellow
Blue = magenta + cyan

The CMYK color model is a subtractive color model, based on the CMY color model, used in color printing,

Its **Red, Green, and Blue primaries** represent the actual wavelengths of light used in the original Wright and Guild experiments that led to the creation of the XYZ color space.





$2^8 = 256$

$\downarrow$

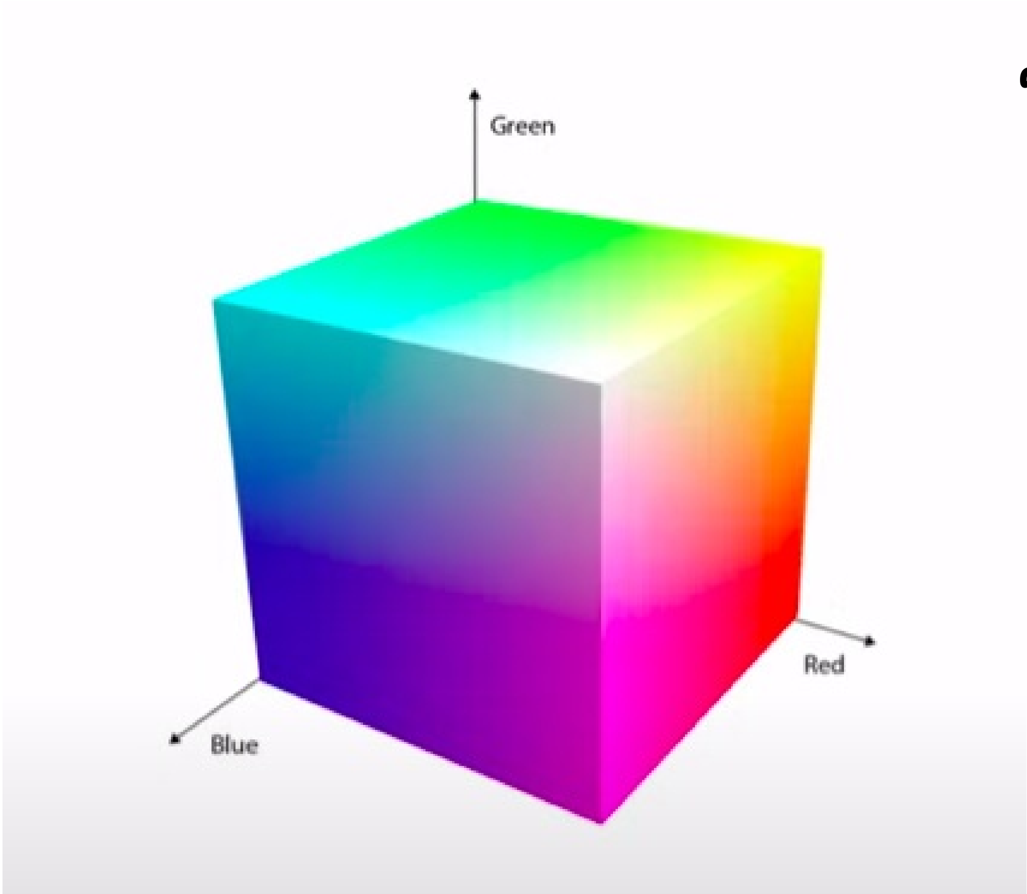
$0 \ldots 255$

8-bits R

8-bits G

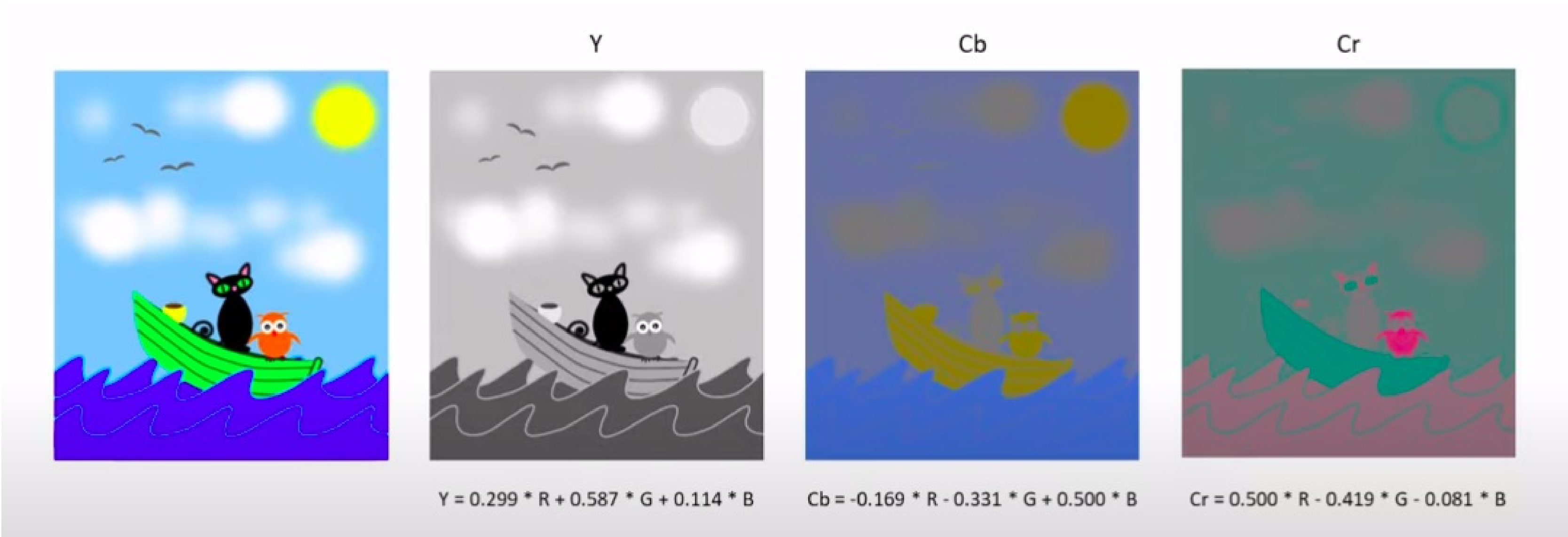
8-bits B

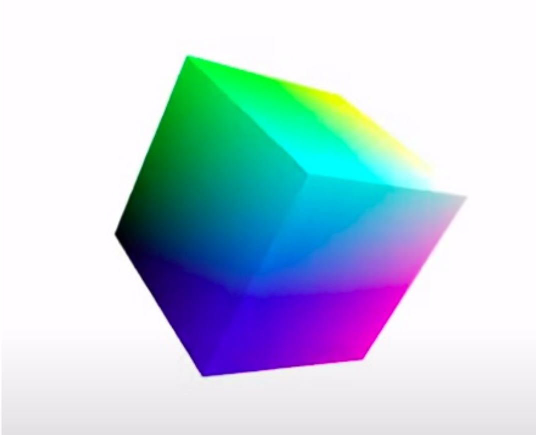
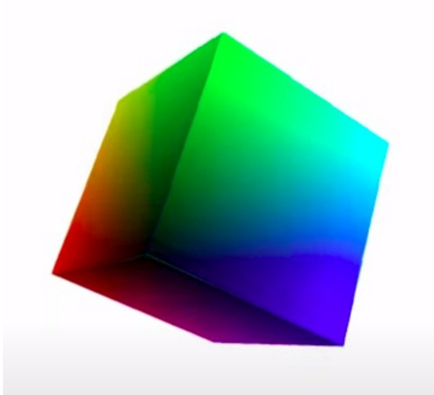
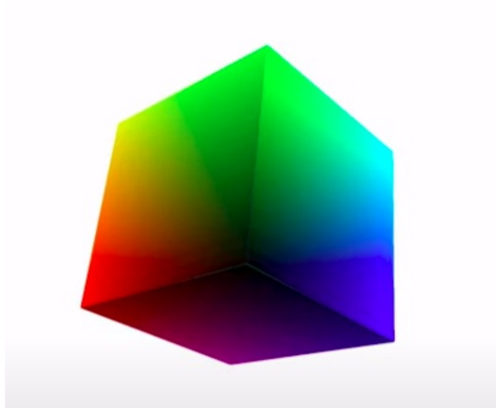
= 24-bits

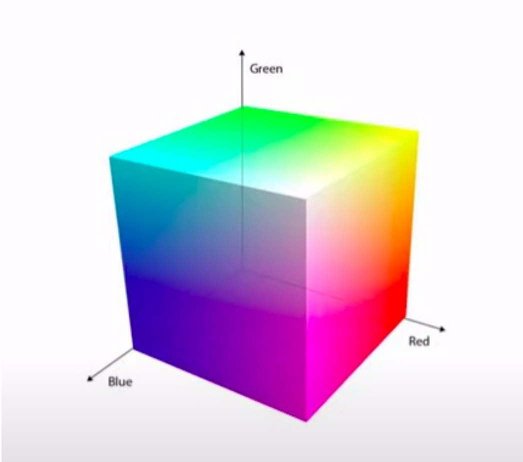
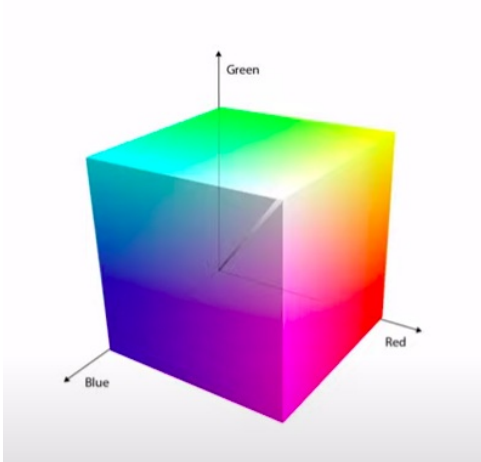
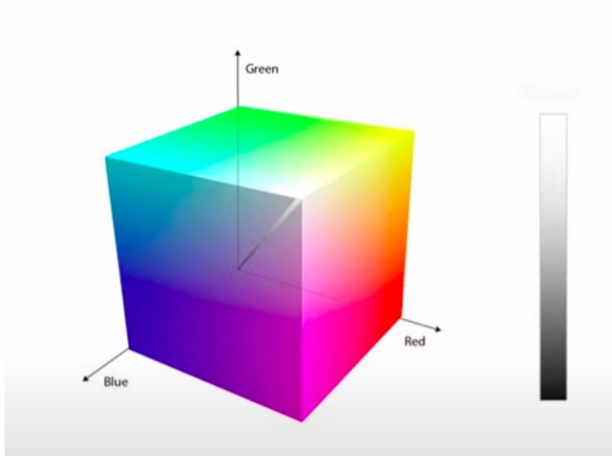


