Image Processing

INT 302

Siyue Yu
siyue.yu02@xjtlu.edu.cn

Contact information

- **Qiufeng Wang**
  - e-mail : qiufeng.wang@xjtlu.edu.cn
  - tel. : +86 512 8188 4765
  - Department of Intelligent Science (INT)
  - Office: **EE510A**
- **Siyue Yu**
  - e-mail : siyue.yu02@xjtlu.edu.cn
  - tel. : +86 512 8816 1507
  - Department of Intelligent Science (INT)
  - Office: **SD553**
- **Teaching assistants (TA)**
  - Jianan Ye, Yijie Hu, Zhaorui Tan
  - e-mail : {jianan.ye20, yijie.hu20, zhaorui.tan21}@student.xjtlu.edu.cn
  - Room: EE 515

Office Hours

- Tuesday morning: 9:30 ~11:30
- it is good practice to send an email to arrange for an appointment
- Or call the tel. first to check whether I’m at the office or not

Practical Information : teaching hours

- The course will be delivered via lectures, and Laboratory assignments.

- Lectures, 2hx12
  - Tuesday afternoon: 14:00 – 16:00
  - Classroom: SA136
- Labs, 4hx3:
  - Thursday Morning: 9:00 - 12:50
  - Labs will be organized by TA, and the exact date will be communicated in due course (e.g., week5, 9, 11)
  - You can go to the lab room personally to continue to finish the lab in other time slots on the timetable.

Practical Information: Course structure

- Introduction, Human Vision System, image acquisition, Spatial resolution of digital images, Effect of quantization, matrix representation of images, Image storage, image color model. Basic relationship between pixels, Distance measures.
- Image Enhancement; histogram processing; image smoothing and sharpening, filtering in spatial and frequency domains; High pass & low pass filtering.
- Image transforms; spatial frequency concepts, Fourier Transform, Discrete Cosine Transform.
- Image Compression; Introduction to source coding, Types of redundancy Concept of lossless-lossy compression.
- Morphological image processing.
- Segmentation, thresholding, texture segmentation.

Note: Some topics may not be in order to maintain coherency and running time requirements.

Practical Information : Lab

- The lab experiments will be carried out using MATLAB.
- For each lab, a brief report shall be prepared, which will be evaluated and will concur to determine the final grade.
- Most projects include analytical and programming parts.
- All projects must be done individually.

Grading

- FIVE Projects – 100 % (15%, 15%, 70%)
  - There will be an in-class quiz in the lab 3.
- Most projects include analytic and programming parts.
- All projects must be done individually.
- Programming Environment – Matlab.
- For now no re-sit possibility

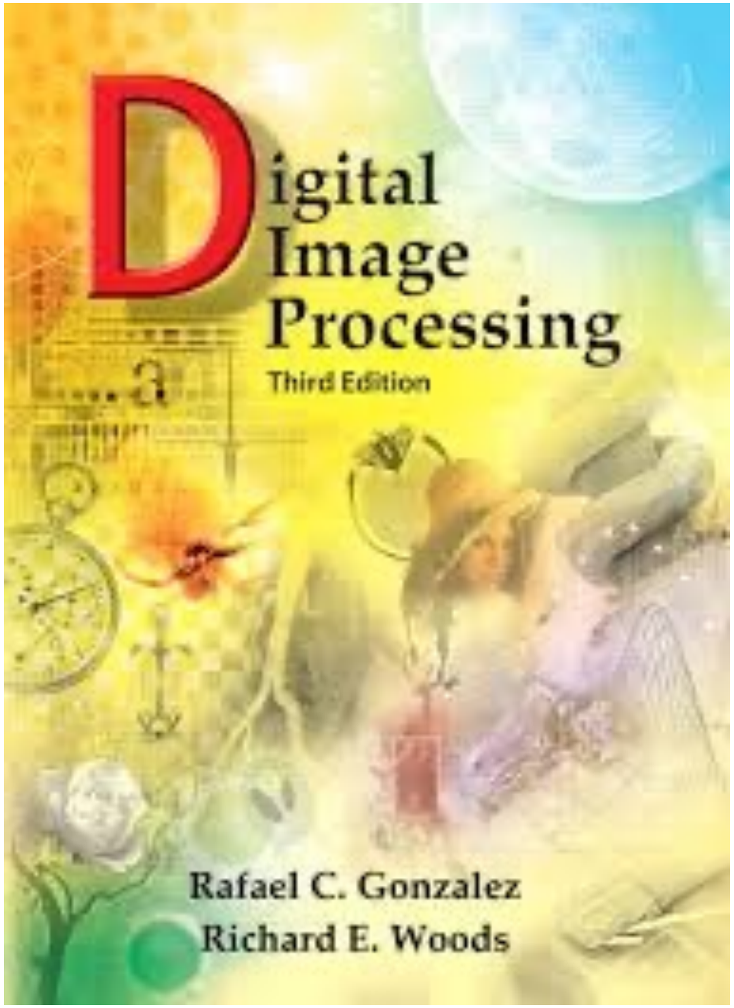
Acknowledgements

The materials to be delivered in this course are prepared according to the following sources :

- Materials taken from Wikipedia and the Internet.
- Digital Image Processing (3rd Edition) by Gonzalez And R.E. Woods

Reference book

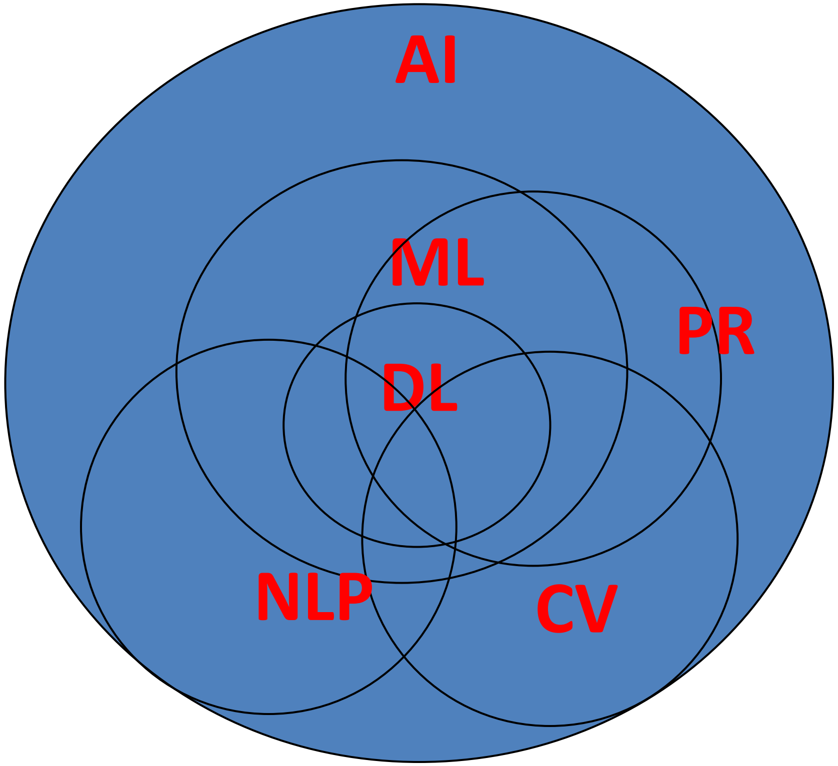
- **Reference**
  - Digital Image Processing (3rd Edition)
  - R.C., Gonzalez And R.E. Woods
  - ISBN: 9780132345637



Introduction
to
digital image processing

Related Disciplines

- Artificial Intelligence (AI)
- Pattern Recognition (PR)
- Machine Learning (ML)
- Deep Learning (DL)
- Computer Vision (CV)
- Speech Recognition
- Natural Language Processing (NLP)
- Others



Artificial Intelligence

Artificial intelligence

From Wikipedia, the free encyclopedia
(Redirected from AI)

"AI" redirects here. For other uses, see AI (disambiguation) and Artificial intelligence (disambiguation).

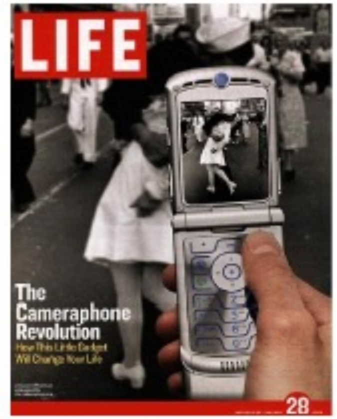
In computer science, **artificial intelligence** (**AI**), sometimes called **machine intelligence**, is intelligence demonstrated by machines, in contrast to the **natural intelligence** displayed by humans. Colloquially, the term "artificial intelligence" is often used to describe machines (or computers) that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving".[1]

As machines become increasingly capable, tasks considered to require "intelligence" are often removed from the definition of AI, a phenomenon known as the AI effect.[2] A quip in Tesler's Theorem says "AI is whatever hasn't been done yet."[3] For instance, optical character recognition is frequently excluded from things considered to be AI, having become a routine technology.[4] Modern machine capabilities generally classified as AI include successfully understanding human speech,[5] competing at the highest level in strategic game systems (such as chess and Go),[6] autonomously operating cars, intelligent routing in content delivery networks, and military simulations.

Some questions

- Why image/video processing is important?
- How the human vision system works ?
- What is light ?
- What is image ?

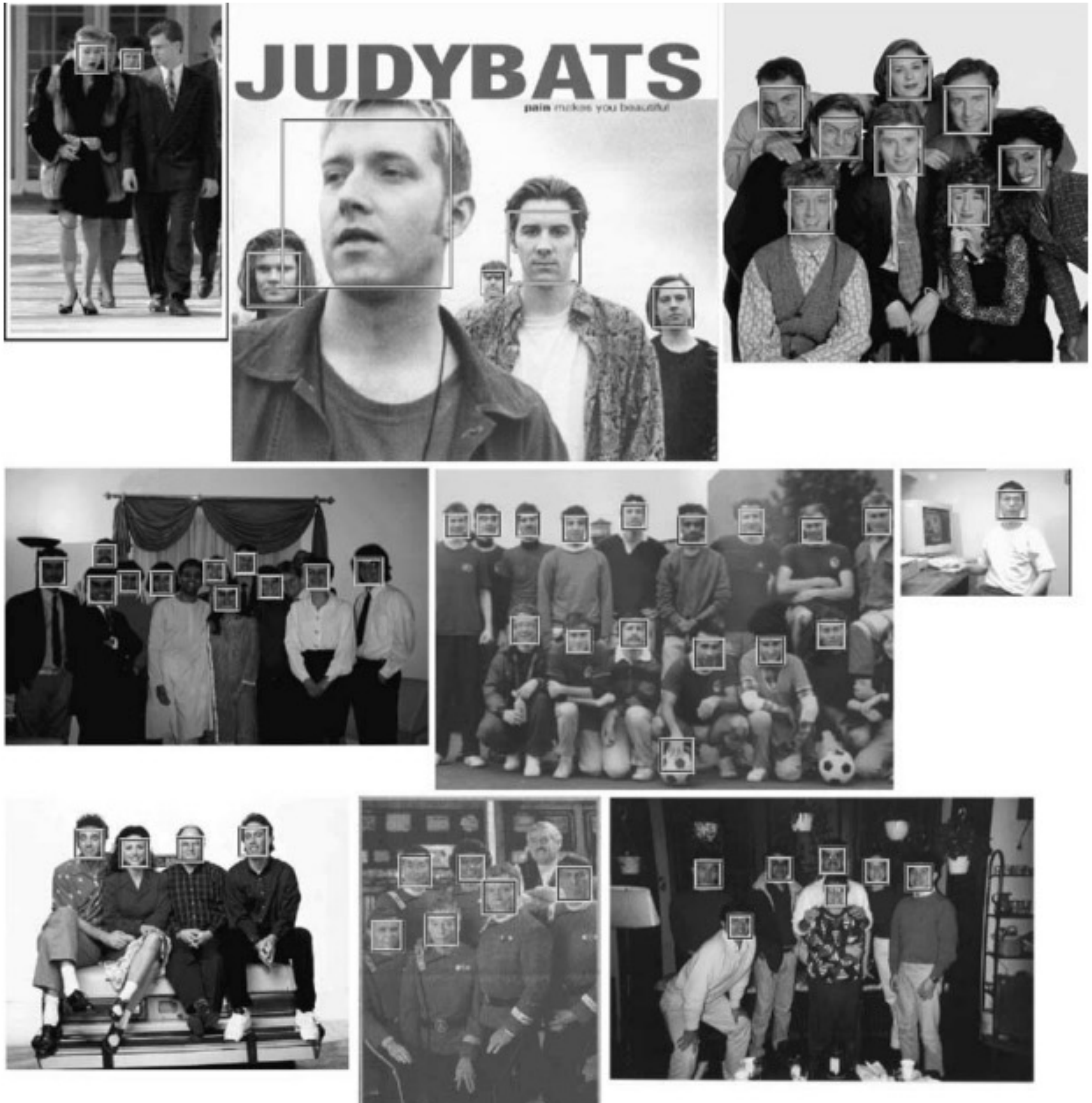
Why image processing is important?



