EZ ME

www.ezcheck.me

- Attendance verification with EZCheck.me app (In the break)
- Download for iPhone

  https://apps.apple.com/us/app/ezcheck-me/id1472247186

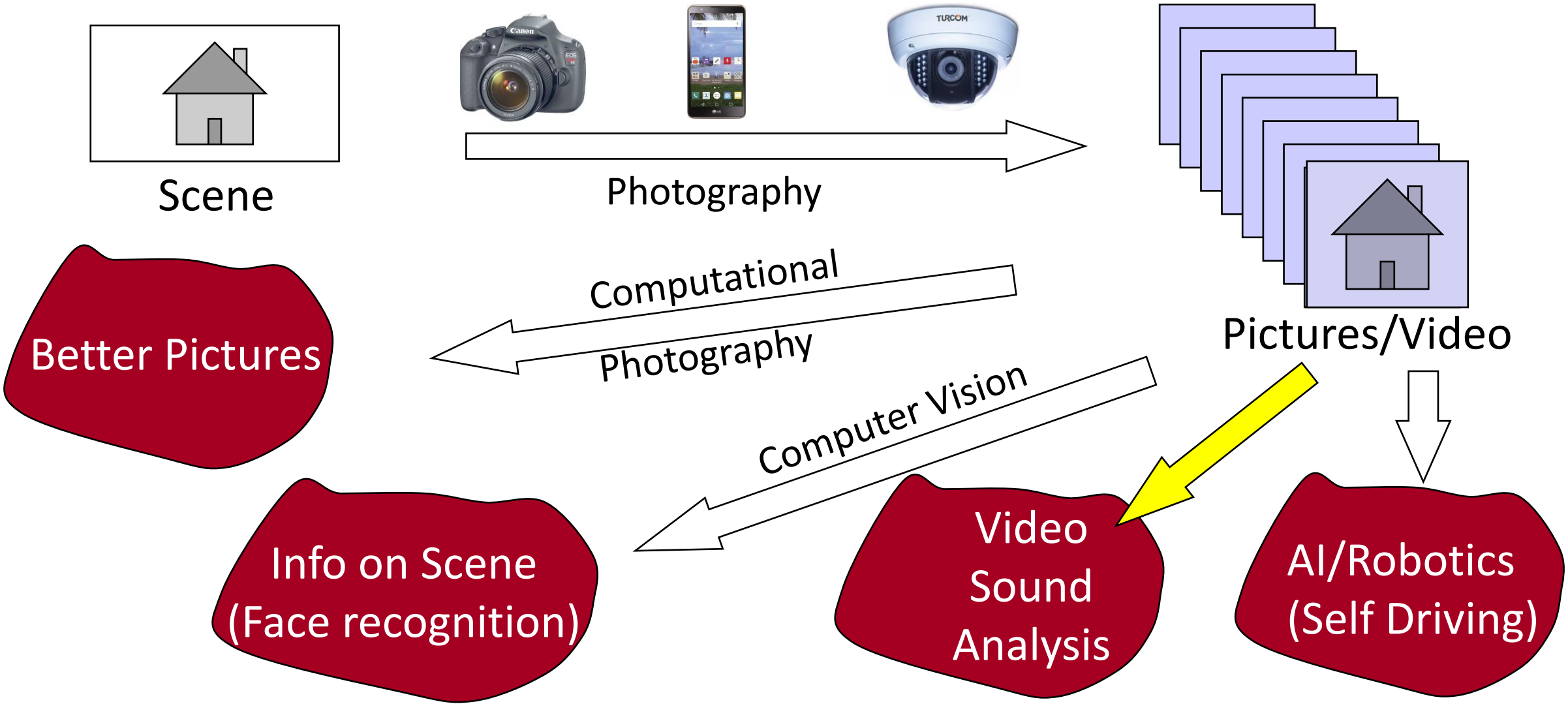
- Download for Android

  https://play.google.com/store/apps/details?id=me.ezcheck

- Please enable location, as we need to verify you attended from the classroom.

Image Processing Course

Computer Vision, Computational Photography



Areas of Computer Science

Classical CS - Invent Algorithms

Complexity, **Algorithms**

Computer Graphics

Data Bases

Computer Architectures

Cryptography

Programming Languages, Compilers

Networking

Cyber

Learn from Data - AI (Artificial Intelligence)

Computer Vision ⟵

Speech Analysis

Natural Language (NLP)

Deep Learning (NN)

Stage 1 – Preprocessing, Tokens

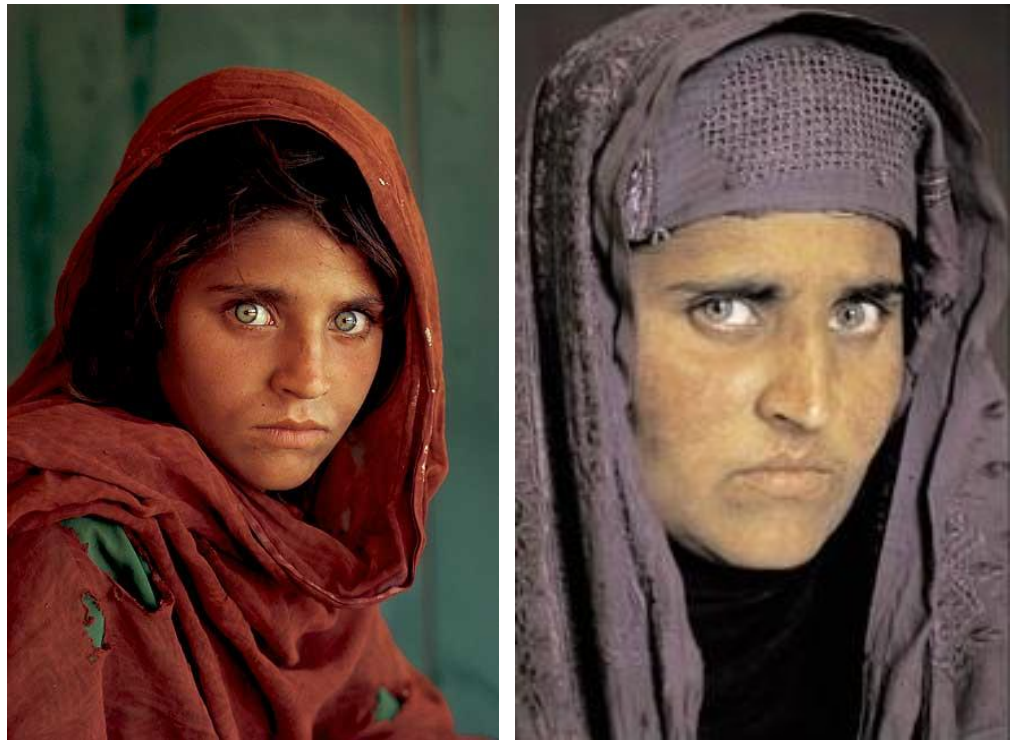
Stage 2 – Deep Learning

Autonomous driving, Medical diagnosis

ChatGPT … Replace Humans

In AI: Different appearance can have same meaning

- We (people) know most of the times to recognize people
- We (people) can not Explain how we do it
  - No one could invent a face recognition algorithm
- AI can learn to recognize people from many examples, doing a better job than humans!