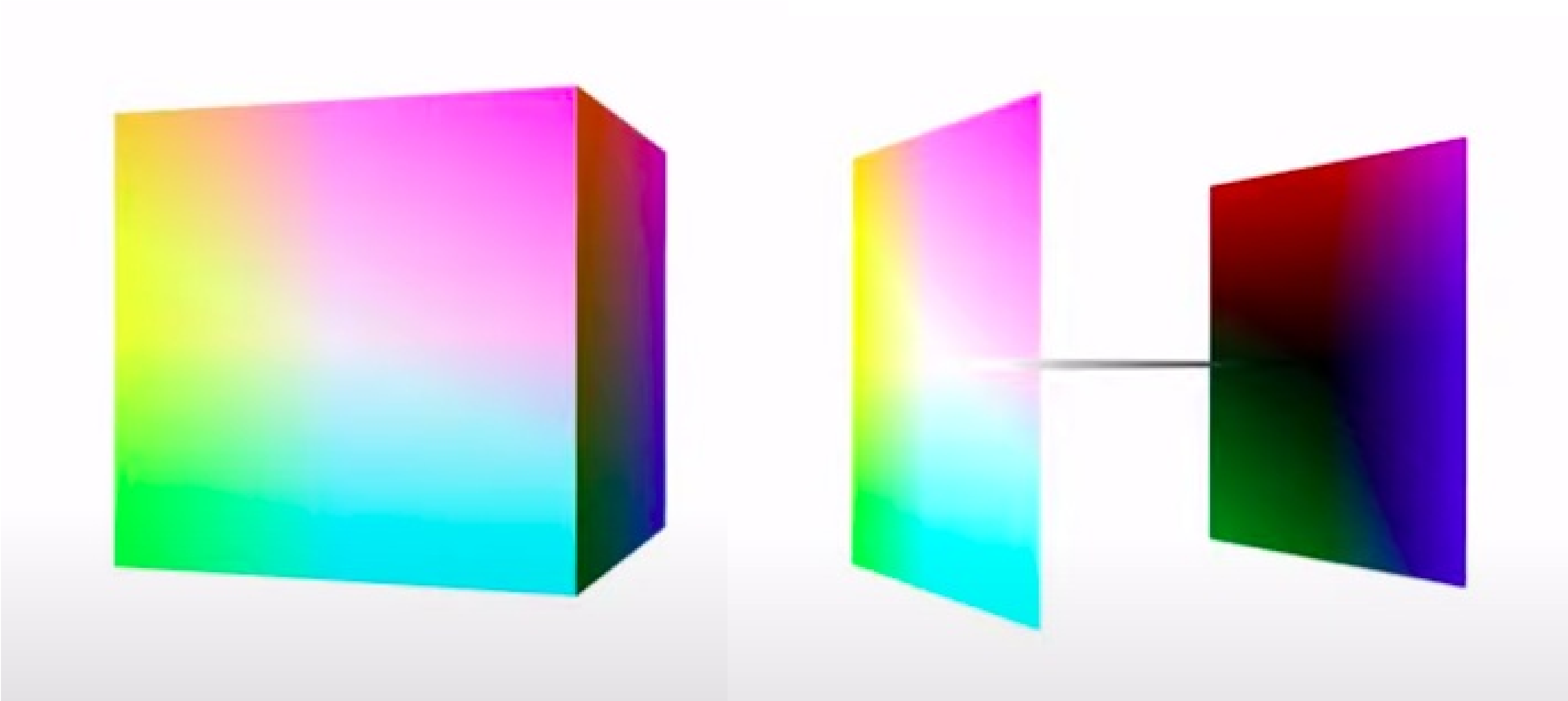
24-bits

16.7 million


Y
Cb
Cr
Y = 0.299 * R + 0.587 * G + 0.114 * B
Cb = -0.169 * R - 0.331 * G + 0.500 * B
Cr = 0.500 * R - 0.419 * G - 0.081 * B


Green
Red
Blue

Green
Red
Blue

Green
Red
Blue



YCbCr → YUV

luminance

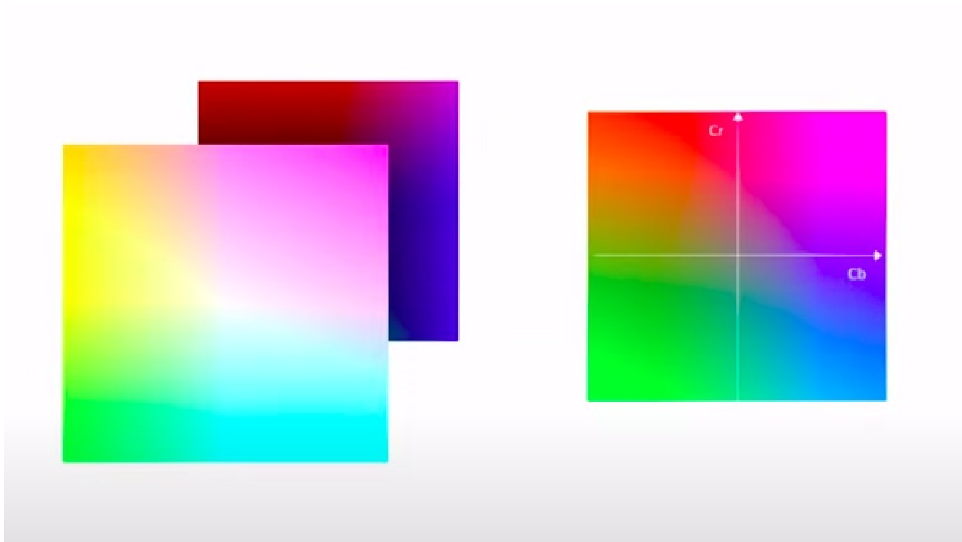


Image Compression

- Last stage in camera's processing pipeline
- Convert signal into YCbCr, compress luminance with higher fidelity than chrominance.

Because of the way humans or the human visual system perceives light, luminance is a bit more important than chrominance, so you ensure that luminance is actually compressed with a higher fidelity which means your reconstruction is better for luminance than for chrominance, so that is one reason why YCbCr is used as a popular colour space before
storage, once again if you do not understand YCbCr, go back to the previous slide look at all
of these links to understand YCbCr is one of the colour space representations that are
available in practice.

- most common compression technique is Discrete Cosine Transform (DCT) that is used in MPEG and JPEG

- videos also use block level motion compensation

Images are divided into frames and set of frames into block and then you store certain frames based on concepts from motion compensation, this is typically used in the MPEG standard which uses, which divides all frames into what are known as i frames, p frames and b frames and then uses strategies to decide how each frame should be coded, that is how videos are compressed.

- Compression quality measured using

Peak Signal-to-Noise Ratio (PNSR)

$$PSNR = 10 \log_{10} \frac{I_{max}^2}{MSE}$$

where MSE (mean square Error)

$$= \frac{1}{n} \sum_{x} \left[I(x) - \hat{I}(x) \right]^2$$

where $\hat{I}$ is compressed version of I

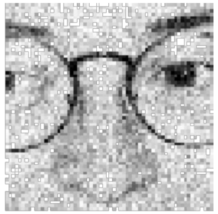
I_{max} = Maximum Intensity

MSE basically calculates the mean square error pixelwise between these two images.