Why image processing is important?

One picture **is worth more than** ***ten thousand words*****.**

Why image processing is important?



Why image/video processing is important?

- Automatic car driving:

Why image/video processing is important?

- Our eye needs help from computers

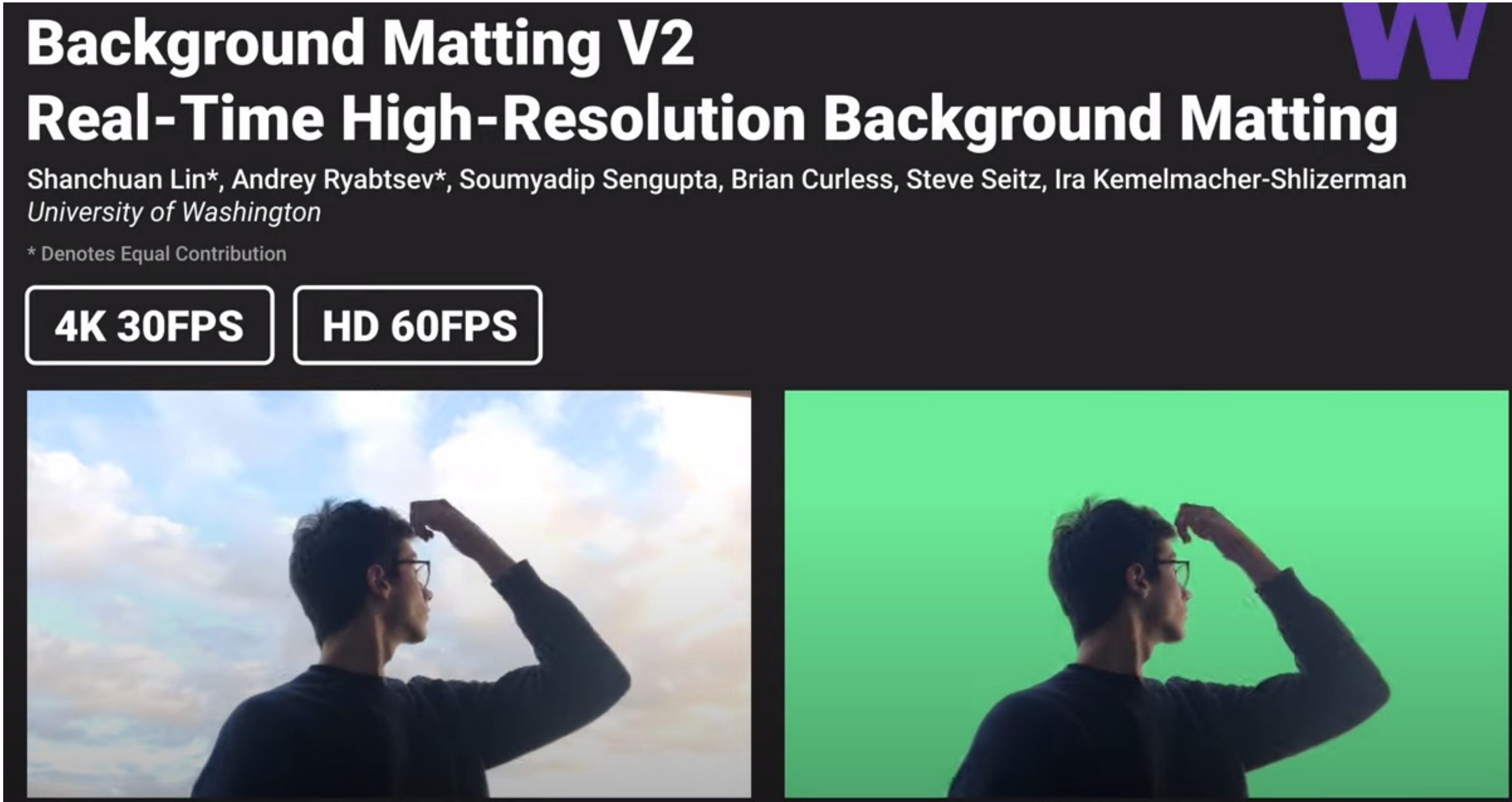
Why image processing is important?

“A surfer riding on a wave”

[Try one demo of **Image Caption**:
https://milhidaka.github.io/chainer-image-caption/](https://milhidaka.github.io/chainer-image-caption/)

Why image/video processing is important?

- https://youtu.be/oMfPTeYDF9g



What is image processing ?

- Digital image processing focuses on two major tasks
  - Improvement of visual information for human interpretation (sometime it is subjective)
  - Processing of image data for storage, transmission and representation for autonomous machine perception (it is objective)

➔Goal in general : produce from an input image to a new image (or attributes), to be processed further or to be interpreted by humans or machines.

Some examples of image processing

- Noise removal: Given an image contaminated by noise, reduce the noise level while distorting the image as little as possible
  - Approach 1: Apply a lowpass filter (LPF). Reduces noise, but blurs important image structure.
  - Approach 2: Apply a nonlinear filter to smooth in flat regions of image, retain structure.

- Edge enhancement: Given an image, produce a new image in which edges are emphasized.
  - Approach 1: Apply a highpass filter. Emphasizes edges, but increases noise.
  - Approach 2: Apply an edge detector.

- Image compression: To store a color photograph at full resolution of about 2500 by 3750 pixels at 32 bits/pixel, it is about 300Mbits. To make the use of such images practical, they must be compressed.
  - Approach 1: Lossless compression, in which a compression ratio of about 2:1 can be achieved.
  - Approach 2: Lossy compression, such as JPEG. Trade off compression ratio with quality of decompressed image.

Some examples of image processing

- Edge detection

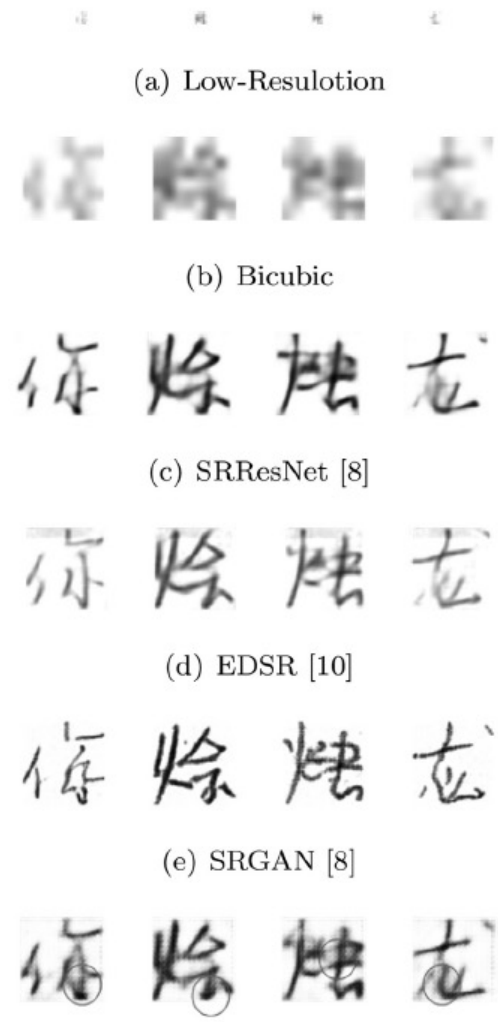
Source Image

Result Image

http://mc.nankai.edu.cn/edge

Some examples of image processing

- Super-resolution


(a) Low-Resulotion
(b) Bicubic
(c) SRResNet [8]
(d) EDSR [10]
(e) SRGAN [8]
(f) The proposed GAC

Zhuang Qian, et al, Generative adversarial classifier for handwriting characters super-resolution, PR 2020

How the human vision system works ?

Human Visual System

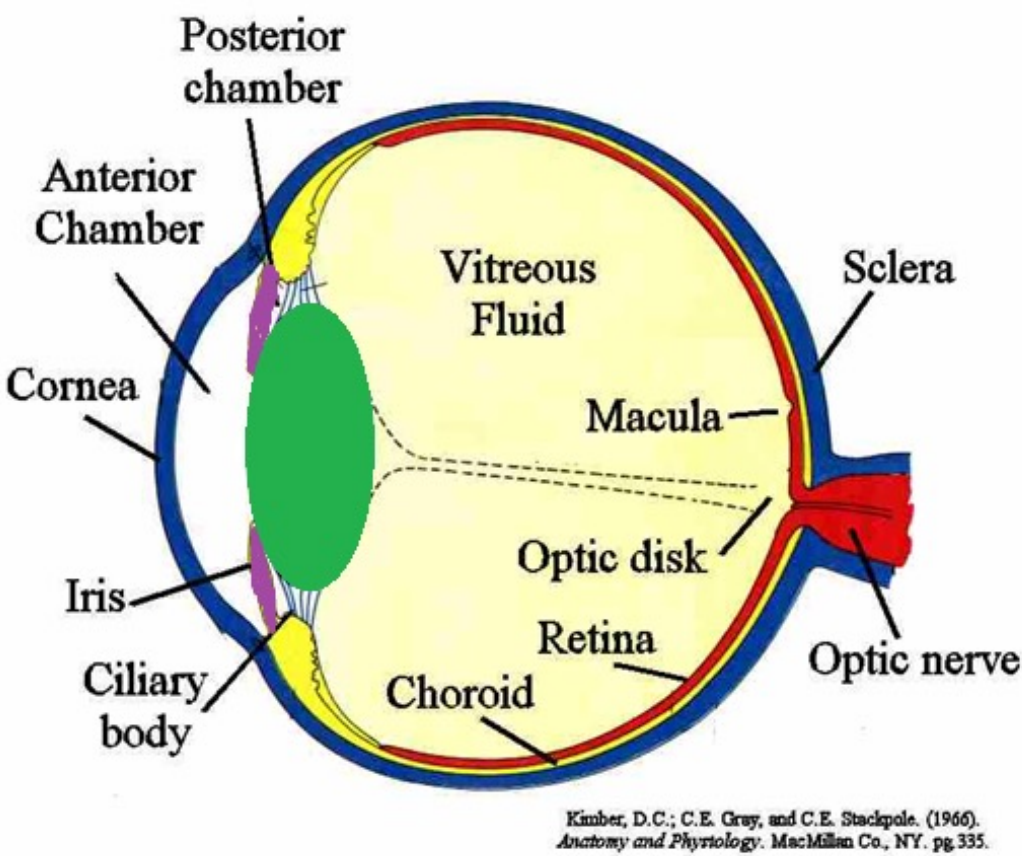
- It takes nature million of years to develop the vision system, so let us learn from it
- Human intuition and analysis
- Mathematical and probabilistic formulations
- AI : simulate the human

Human Vision

- **Allows us to do amazing things like:**
  - Recognize people and objects
  - Navigate through obstacles
  - Understand mood in the scene
  - Imagine stories

- **But still it is not perfect …**
  - Ignores many details
  - Suffers from Illusions

Human Visual System (HVS)