1984 Same Person 2000

National Geographic: “Afghan Girl”

Nature: Vision = Intelligence = Moving

- Only intelligent and moving organisms can see!
  - Bacteria & Plants do not see
- Visual recognition at early development
  - Babies recognize and track the mother very early
- Most of the human brain is involved in visual processing

Predator or Prey - Vision Used for Survival

Predator: Hunts its food
Eyes directed forward

Prey: Escape Predators
Eye directed sideways

Applications: Image Enhancement

Dehazing

High Dynamic Range Imaging

Panoramic Stereo Mosaics

Developed at HUJI

A previous course exercise

Video Synopsis (Also by Huji)

Original: 9 hours

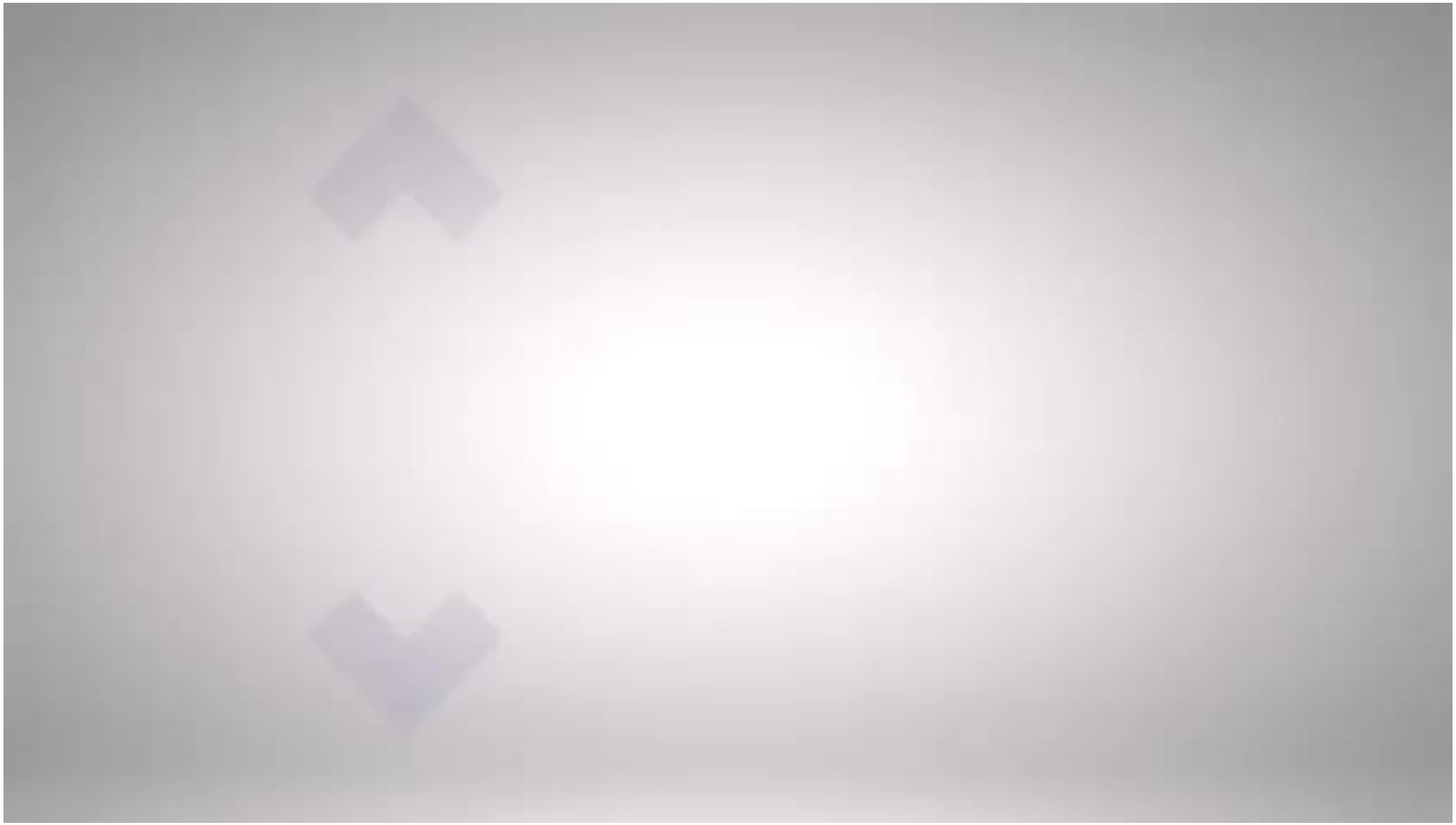
Synopsis: 30 seconds

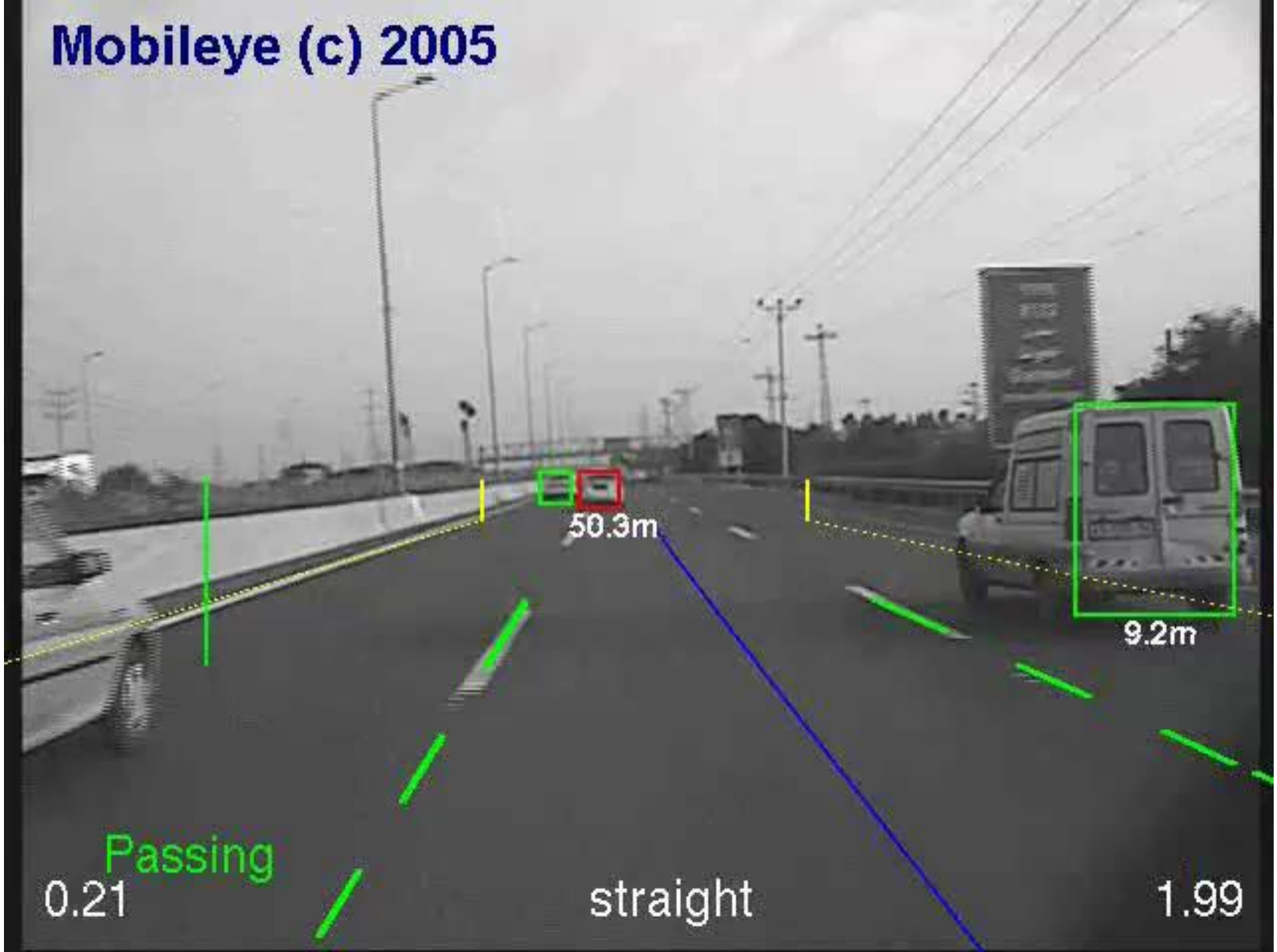
- Commercialized by Briefcam
- Acquired by Canon 2018