Thus, PSNR measures the quality of image compression

PSNR = 40 dB

PSNR = 30 dB



PSNR = 20 dB

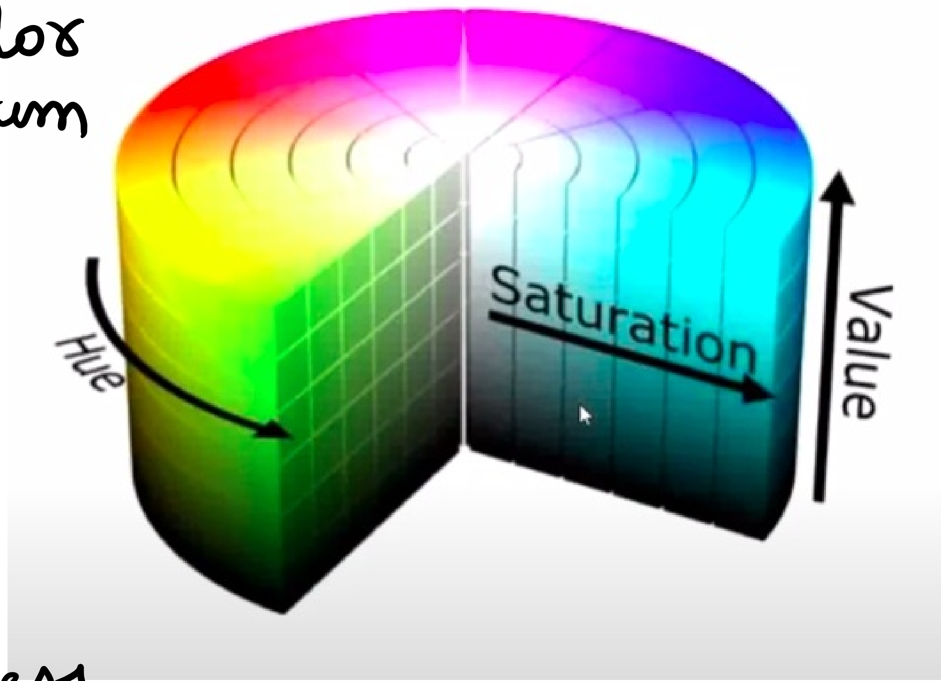
PSNR = 10 dB

PSNR = 0 dB

HSV Color Space Representation

- 3 Channels: Hue, Saturation, and Value

- Hue decides the color in the entire spectrum

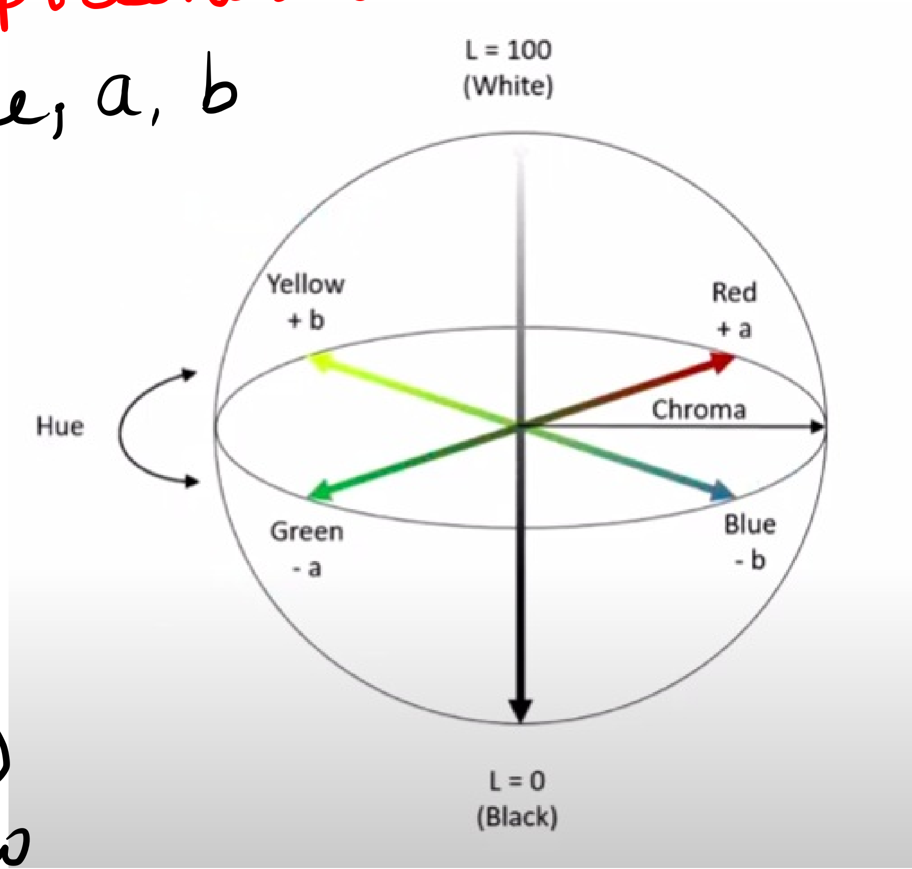


- Saturation decides purity and strength of the color (lower means faded)

- Value decides brightness of the pixel (0 is black and 255 is the brightest color)

Lab Color space Representation

- 3 channels: Luminance, a, b



- Luminance is the brightness/intensity of the pixel
- a: how green or red is the pixel (green - red spectrum)
- b: how blue or yellow is the pixel (blue - yellow sepectrum)