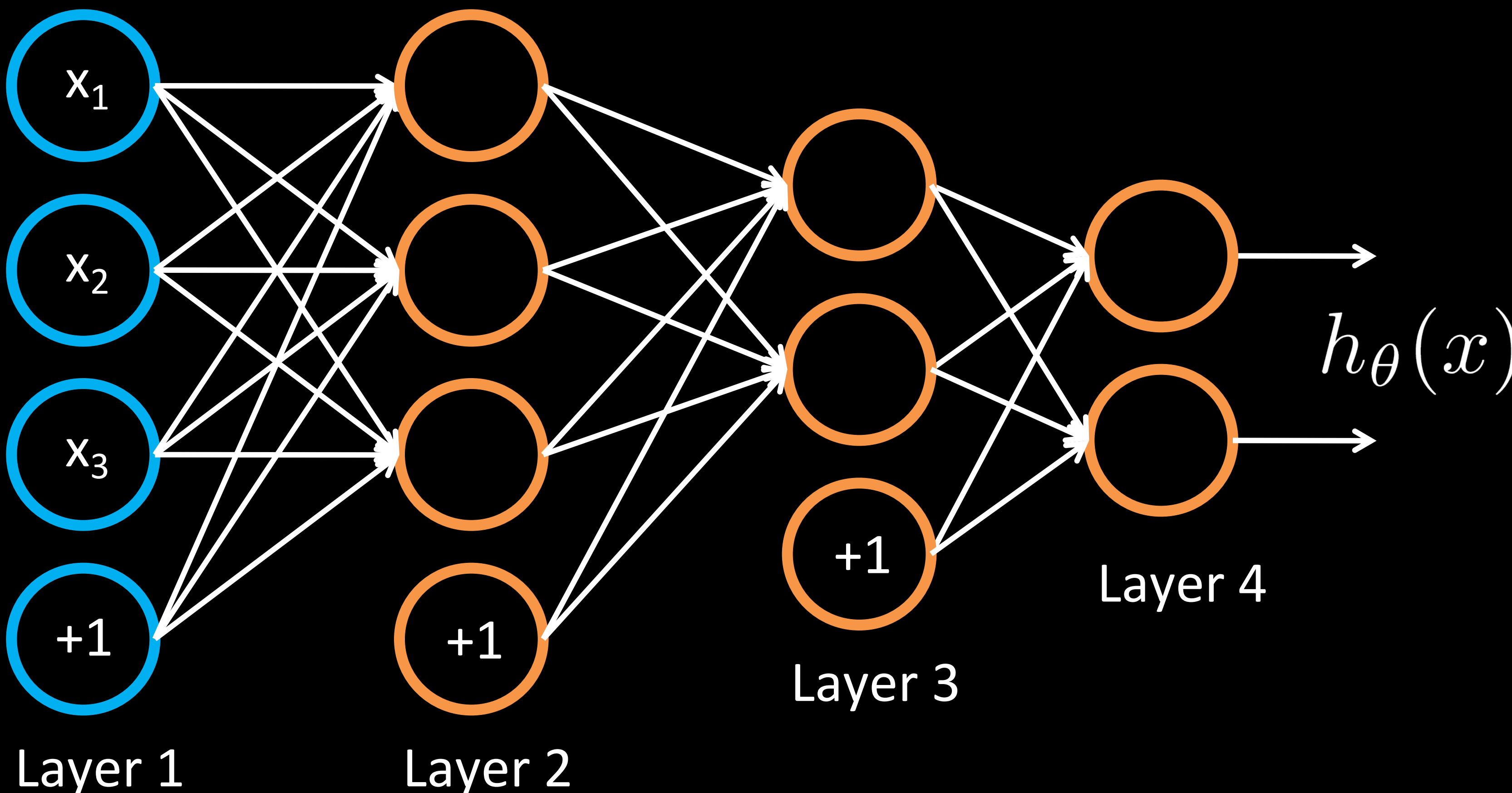
Cross section of right human eye

Eye physiology

- Front of eye covered by a tough and transparent surface (***cornea:角膜***), Remaining outer cover is opaque (***sclera***) "the white of the eye."
- The ***choroid*** is a layer containing blood capillaries, serve as nutrition
  - Light entering the cornea is focused to the retina by a ***lens*** that changes shape under muscular control
  - The ***iris(虹膜,pupil:瞳孔)*** acts as a diaphragm to control the amount of entering light
- The ***retina* (视网膜)** is composed of a network of light receptors /sensors: rods and cones
  - Rods: general and black-and-white view
  - Cones: color and details
  - transform into electrical impulses that ultimately are decoded by the brain

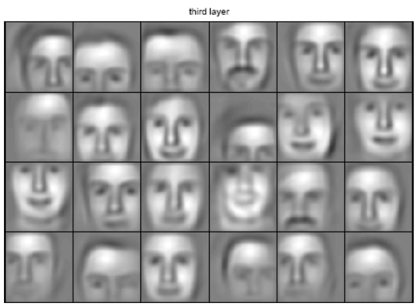
Deep Neural Network

Example 4 layer network with 2 output units:



Why feature hierarchies

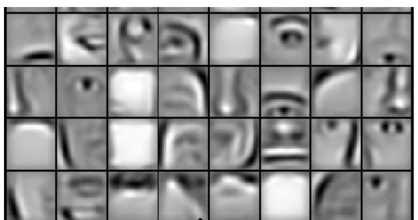
④



object models

Deep layers (4~5)

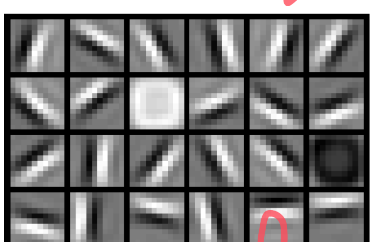
③



object parts
(combination
of edges)

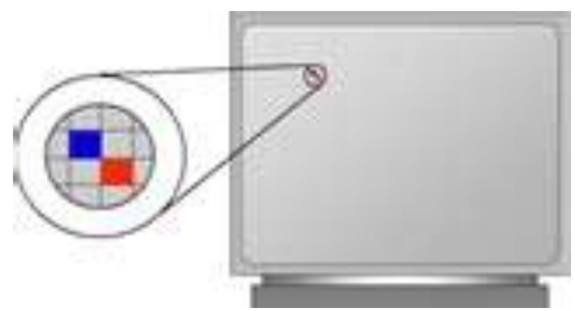
3~4

②



edges

256×256



layer 1~3

①

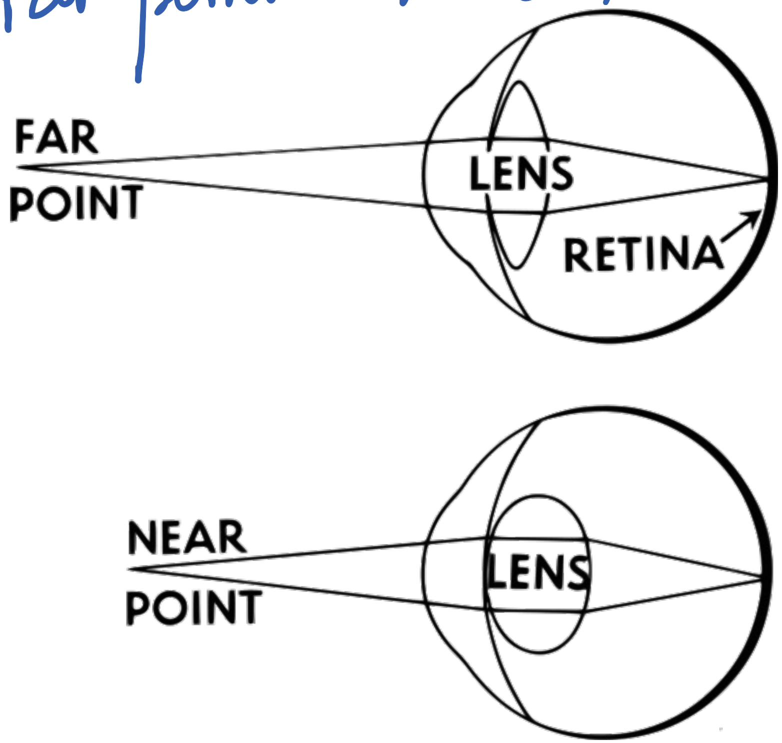
pixels

512×512
width height

Near point → Lens thick → focal lengths smaller (Near-sightedness)

Image formation in the eye

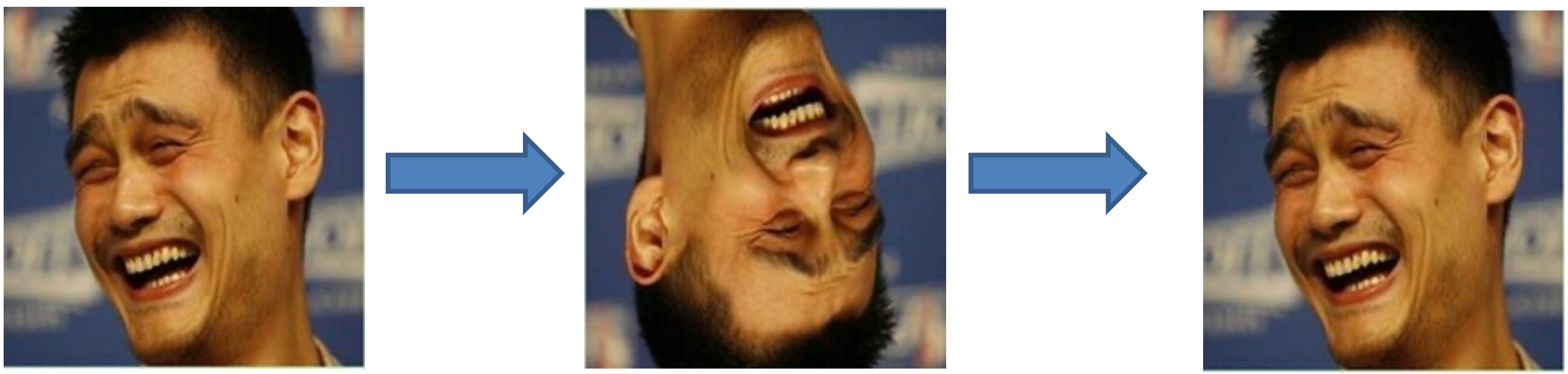
For point → Thiner lens → focal lengths longer (Far-sightedness)



- The lens of humans eye is flexible and its shape is controlled by the tension of the fibers of the ciliary body
- The *focal length* of the lens is: 17-14 mm (min. to max. refractive power)
- **Accommodation** is the process by which the eye changes optical power to maintain a clear image or focus on an object as its distance varies.

Farsightedness
Nearsightedness

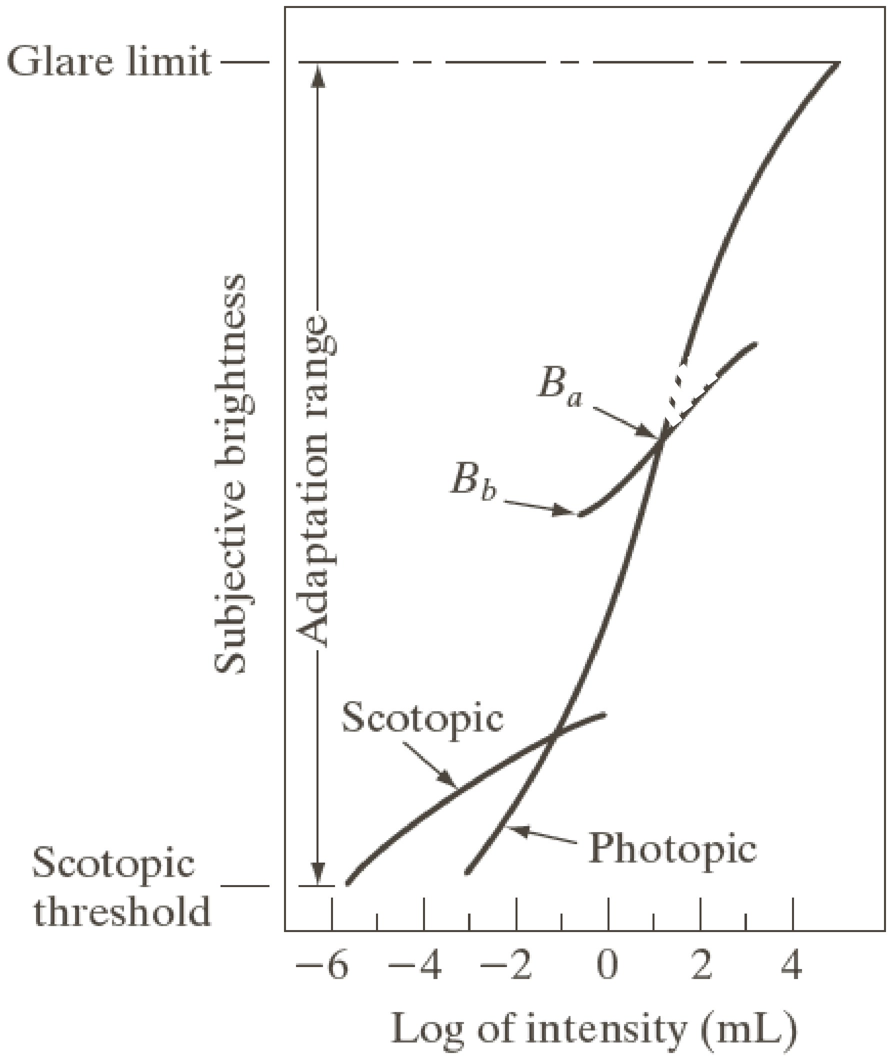
Human Visual System (HVS)



Brightness adaptation and discrimination

- Range of intensity levels to which the HVS can adapt is in the order 10^{10} of magnitude from scotopic threshold to glare limit
- The HVS cannot operate over such a range simultaneously; rather, it accomplishes this large variation by changing its overall sensitivity (brightness adaptation)
- Subjective brightness (intensity as perceived by HVS) is a log function of incident light intensity

Brightness adaptation and discrimination



- Ba = *brightness adaptation level*
- All stimuli of intensity lower than Bb are perceived as indistinguishable black
- Upper portion of the curve (dashed) not actually restricted but, if extended too far, simply raises the brightness adaptation layer