BriefCam

Canon

Hebrew University at Night




Mobileye (c) 2005
50.3m
9.2m
Passing
0.21
straight
1.99

The Real Big Data:
(i) Video (ii) Without Photographers

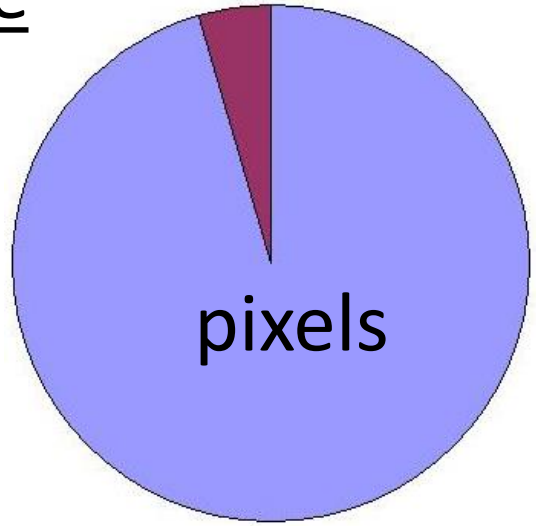
- Surveillance Cameras
- Industrial/car cameras
- Wearable Cameras

<u>Video is ~85% of the digital universe</u>

Surveillance ~70%

Characters (Text, numbers):

Only ~15%



Computer Vision in 2024

Topic Started in ~1964, Nothing Worked for 50 Years

Started to Work ~2014 (Neural Networks)

- Computer vision products are commonly used now :
  - Face Detection is included in every camera
  - Face Recognition in Facebook, Google, & Passport Control
  - Autonomous Driving (MobilEye..) – Robotaxi: Tel-Aviv Oct 2022!
  - Medical Diagnosis
  - OCR (Optical Character Recognition)
  - Image and Video Generation (Deep Fake…)

Image Processing: 2024-25

Teacher: **Shmuel Peleg** peleg@mail.huji.ac.il
Assistant: **Leeyam Gabay** leeyam.gabay@mail.huji.ac.il
Tzars: **Eliahu Horwitz**

Textbooks: Web, Wikipedia, Szeliski (Free download https://szeliski.org/Book/), Moodle

Grading (Average of final grade in course is normalized to 83-85 **for those that attended**):
3 exams: 1 hour & 2 questions each. From the 6 questions, we use the best 5!
Estimated Exam Dates (All at 6-7pm): (1) Sun 1/12; (2) Sun 5/1; (3) Sun 2/2;

Expected Exercises (Python, GITHUB)
5 individual programming. Show you friends & family… ChatGPT, CoPilot,…

Final Grade: 3 Exams: 75% (25% each); Exercises: 23% (non-equal weight); Attendance: 2%

אינפורמציה חשובה

- אין מועד ב' כי משקל כל בוחן פחות מ 26% מהציון הסופי
  - יינתן מועד מיוחד (ג') למבחן חסר באישור ועדת הוראה (מילואים, מחלה, ...)
- אנחנו נלחמים בהעתקות
  - בשנה שעברה הוגשו תלונות נגד 10% מהסטודנטים!
- בונוס נוכחות: 1 נקודה על 50% נוכחות. 2 נקודות על 70% (חישוב אישי להיעדרות מוצדקת מעל 30%)

Relevant “Vision-AI” Courses

1st Semester

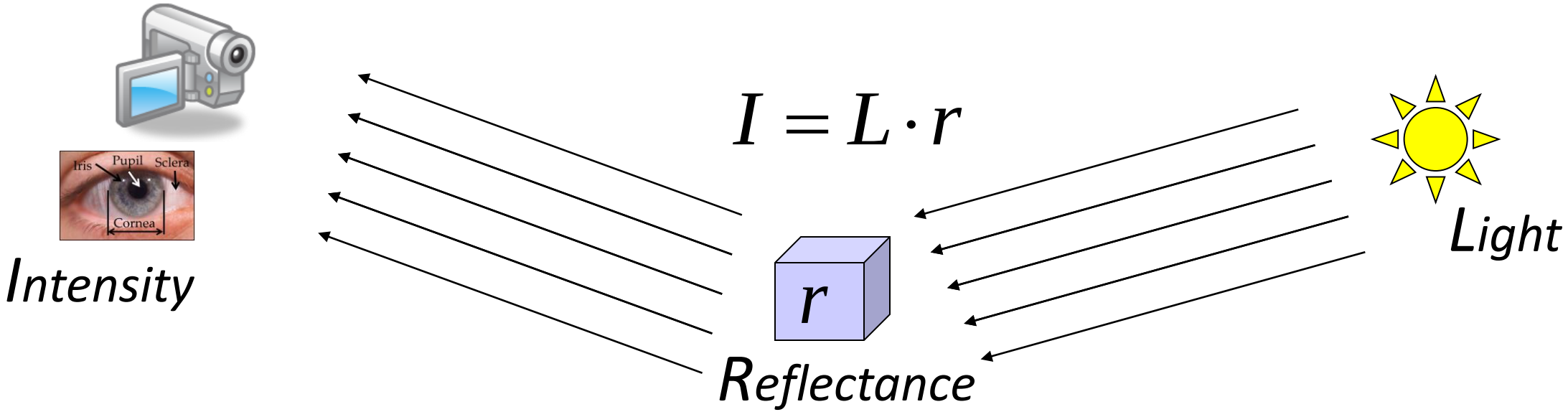
- **Image Processing (Peleg)**
- IML (Introduction to Machine Learning) (Hoshen)
- Medical Image Processing (Joskowicz)