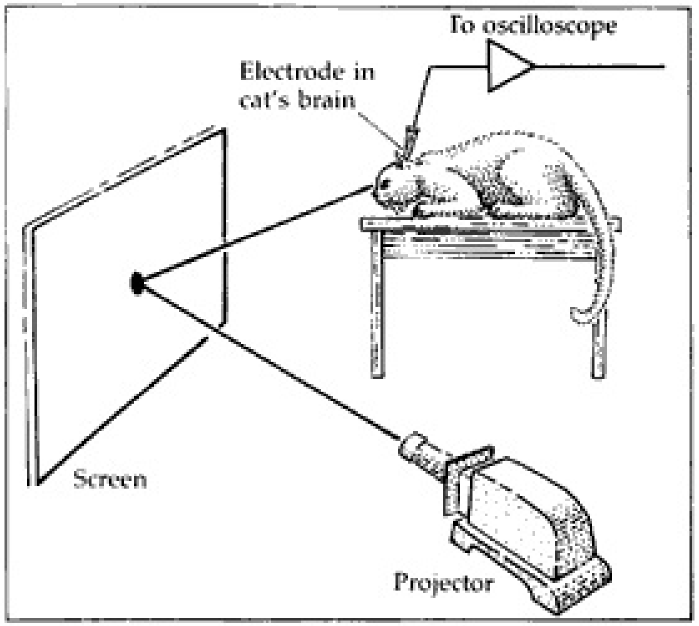
History of Computer vision

50s
60s inventory
70s

80s low level understanding

90s high level understanding
2000s

{developments in last decade



STIMULATION OF RETINA with patterns of light. The eyes of an anesthetized, light-adapted cat (or monkey) focus on a screen onto which various patterns of light are projected. Alternatively, a TV screen is used, with patterns generated by a computer. An electrode records the responses from a single cell in the visual pathway. Light (or shadow) falling onto a restricted area of the screen may accelerate (excite) or slow (inhibit) the signals given by a neuron. By determining the areas on the screen from which a neuron's firing is influenced, one can delineate the receptive field of the cell. The positions of cells in the brain and the tracks of electrode penetrations can be reconstructed histologically after the experiment.

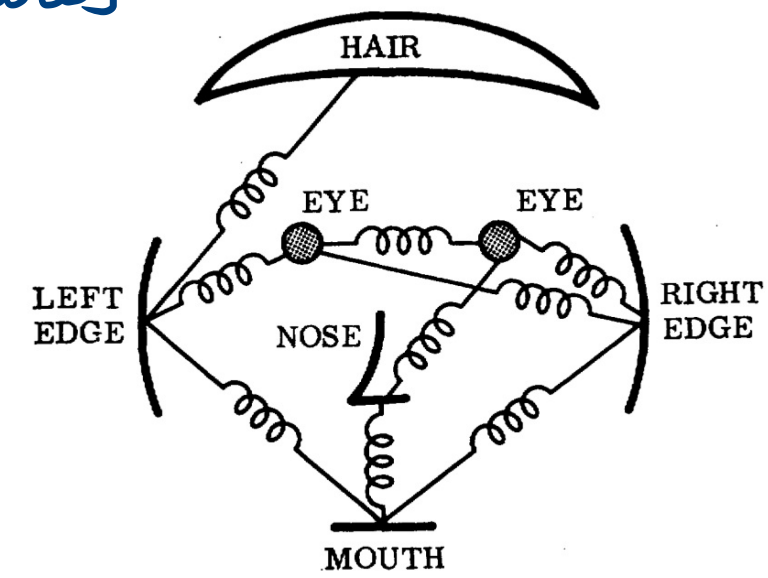
In 50s, David Hubel and Torsten Wiesel published their work called "Receptive fields of single neuron in the cat's striate cortex"

This experiment was conducted to understand how mammalian vision cortex functions

In 1959, Russell Kirsch and his colleagues were first time represented an image as a set of 1's and 0's

5X5 cm2 (162 X162) pixels

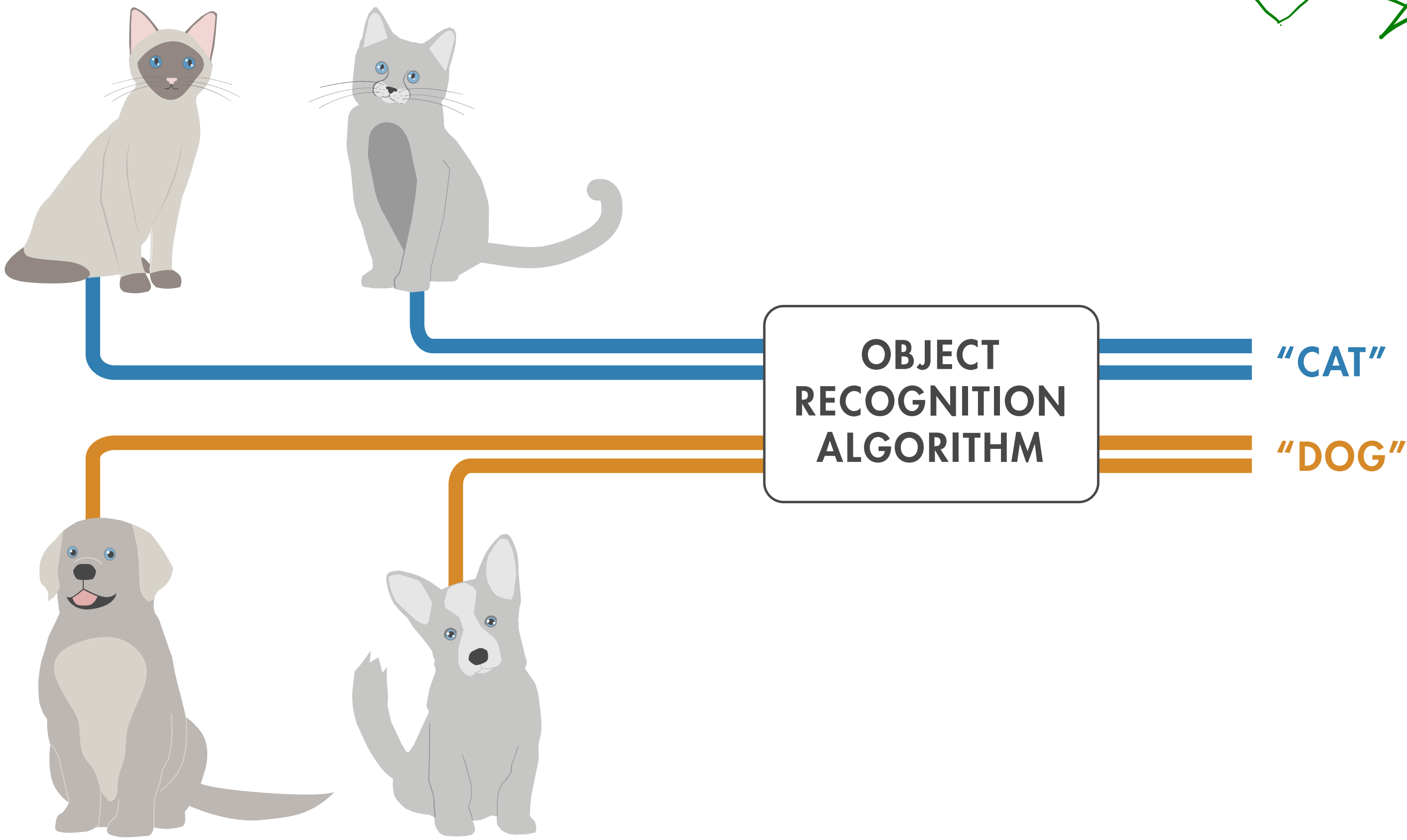
In 1973, pictorial structures model by Fischer and Elschlager which was again reinvented in early 2000s