Brightness adaptation and discrimination

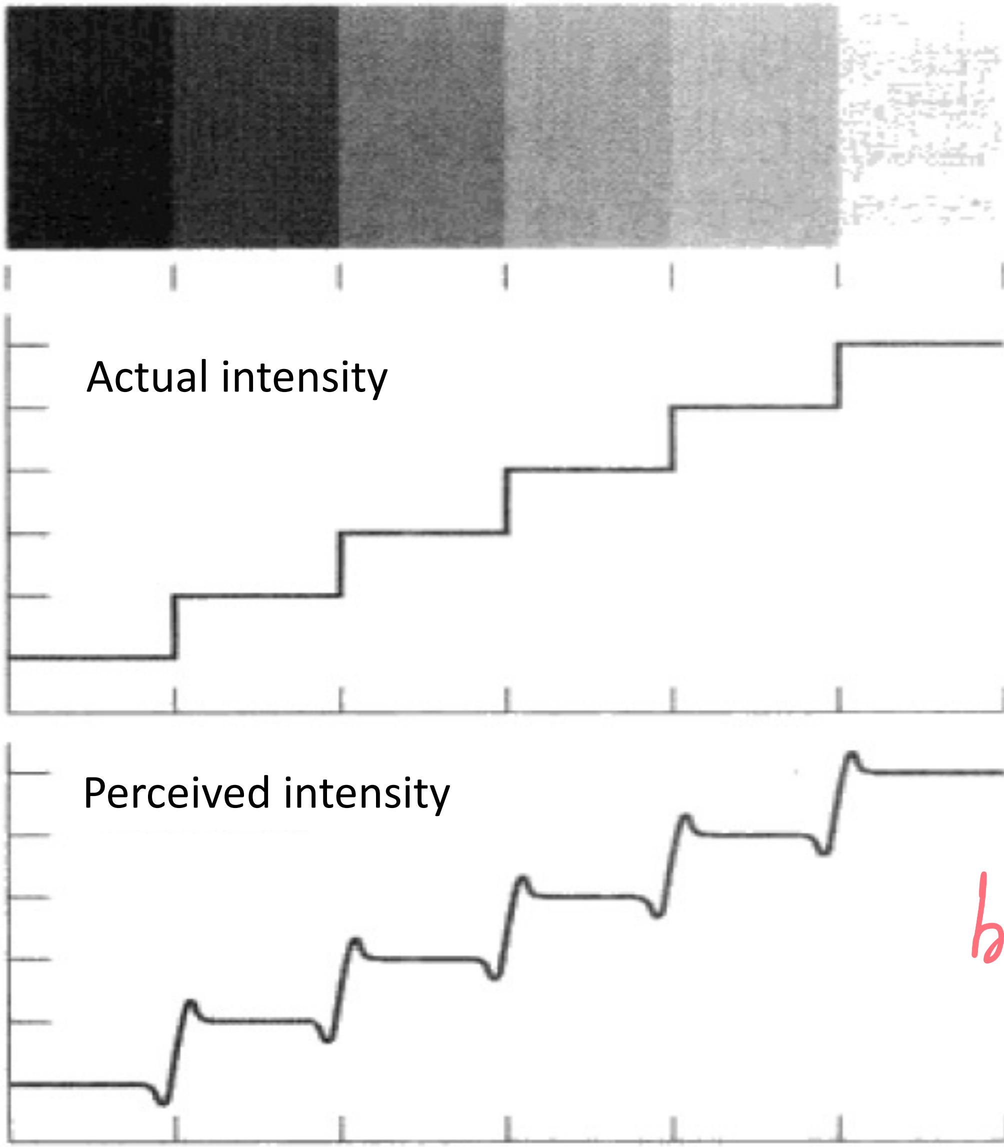
- The total range of intensity levels it can discriminate simultaneously is rather small compared with the total adaptation range
- At low illumination levels (scotopic vision) the brightness discrimination is poor whereas it is better when vision is function of cones
- If the intensity of a light spot source is not flashing, but is allowed to vary incrementally from never being perceived to always been perceived, the typical observer can discern *one-two dozen intensity levels*

HOWEVER

simultaneous contrast phenomenon

- The two small squares have exactly the same intensity, but the right-hand appears darker as its background is lighter
- This shows that *perceived brightness is not a simple function of intensity*

Mach bands phenomenon



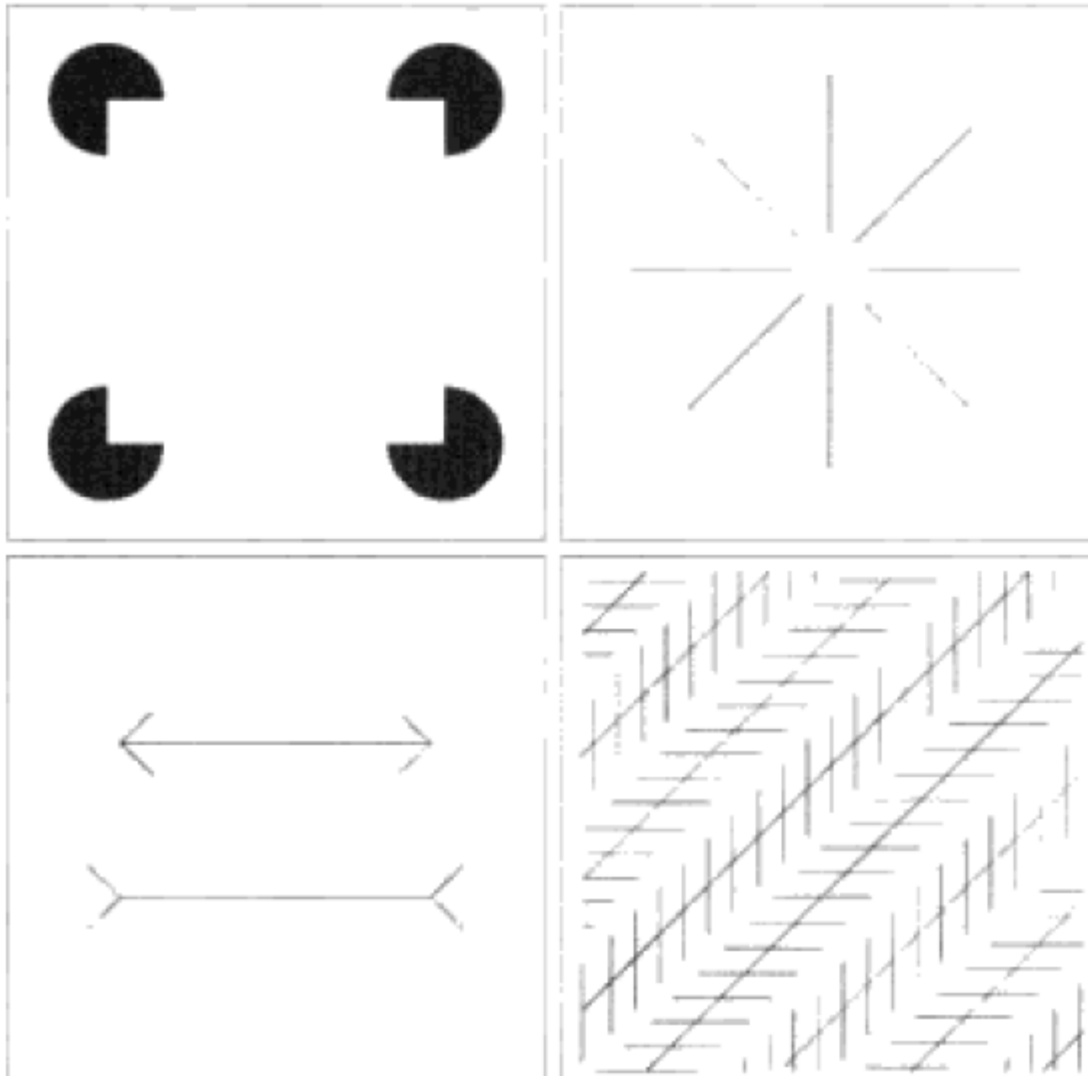
Along the boundary between adjacent shades of grey in the Mach bands illusion, lateral inhibition makes the darker area falsely appear even darker and the lighter area falsely appear even lighter.

The vision system tend to undershoot and overshoot around the regions of different intensities.

Light → dark brightness lower

Dark → Light brightness higher

Some optical illusions



What is light?

What is light ?

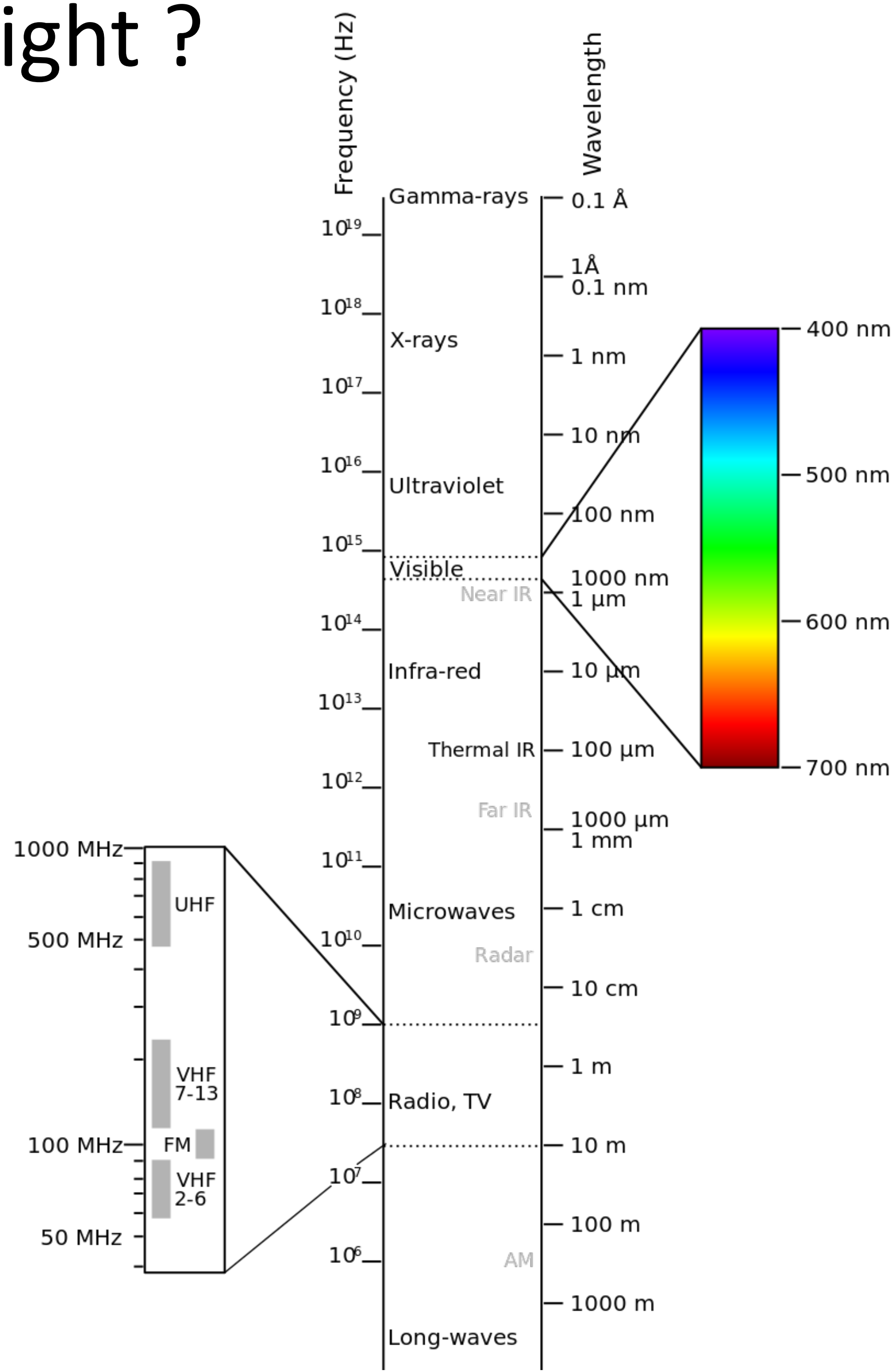
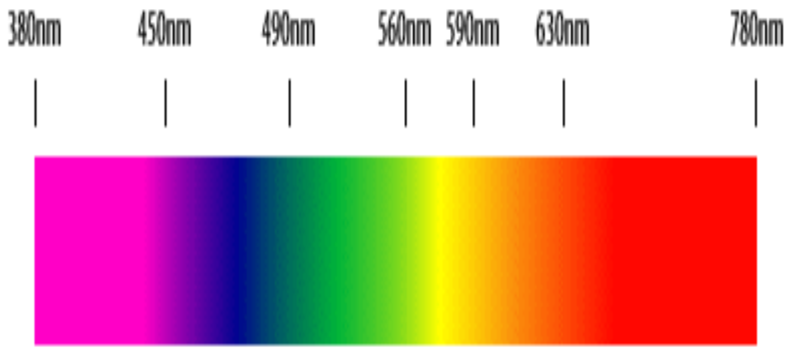