2nd Semester

- IML (Introduction to Machine Learning) (Stanovsky)
- Introduction to Deep Learning (Fattal)
- Advanced Course in Machine Learning (Hoshen)
- **3D Computer Vision – 3D Geometry (Werman)**
- SLAM (**S**imultaneous **L**ocalization **A**nd **M**apping) - Video Navigation (Vivanti)
- Introduction to Speech Processing (Adi)

Image Formation: Luminance

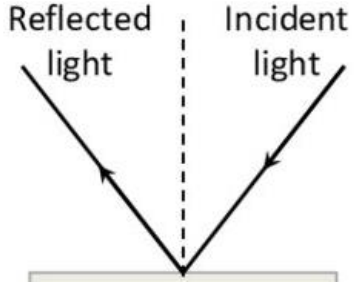
- Light is emitted by light sources
- Light is reflected from objects
- Reflected light is sensed by eye or by camera

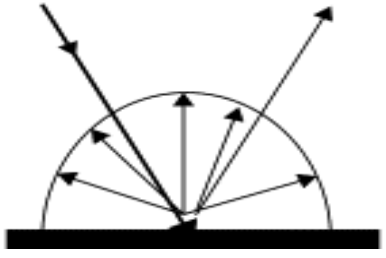


Depends on lighting & viewing directions; color

Reflectance: Shining (Specular) vs. Matt (Diffuse)

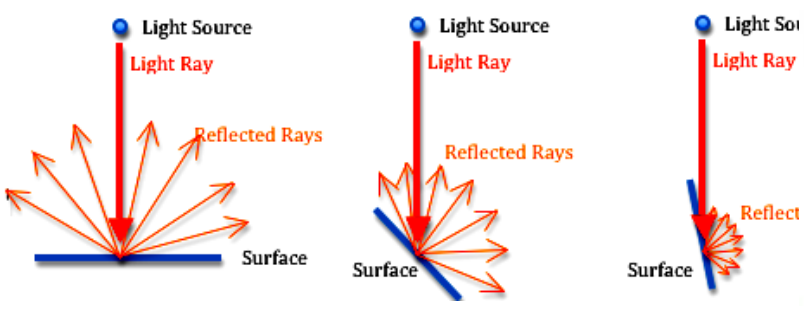
A ray is reflected to have same angle with normal