- They wanted visual object's description that somebody should be able to find out the object in a photograph.

- So outcome defines an object as a combination of individual components and the connections between those components.

1971 - 1978 (winter of AI)



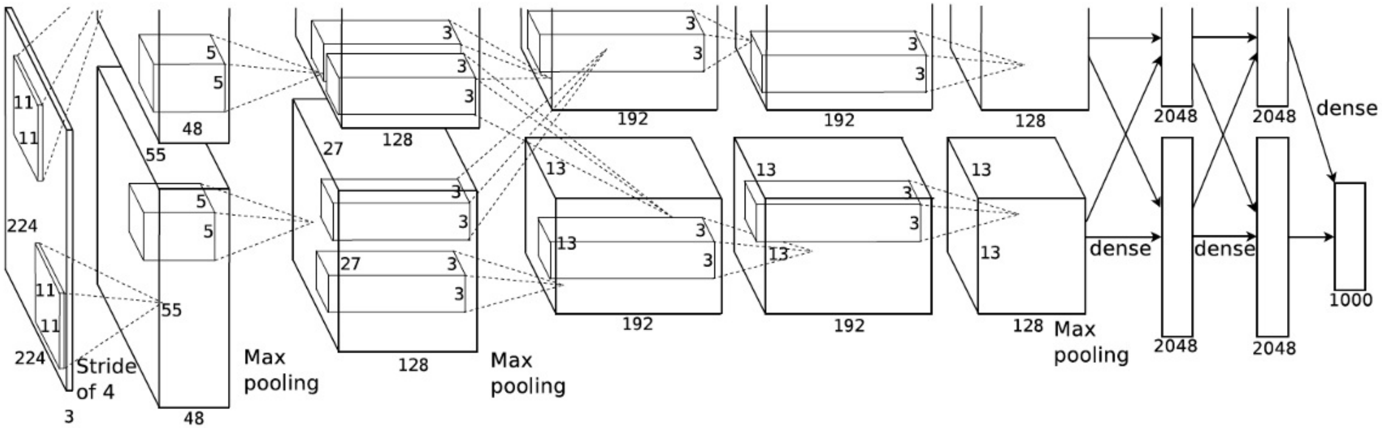
Efforts are focused that time on object recognition using shape understanding, envision objects as summation of parts

Deep Learning Era (2010 - till date)