Diagram of the electromagnetic spectrum

From Wikipedia.org

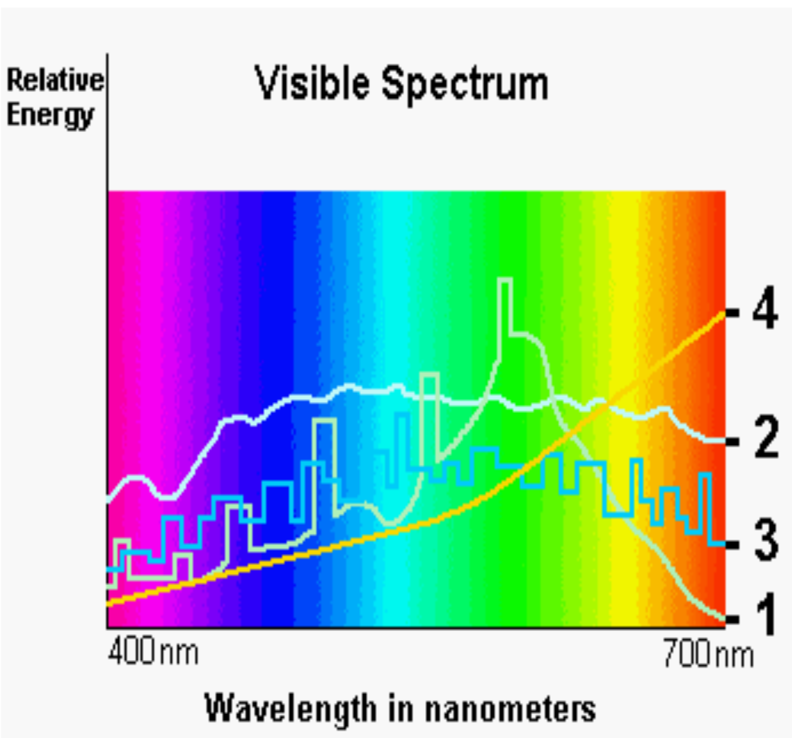
Visible light



- Webster's dictionary definition of light : *[visible] Light is radiant energy which, by its action on the organs of vision, enables them to perform their function of sight*
- EM radiation lying in the relatively narrow band of about 350 to 780 nm

Visible light

- A light source could be characterized by the rate of radiant energy (*in watt*) emitted at a particular wavelength



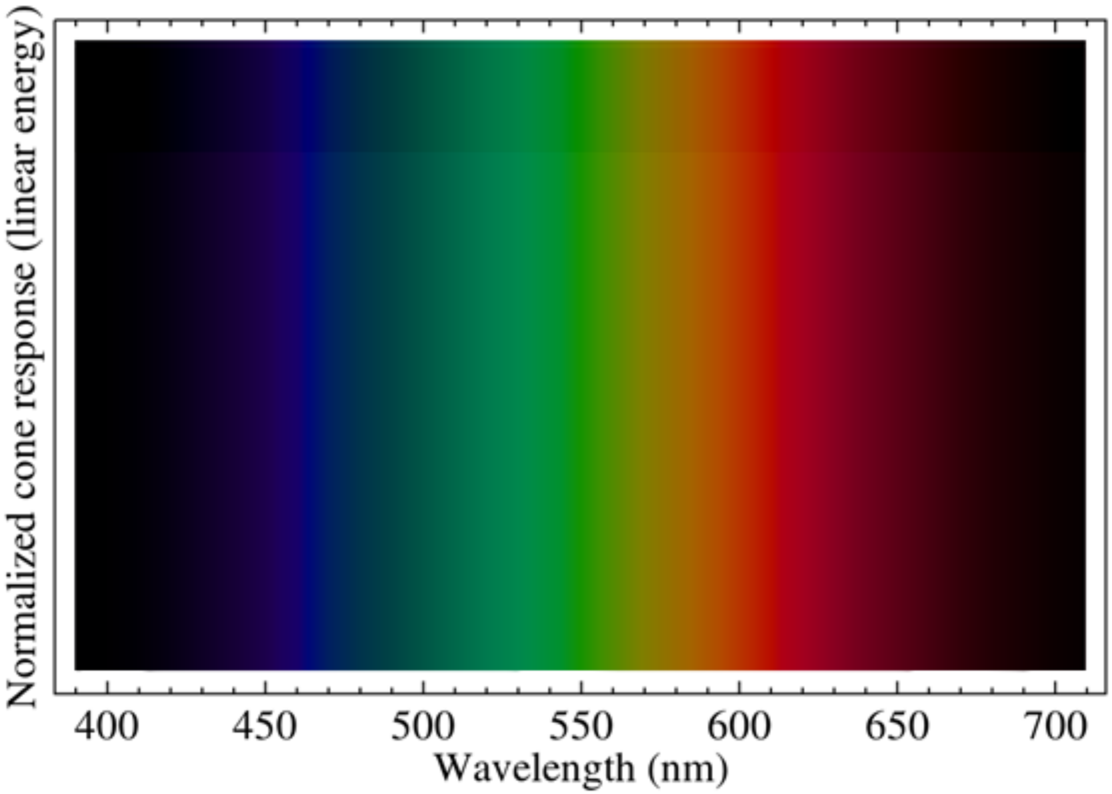
1. fluorescent

2. daylight
3. HMI (Hydrargyrum medium-arc iodide lamp)

4. tungsten

How do we sense colors ?

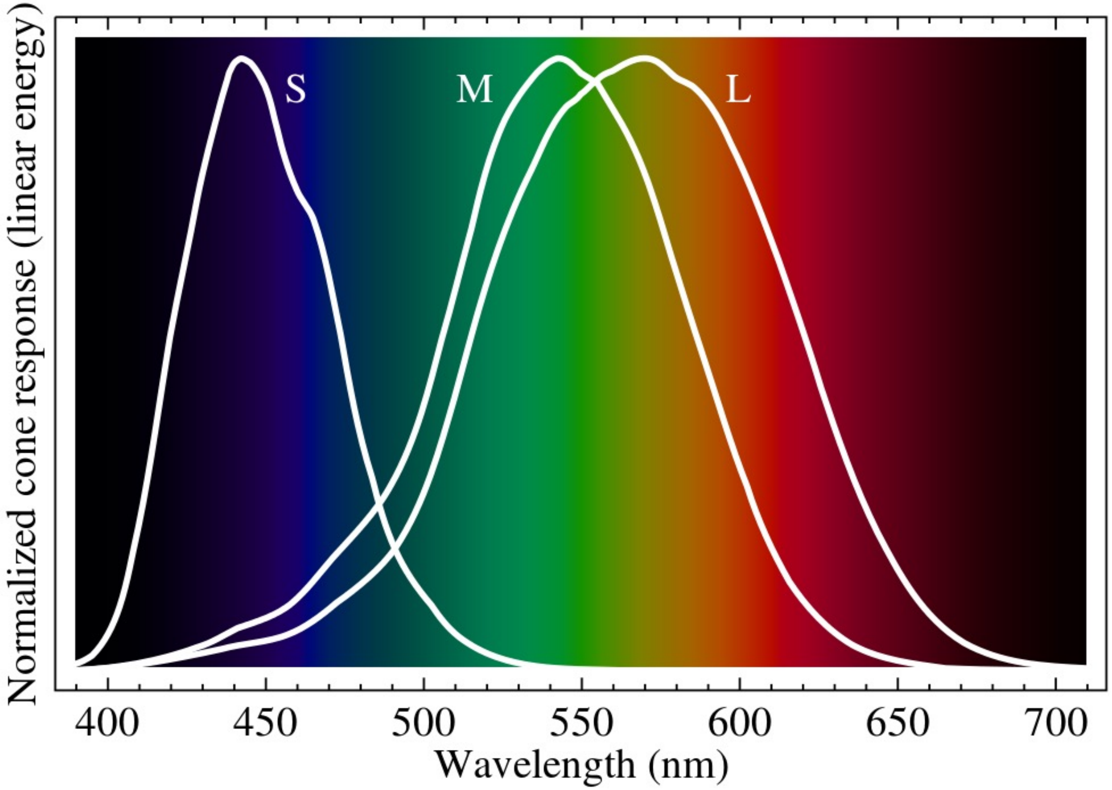
- Human eye can distinguish about **10 million different colors**.



How do we sense colors ?

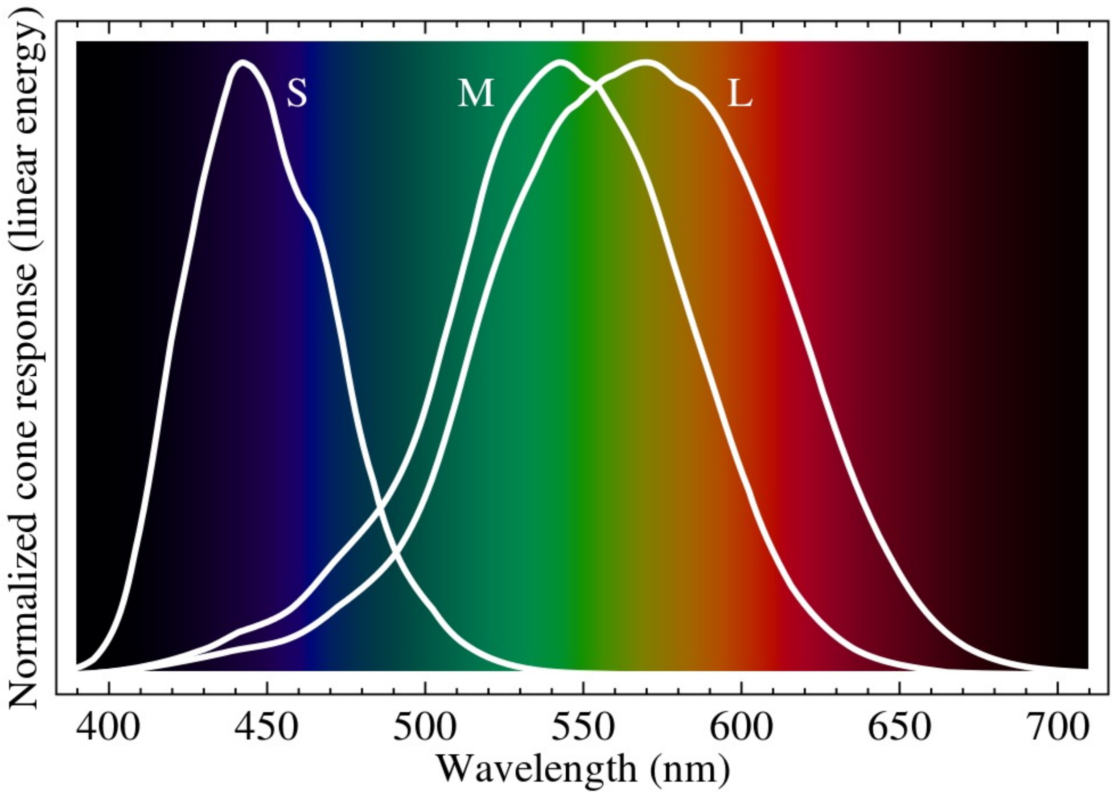
- The retina has **three sensing cons**: red(Long wave), green(Medium wave), and blue (Short wave)

Normalized response spectra of human cones to monochromatic spectral stimuli, with wavelength given in nanometers



How do we sense colors ?