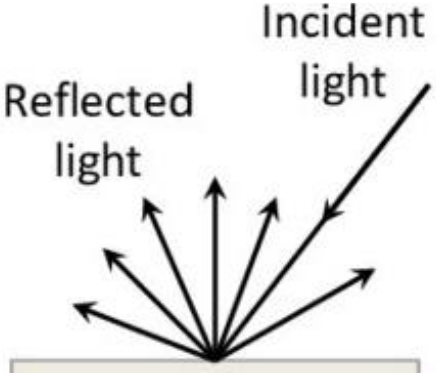




Mixed surface

Each surface point has equal reflection to all directions

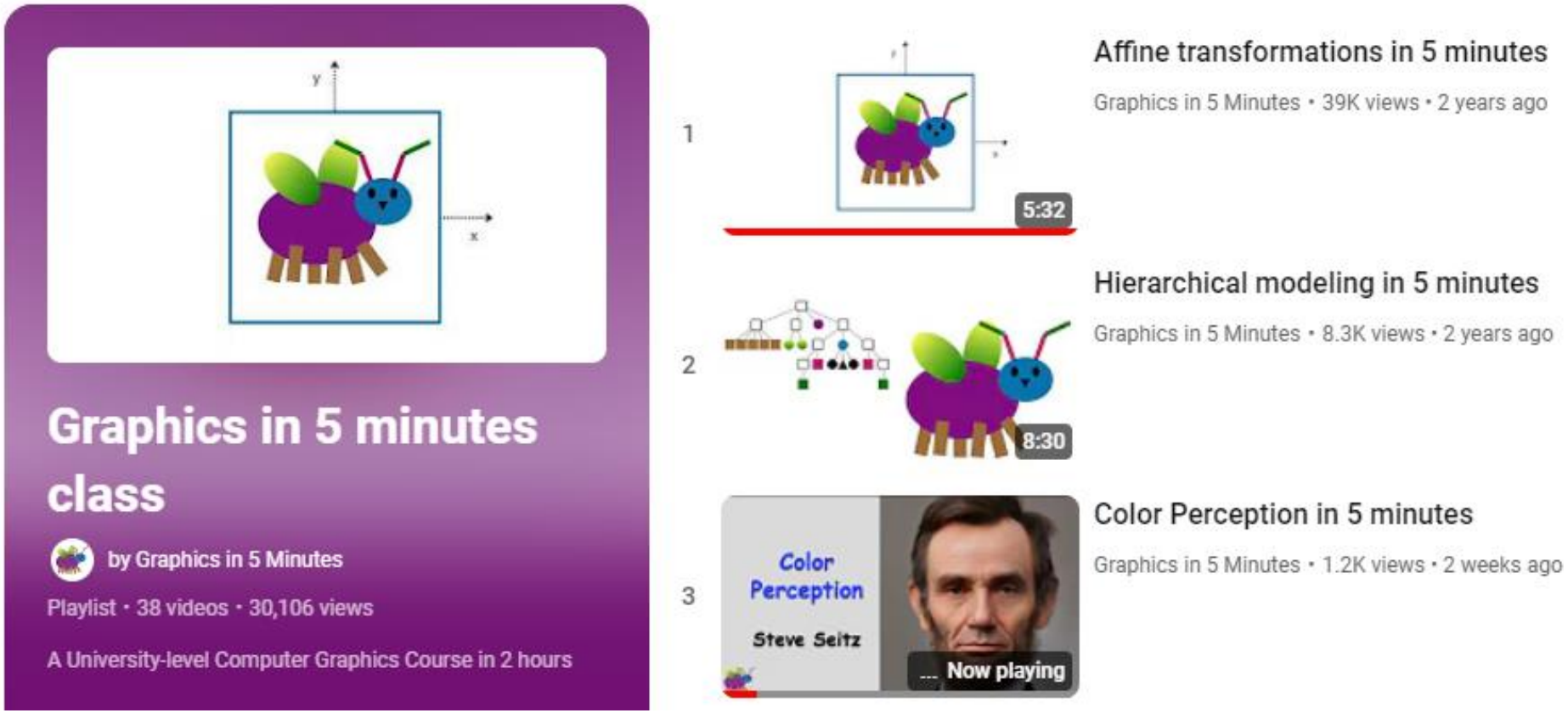




Graphics in 5 Minutes Class

YouTube video clips by Steve Seitz

https://www.youtube.com/playlist?list=PLWfDJ5nla8UpwShx-lzLJqcp575fKpsSO



5 Minute Clips

- Specular reflection
- Diffuse Reflection

10

Specular reflection in 5 minutes

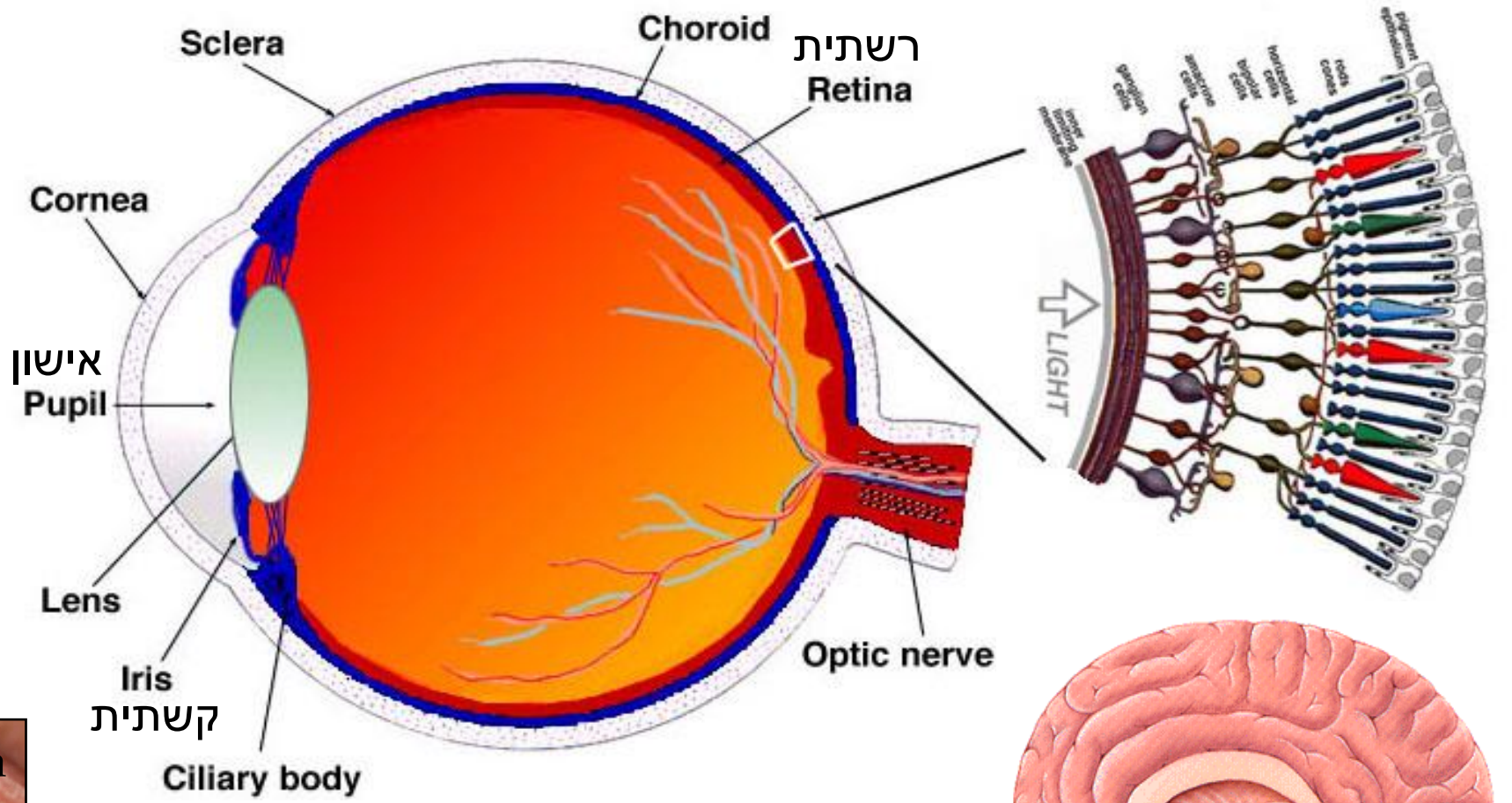
Graphics in 5 Minutes • 4.9K views • 2 years ago