In 2010, ImageNet arrives

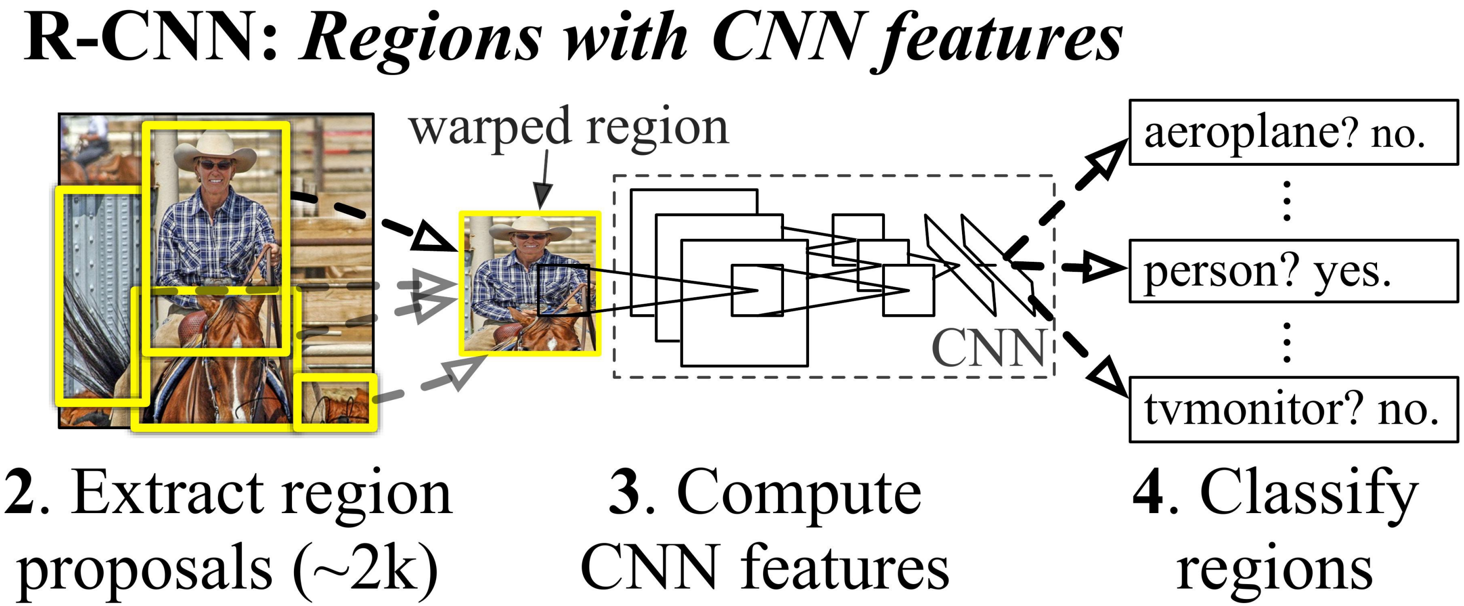
Shallow model

Extract some features out of the images and then used Machine Learning models such as support vector machine to do object recognition.



In 2012, Alexnet was the first Convolution neural network

In 2013, regions CNNs (R-CNN) were developed for object detection

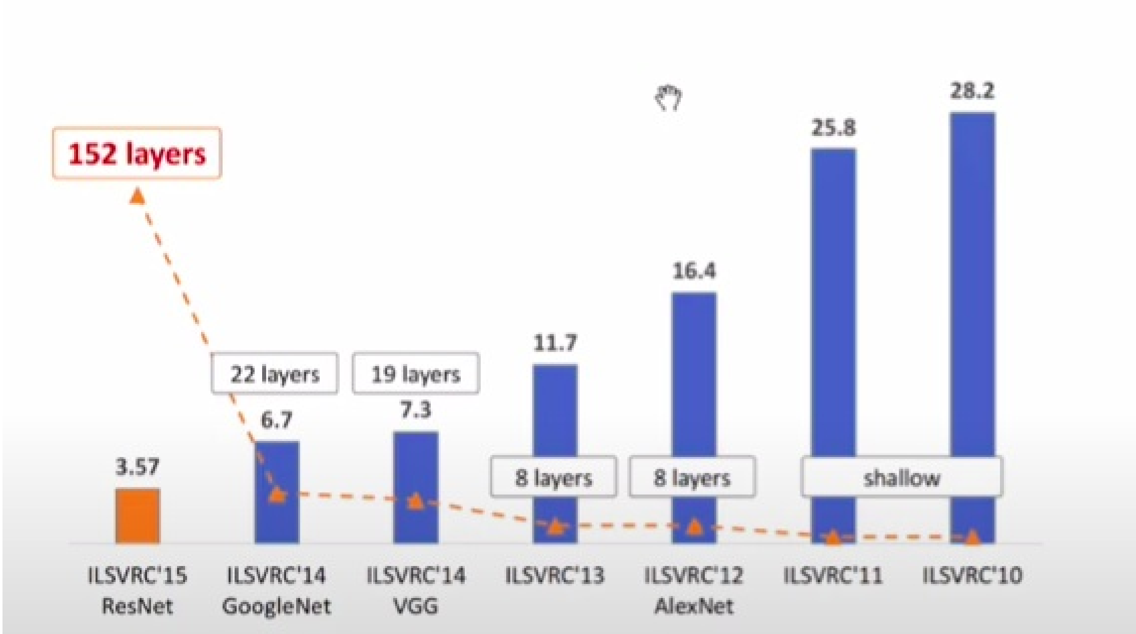


In 2014, InceptionNet and VGG models arrive, Human pose estimation CNNs, Deep generative models GANs, VAEs

GANs → Generative Adversarial Networks

VAEs → Variational Auto Encoders

In 2015, Residual networks or ResNets arrived and CNNs matched human performance on ImageNet



In 2016, moving beyond region based CNNs for object detection, single stage methods such as You Only Look Once (YOLO) and Single Shot Detector (SSD) were developed

The cityscapes dataset arrived, the visual genome dataset arrived and 2017 was the start of a higher level abstraction in Understanding images which is scene graph generation.

And in 2018 and 2019, higher level of abstraction such as the visual common sense reasoning dataset

Where we not only give answer to a question on an image but can also give a rational to that answer and task such as Panoptic Segmentation has been developed.

So we can see, in this journey, the focus is going from low level image Understanding to higher abstractions of the world we see around us from images.

History of applications

1970s: Optical Character Recognition (OCR)