- **White** is sensed when the light stimulates all **three types of cones** cells in the eye in **nearly equal amounts**.
- Yellow is perceived when the L cones are stimulated slightly more than the M cones



Perceptual descriptors of light sensation

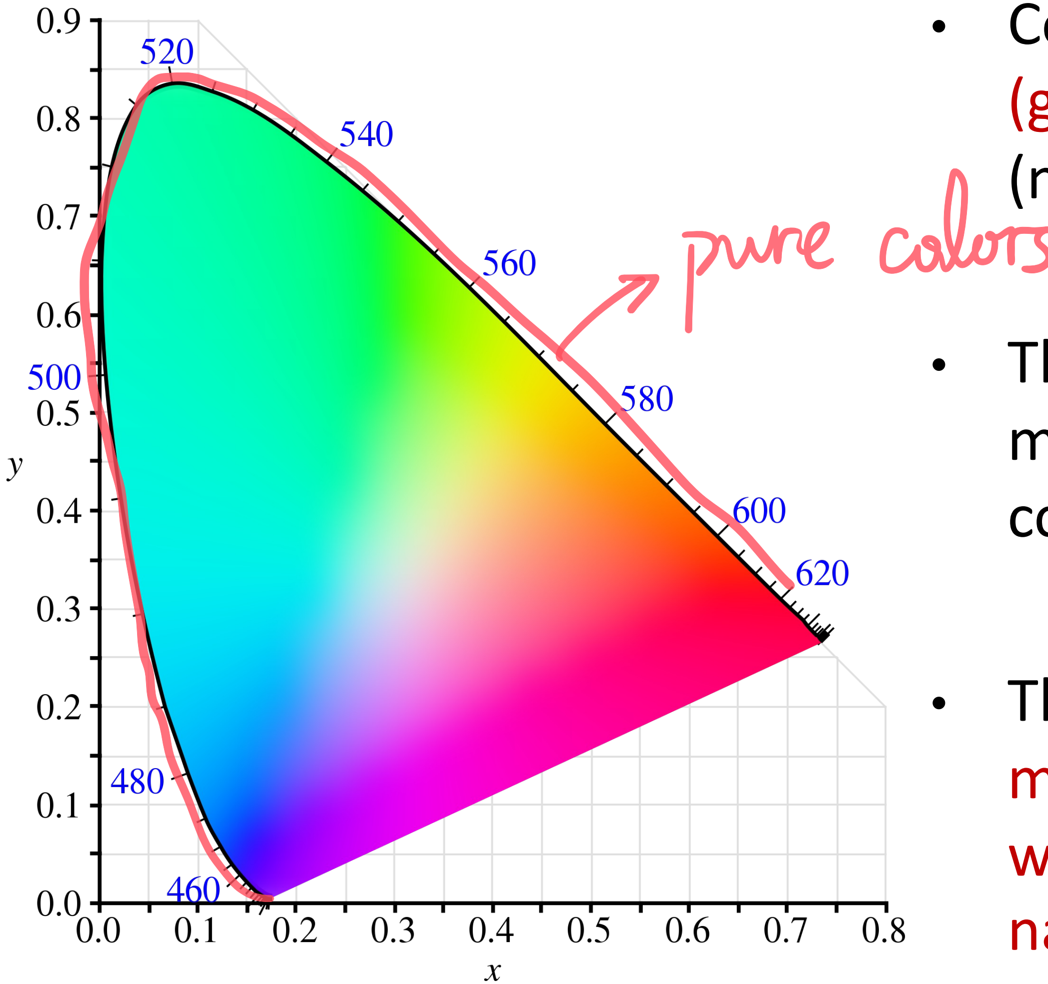
- Three common perceptual descriptors of light sensation:
  - **Hue**: perception of difference between colors, an attribute associated with the dominant wavelength in a mixture of light waves
  - **Saturation**: vividness of hue, whiteness of light. Pure spectrum colors are fully saturated.
  - **Intensity**: perceptual measure of physical intensity (brightness)
- The concept of color can be divided into two parts: brightness and chromaticity. (hue & saturation)
  - For example, the color white is a bright color, while the color grey is a less bright version of white ➔ the chromaticity of white and grey are the same while their brightness differs.

Color space chromaticity diagram

- Since cons respond to three different ranges of wavelengths a full plot of all visible colors is a three-dimensional figure.

- How can we get all visible colors?

Color space chromaticity diagram

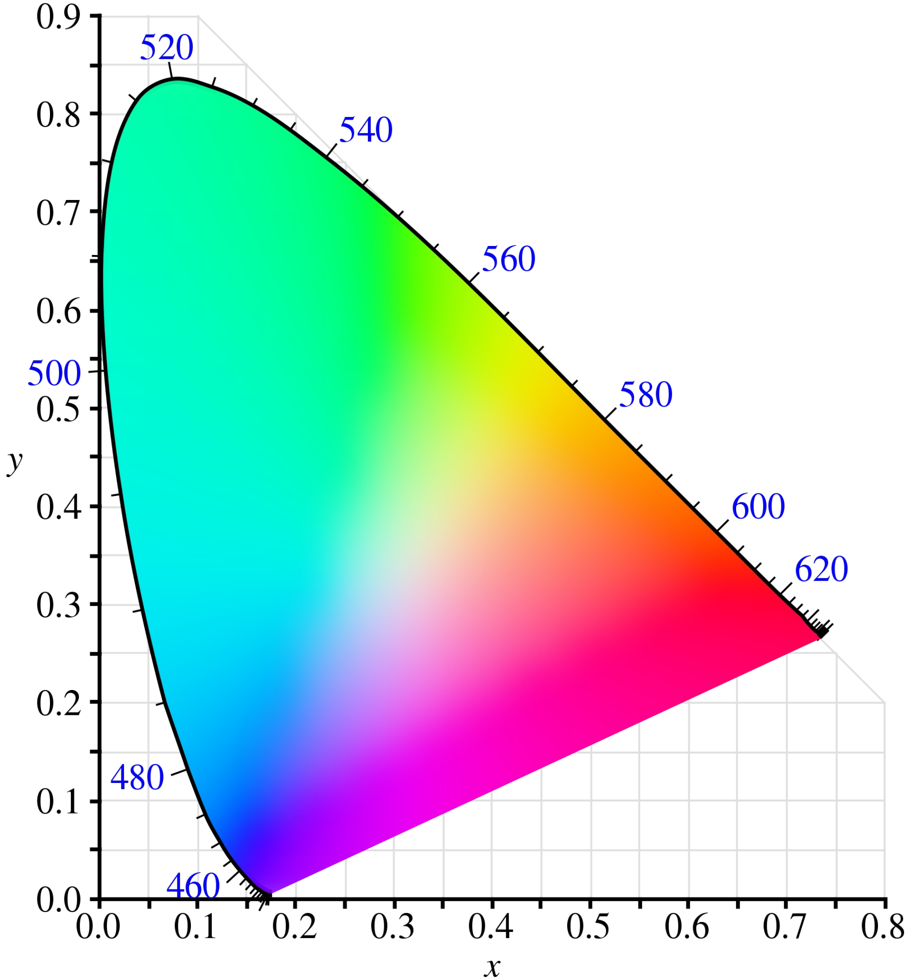


- Color coefficients: x (red) , y (green) and z (blue) (normalization: x+y+z=1)

- This diagram displays the maximally saturated bright colors that can be "perceived".

- The outer boundary is the monochromatic locus, with wavelengths shown in nanometers.

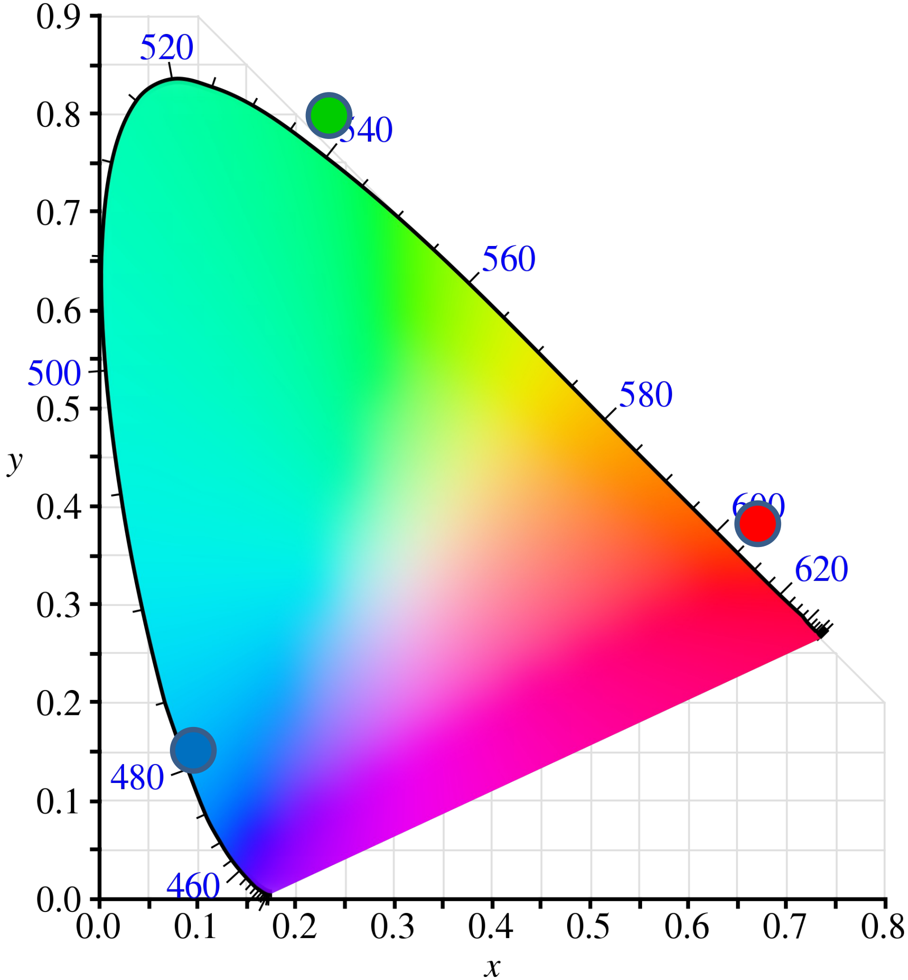
Color space chromaticity diagram



- The line between any two points represents the colors that could be obtained by combining different percentages of these two colors

- So what if we have three elementary points ? What is the range of colors that these three points could represent ?

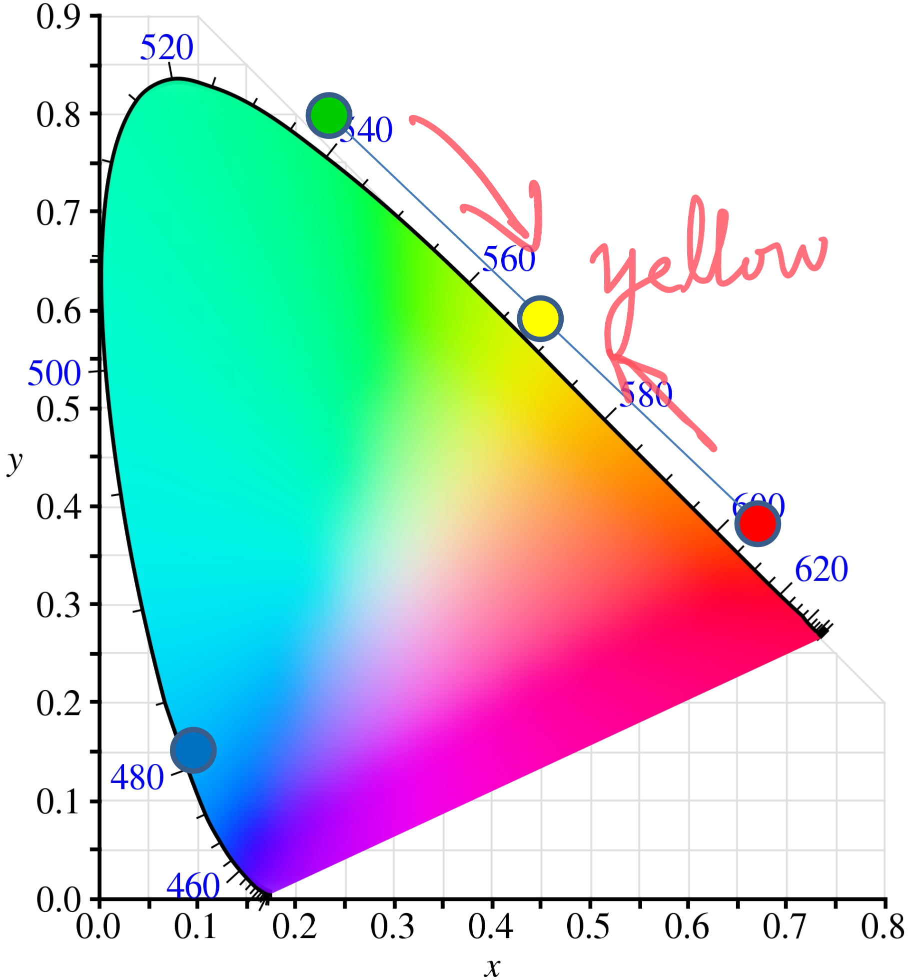
Color space chromaticity diagram



- The line between any two points represents the colors that could be obtained by combining different percentages of these two colors

- So what if we have three elementary points ? What is the range of colors that these three points could represent ?