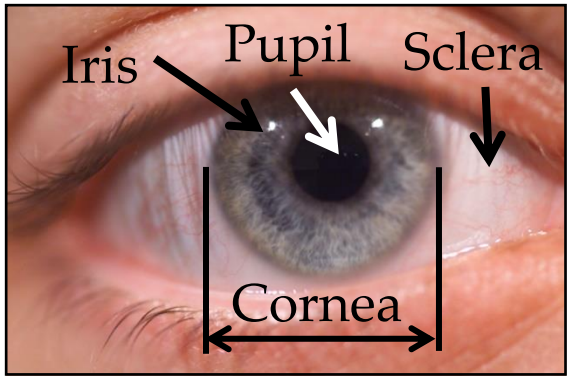
11

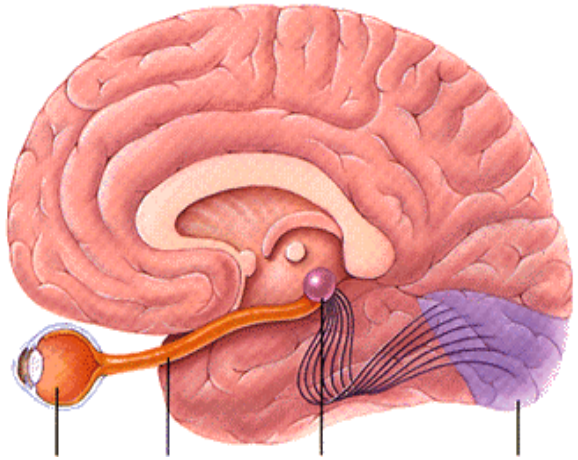
Diffuse reflection in 5 minutes

Graphics in 5 Minutes • 5.5K views • 2 years ago

Visual Perception: The Human Eye

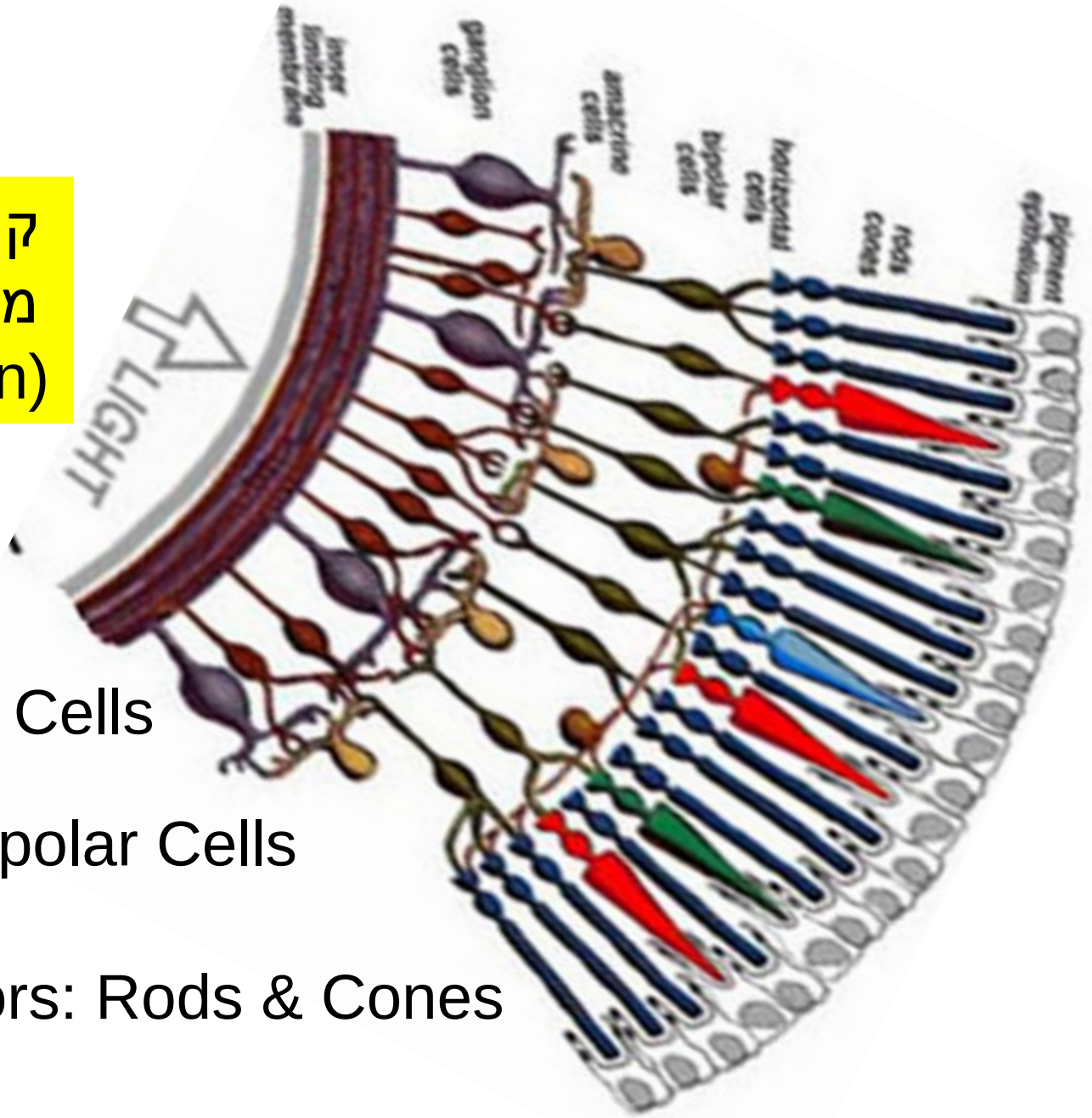






The Retina

≈ 10^8 Rods (B/W) קנים
≈ 10^7 Cones (RGB Color) מדוכים
≈ 10^4 Nerves (≈1:10^4 Reduction)



Ganglion Cells

Bipolar Cells

Receptors: Rods & Cones

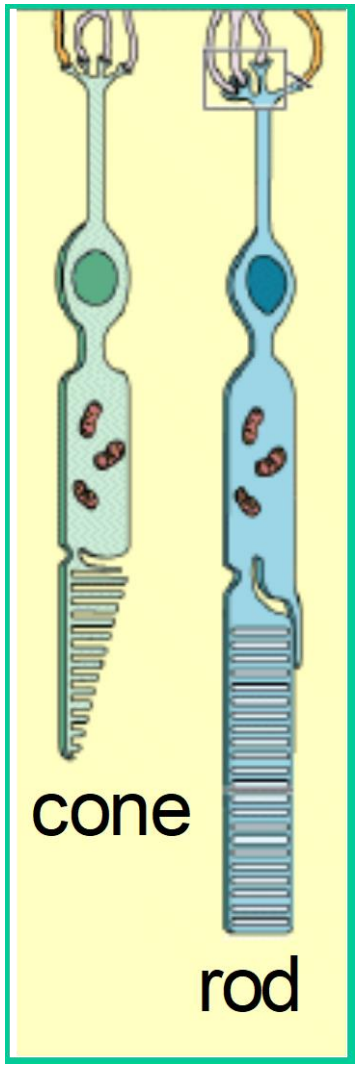
Two types of light-sensitive receptors

≈10^8 Rods - gray-scale vision

- rod-shaped
- highly sensitive
- operate at low light

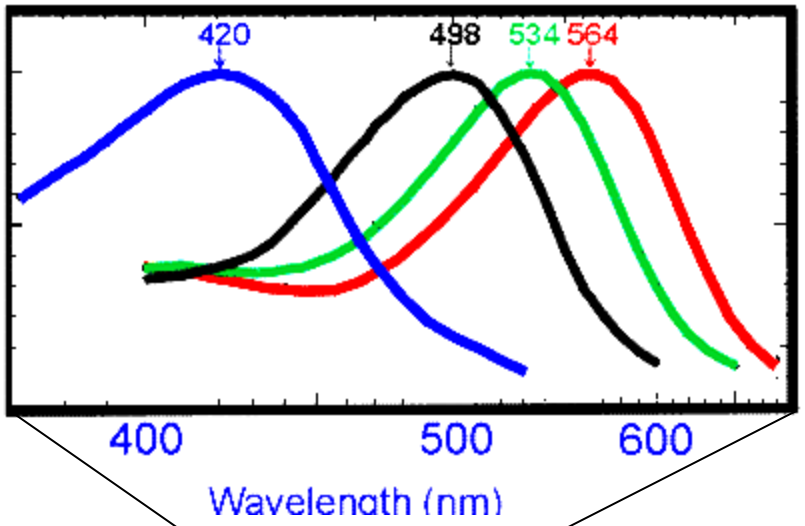
≈10^7 Cones - color vision

- cone-shaped
- less sensitive
- operate in bright light



Colors - Electromagnetic Radiation

Rods/Cones Integrate
a frequency range



- Rods
- **Blue** Cones (high Freq.)
- **Green** Cones
- **Red** Cones (low freq.)

High Frequency
Short Wavelength

Low Frequency
Long Wavelength

350-780 nm

Cosmic Rays | X Rays | UV | **Visible** | IR | Microwave | TV/Radio

- Maximum Sun Energy: 450 nm. Is it luck?
- Best Atmospheric Transmittance: Visible Range