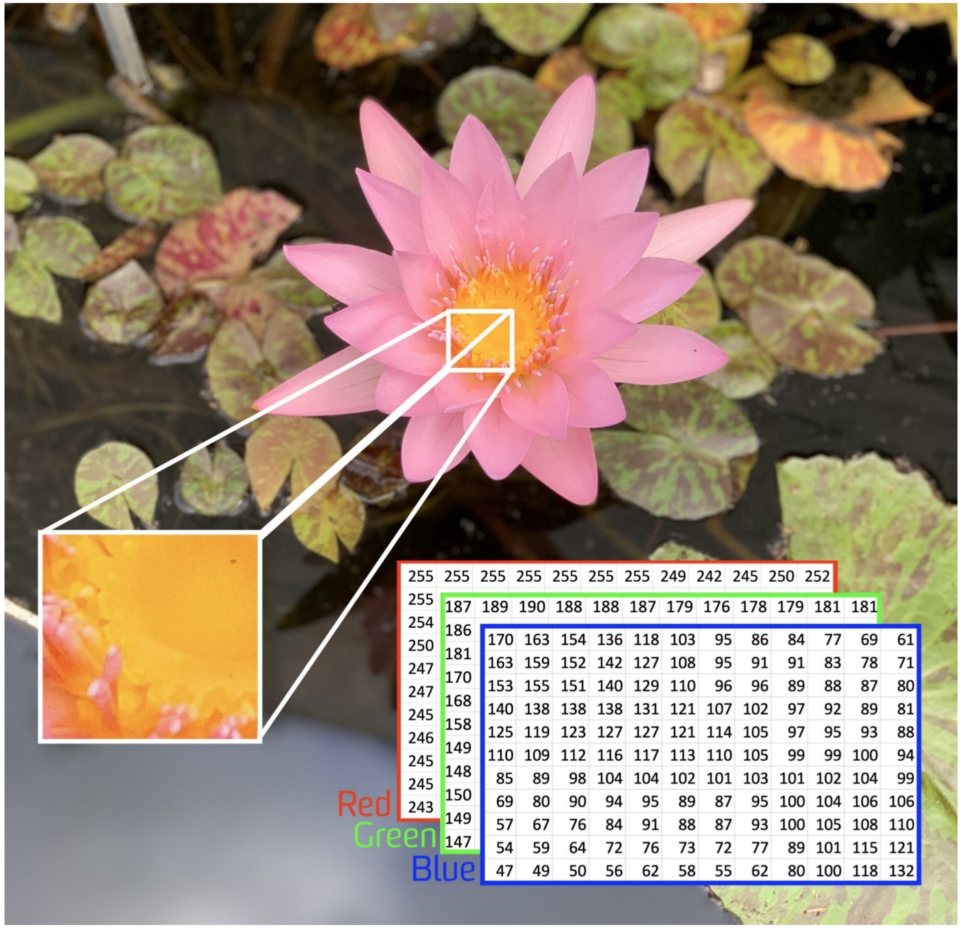
1980s : Machine vision, Smart Cameras

1990s : machine vision in manufacturing environment, biometrics, medical imaging recording devices, video surveillance

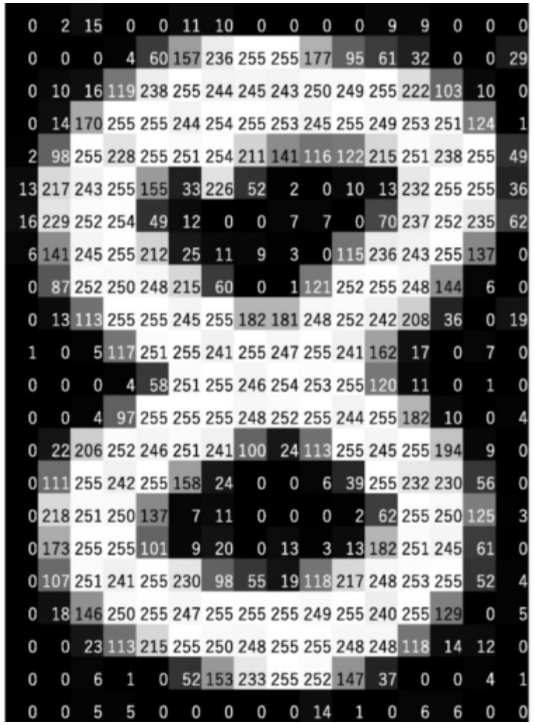
2000s: more biometrics, better medical imaging, object/face detection, autonomous navigation, Google Goggles, vision on social media

2010s : Everything around us

Color Image as a matrix



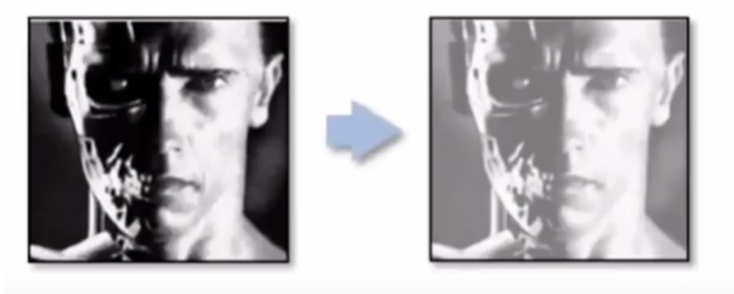
Grayscale image as a matrix



0	2	15	0	0	11	10	0	0	0	0	9	9	0	0	0
0	0	0	4	60	157	236	255	255	177	95	61	32	0	0	29
0	10	16	119	238	255	244	245	243	250	249	255	222	103	10	0
0	14	170	255	255	244	254	255	253	245	255	249	253	251	124	1
2	98	255	228	255	251	254	211	141	116	122	215	251	238	255	49
13	217	243	255	155	33	226	52	2	0	10	13	232	255	255	36
16	229	252	254	49	12	0	0	7	7	0	70	237	252	235	62
6	141	245	255	212	25	11	9	3	0	115	236	243	255	137	0
0	87	252	250	248	215	60	0	1	121	252	255	248	144	6	0
0	13	113	255	255	245	255	182	181	248	252	242	208	36	0	19
1	0	5	117	251	255	241	255	247	255	241	162	17	0	7	0
0	0	0	4	58	251	255	246	254	253	255	120	11	0	1	0
0	0	4	97	255	255	255	248	252	255	244	255	182	10	0	4
0	22	206	252	246	251	241	100	24	113	255	245	255	194	9	0
0	111	255	242	255	158	24	0	0	6	39	255	232	230	56	0
0	218	251	250	137	7	11	0	0	0	2	62	255	250	125	3
0	173	255	255	101	9	20	0	13	3	13	182	251	245	61	0
0	107	251	241	255	230	98	55	19	118	217	248	253	255	52	4
0	18	146	250	255	247	255	255	255	249	255	240	255	129	0	5
0	0	23	113	215	255	250	248	255	255	248	248	118	14	12	0
0	0	6	1	0	52	153	233	255	252	147	37	0	0	4	1
0	0	5	5	0	0	0	0	0	14	1	0	6	6	0	0

Image Transformation (Image as function)

- we can think of a grascale image as a function $f: R^2 \rightarrow R$ giving the intensity at position (x,y)



- A digital image is a discrete (sampled, quantized) version of this function