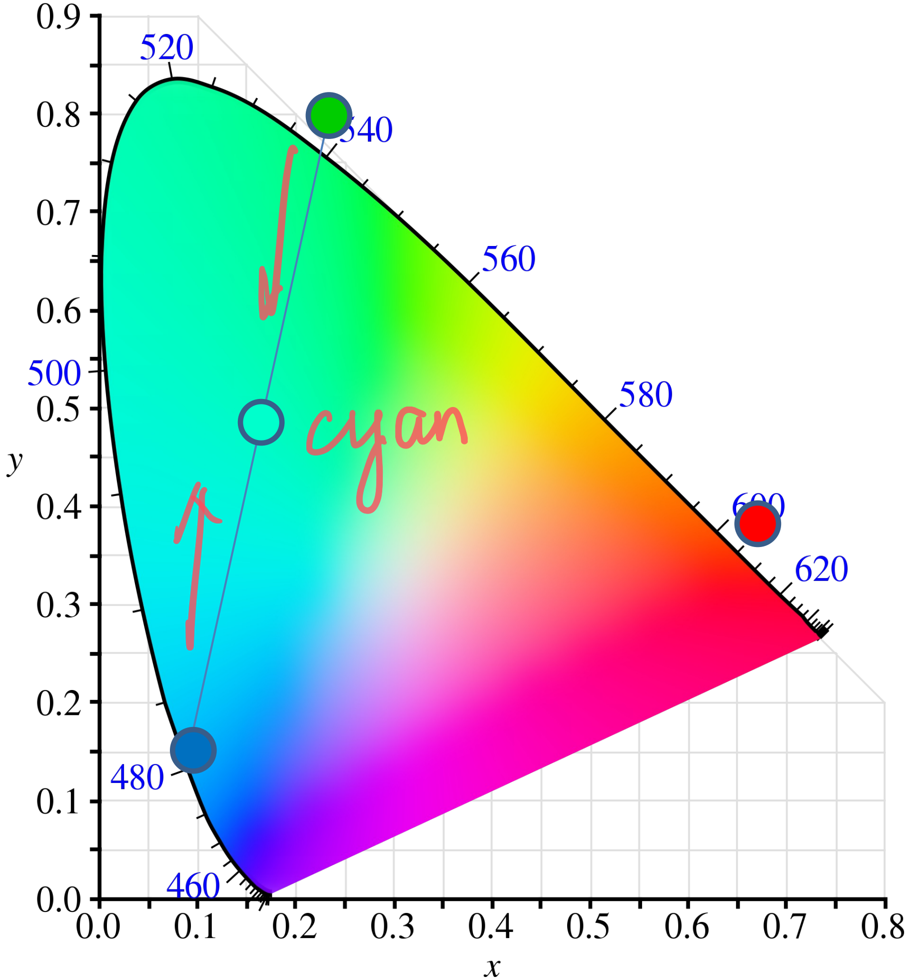
Color space chromaticity diagram



- The line between any two points represents the colors that could be obtained by combining different percentages of these two colors

- So what if we have three elementary points ? What is the range of colors that these three points could represent ?

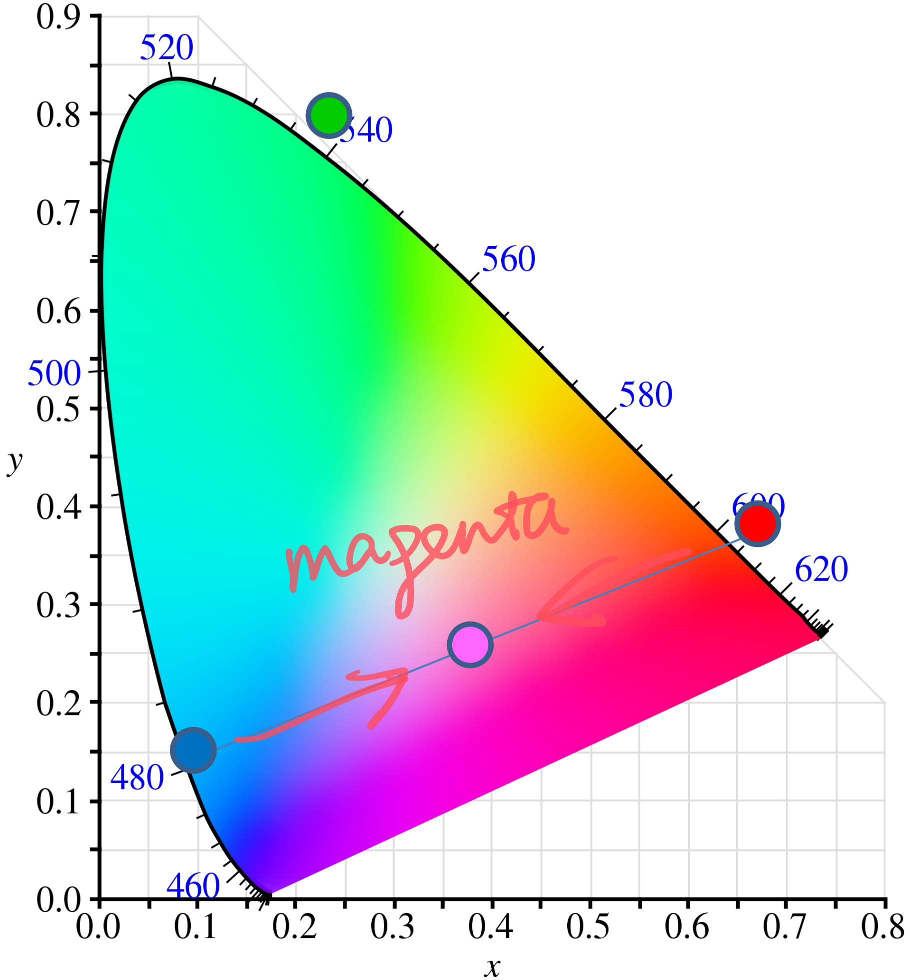
Color space chromaticity diagram



- The line between any two points represents the colors that could be obtained by combining different percentages of these two colors

- So what if we have three elementary points ? What is the range of colors that these three points could represent ?

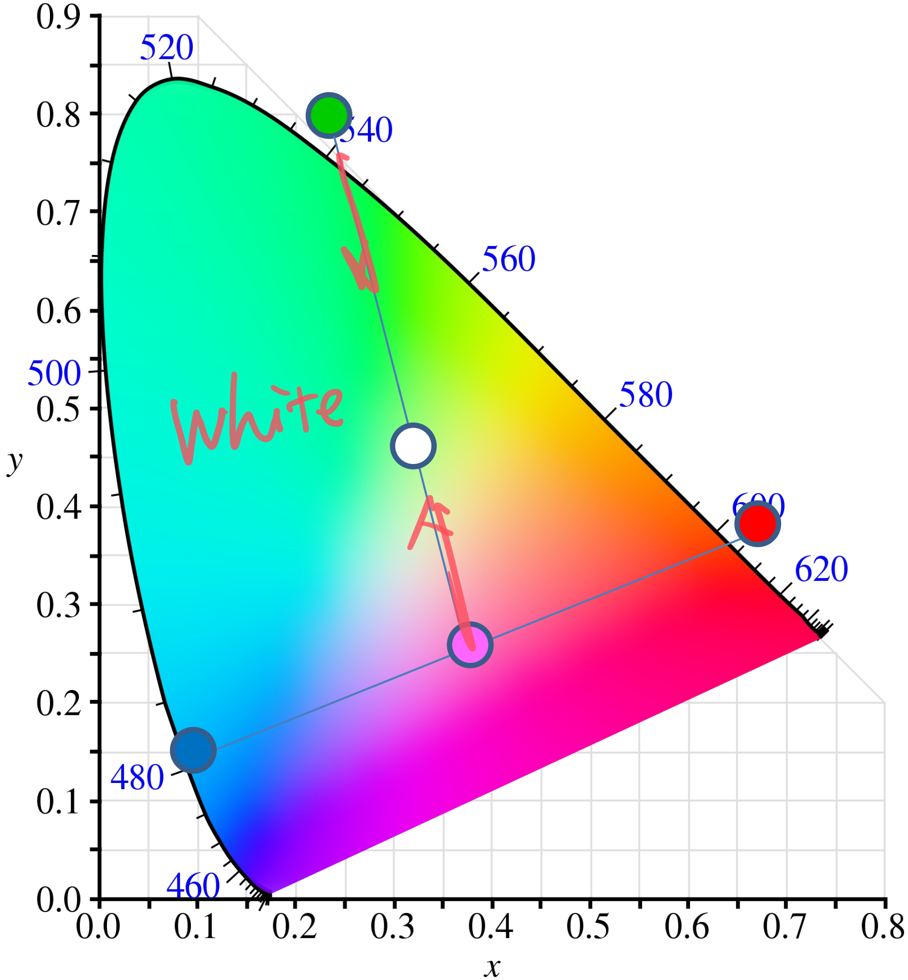
Color space chromaticity diagram



- The line between any two points represents the colors that could be obtained by combining different percentages of these two colors

- So what if we have three elementary points ? What is the range of colors that these three points could represent ?

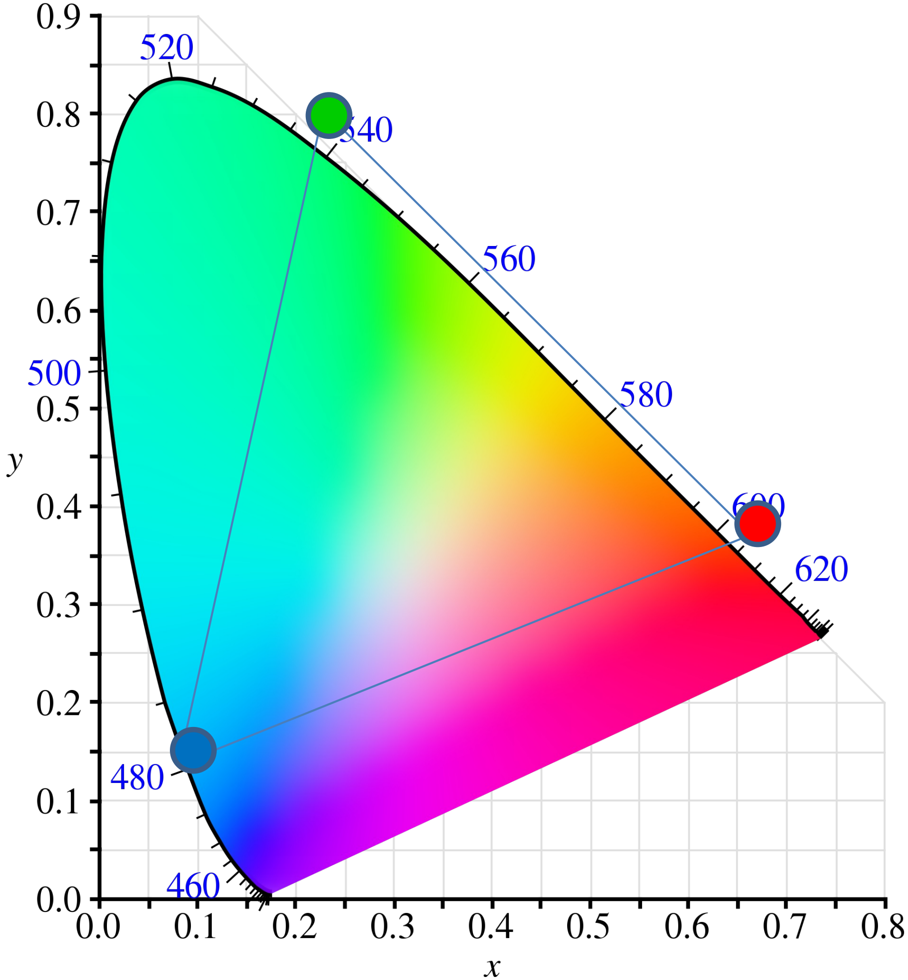
Color space chromaticity diagram



- The line between any two points represents the colors that could be obtained by combining different percentages of these two colors

- So what if we have three elementary points ? What is the range of colors that these three points could represent ?

Gamut map



- Gamut map is a subset of colors that can be produced by a device.

- Colors that a screen displays are specified using RGB, so the colors outside the RGB gamut are not displayed properly.