5 Minute Clips

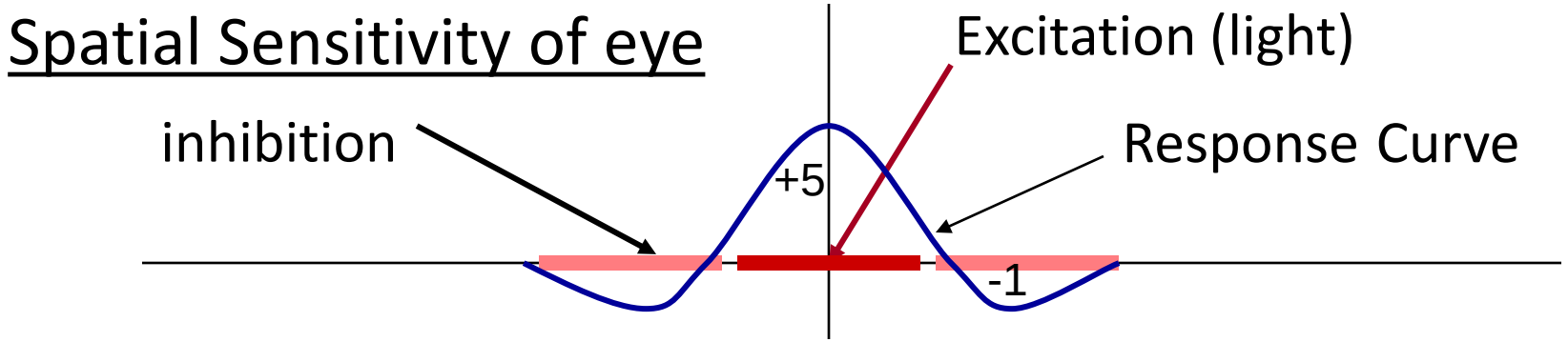
- Color Perception

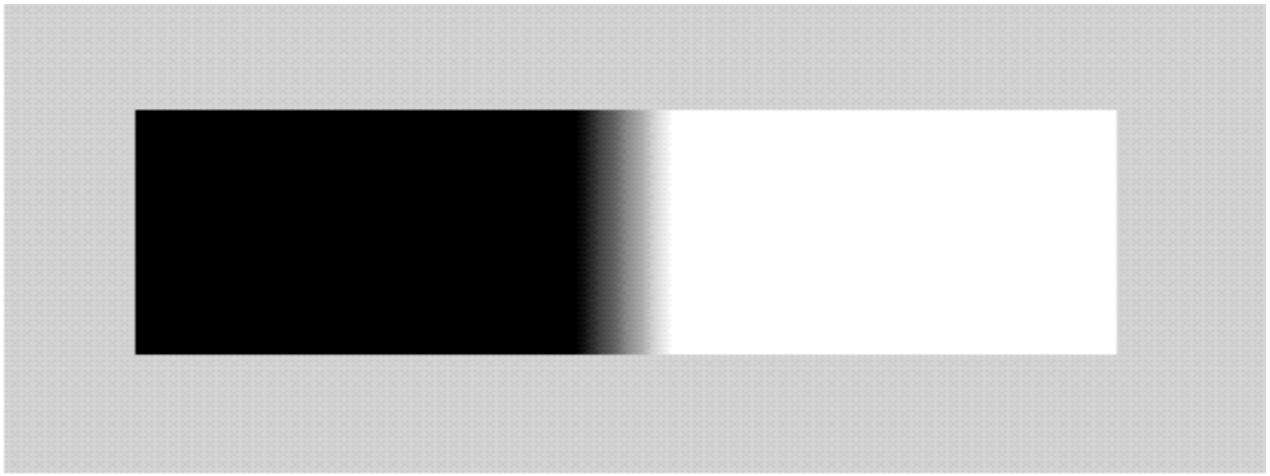
3

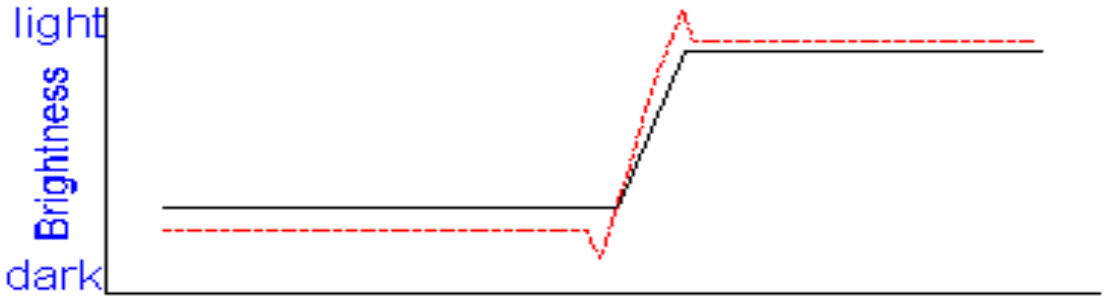
Color Perception in 5 minutes

Graphics in 5 Minutes • 1.2K views • 2 weeks ago