(Coefficients won't be inside 0~1 outside the RGB gamut)

Three-Color Theory

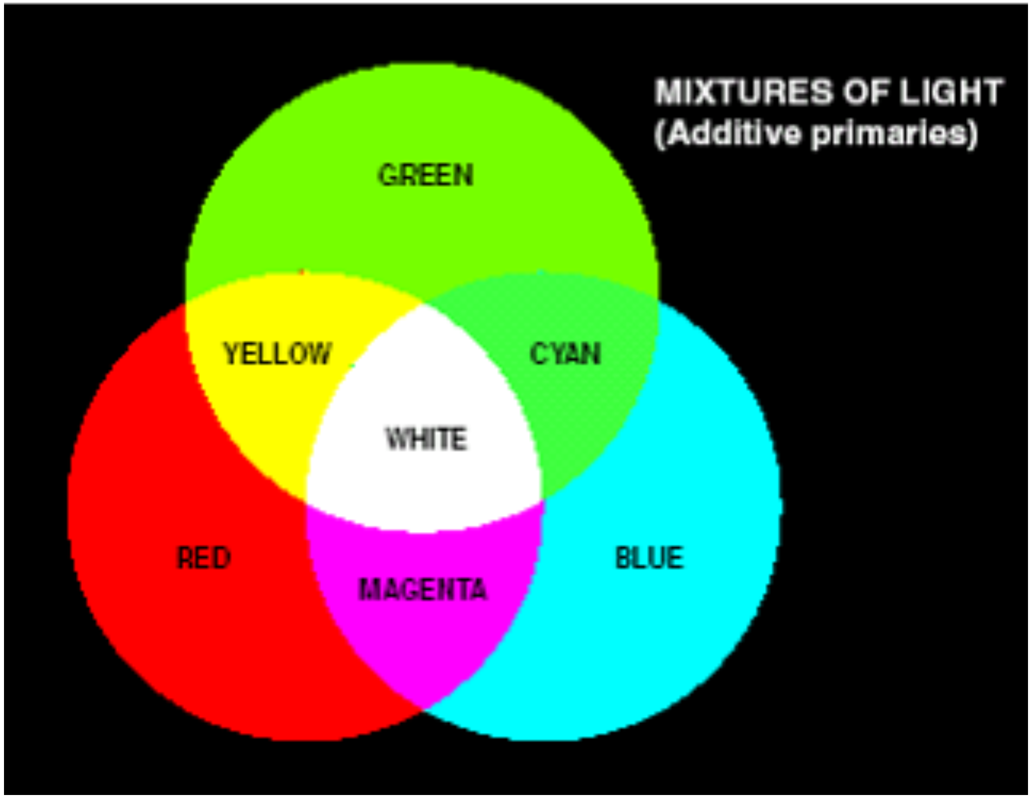
- For the purpose of producing colors three primary colors can be selected
- By mixing appropriate amounts of three colored a specific set of colors can be produced,
- However, three primary colors cannot reproduce all possible colors.
  - Why ? : The colour space limitations
  - What should we do to reproduce bigger set ?

Three-Color Theory

- Primary colors can be added to produce other colors, such as the secondary colors (in the additive model, for light) :
  - magenta : red plus blue
  - cyan : green plus blue
  - yellow : red plus green
  - Mixing equally the three primaries produces white

- Colour TV/monitor is an example of the additive nature of light colours

Three-Color Theory



Colour Models

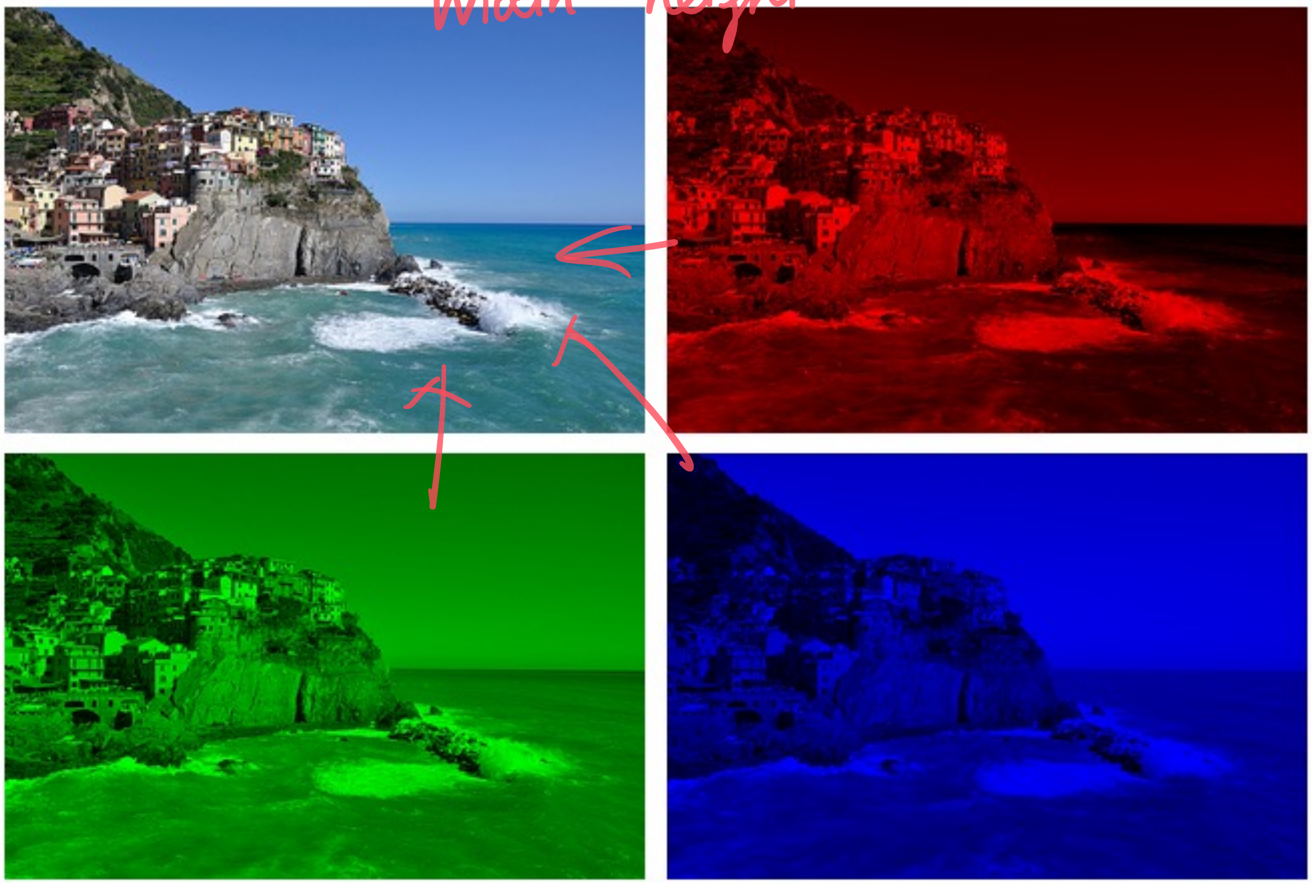
Colour Models

- Common colour models:
  - RGB model for colour monitors and colour video cameras
  - CMY (cyan, magenta, yellow) and CMYK (cyan, magenta, yellow, black) models for colour printing
  - HSI (hue, saturation, intensity) model corresponds well with human perception of colour
  - YCbCr (luminance, colour differences) colour space is used in analog and digital video

RGB Image Components

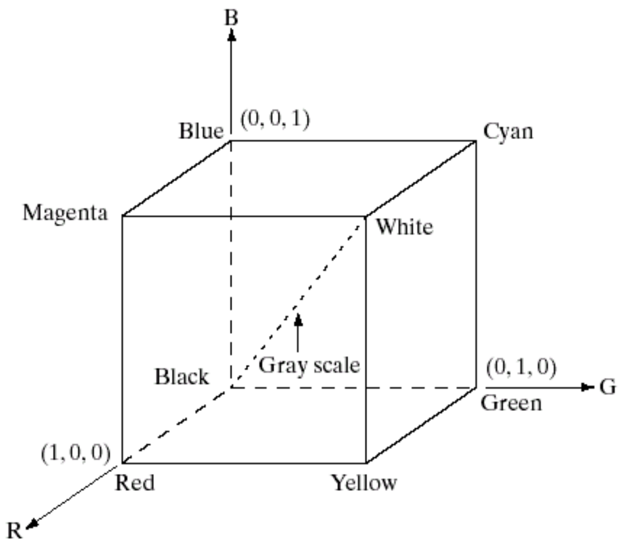
512 x 512 x 3

Width height



RGB Color Space

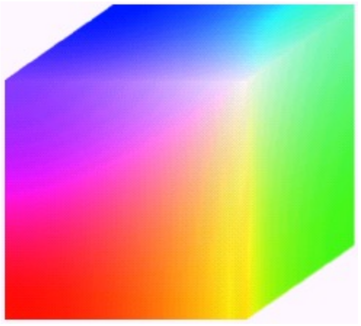
- Colour values normalized to [0,1]
- Cartesian coordinates are for the three colour components : red, green, blue
- **Primary colours**: are placed at the corners of the colour cube
- **Secondary colours**: cyan, magenta, and yellow at the other corners of colour cube
- **Black** is at the origin
- **White** is at the corner farthest from the origin



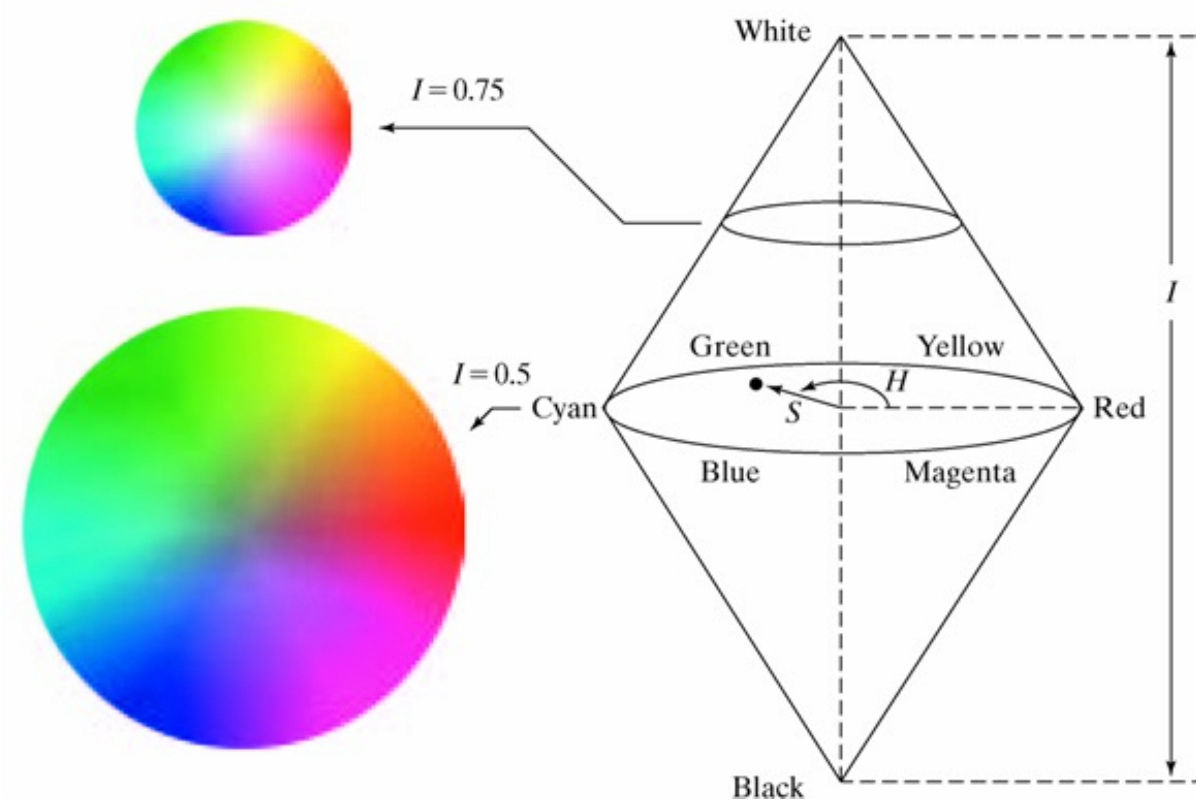
RGB Color Space

- **Gray scale** (points of equal RGB values) extends along the line from black to white→ **gray is fully described by one value.**
- Full-colour pixels –in most applications-needs 24-bit (**colour depth**), having 8 bits in each colour
- Total number of colours in 24-bit RGB image is

$$\left(2^8\right)^3 = 16,777,216$$



HSL Colour Space: Circular Color Planes



HSI Image Components

Hue

Saturation

Intensity

CMY and CMYK Color Spaces

- Most devices that deposit colored pigments on paper, such as color printers, require CMY data input ➔ perform an RGB to CMY conversion internally
- The conversion operation:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- Equal amounts of the pigment primaries, cyan, magenta, and yellow, should produce black. In practice, combining these colors for printing produces a muddy-looking black. So a fourth color, *black,* is added, giving rise to the CMYK color model.
- Mapping between RGB, HSV, YCrCB

Colour Models

- Common colour models:
  - RGB model for colour monitors and colour video cameras
  - CMY (cyan, magenta, yellow) and CMYK (cyan, magenta, yellow, black) models for colour printing
  - HSI (hue, saturation, intensity) model corresponds well with human perception of colour
  - YCbCr (luminance, colour differences) colour space is used in analog and digital video