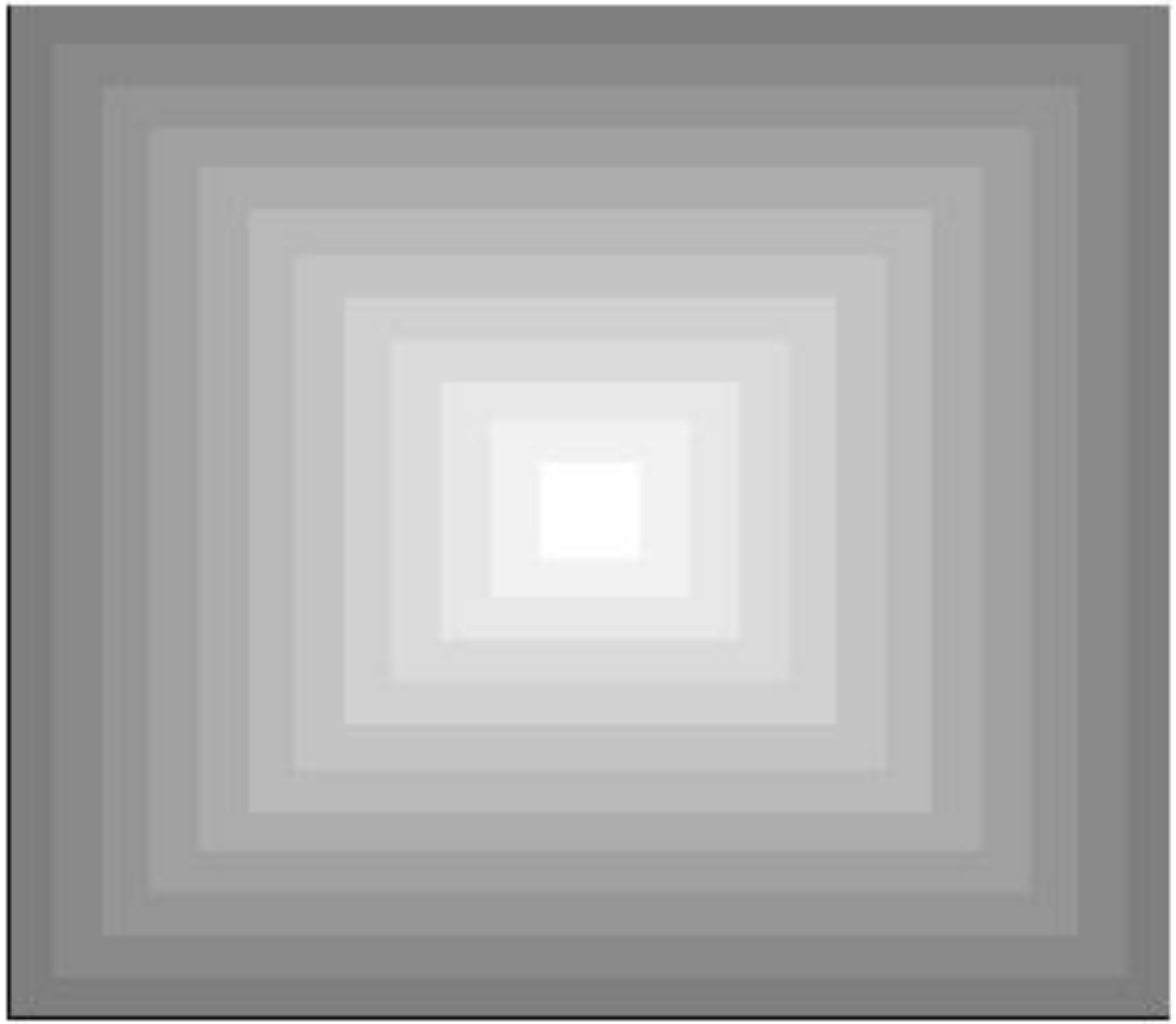
Visual Perception: Mach Bands

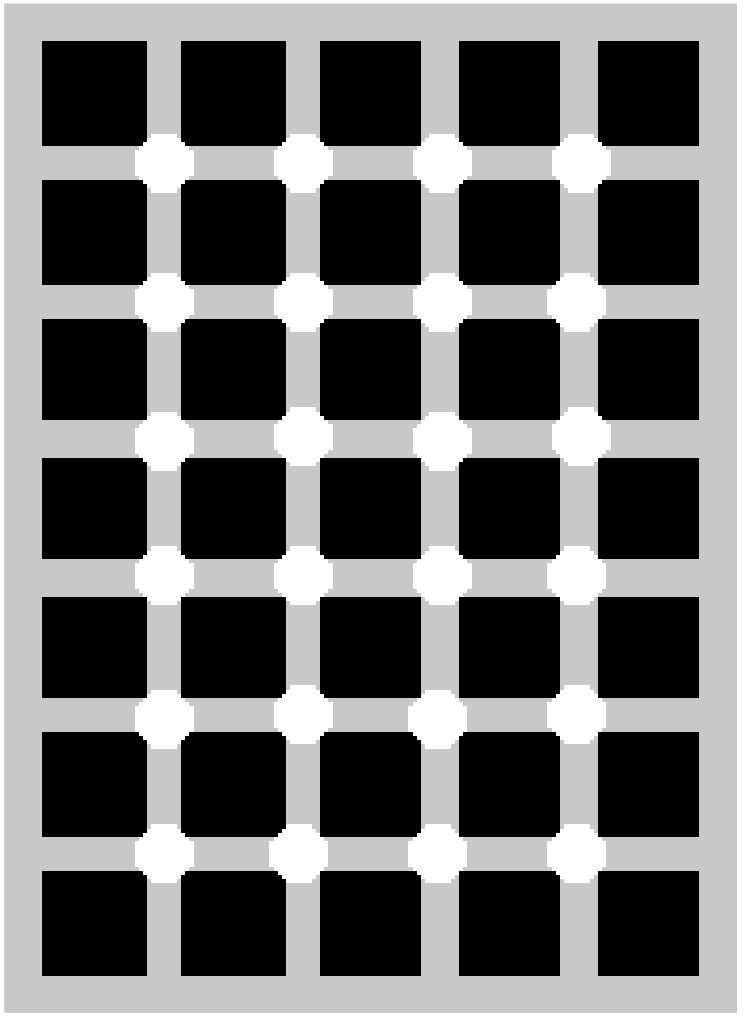




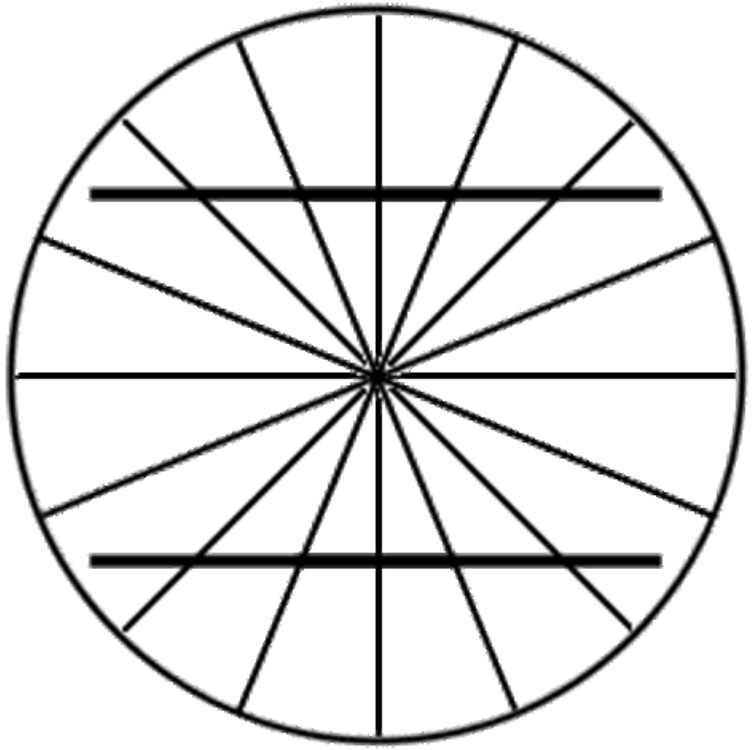


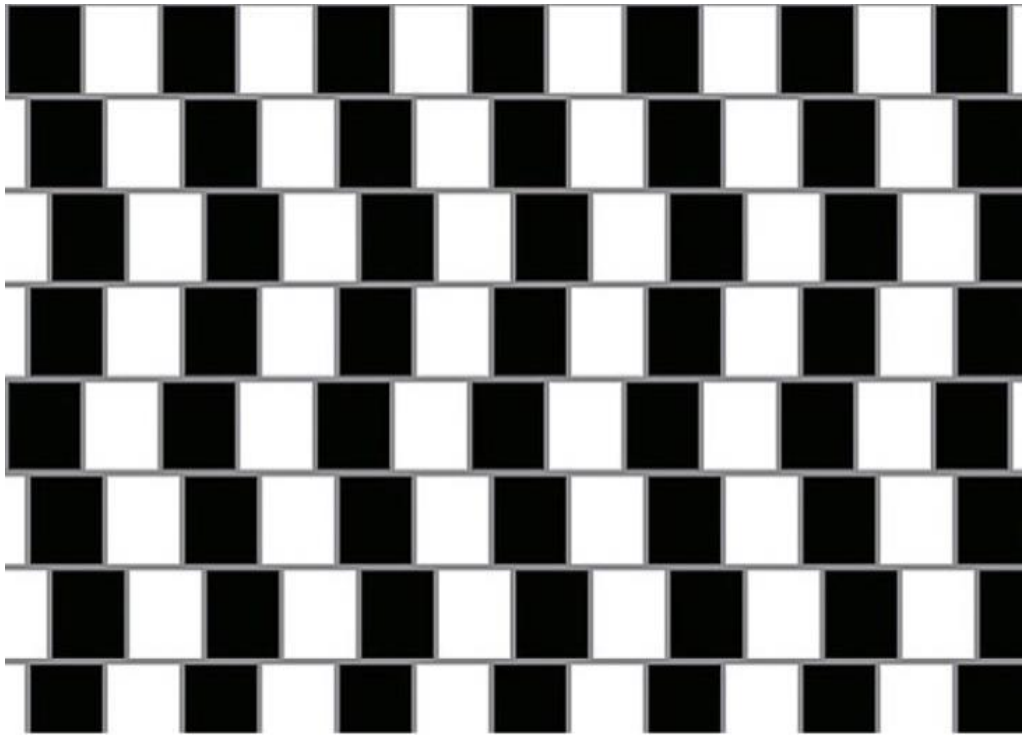
Mach Bands (2)





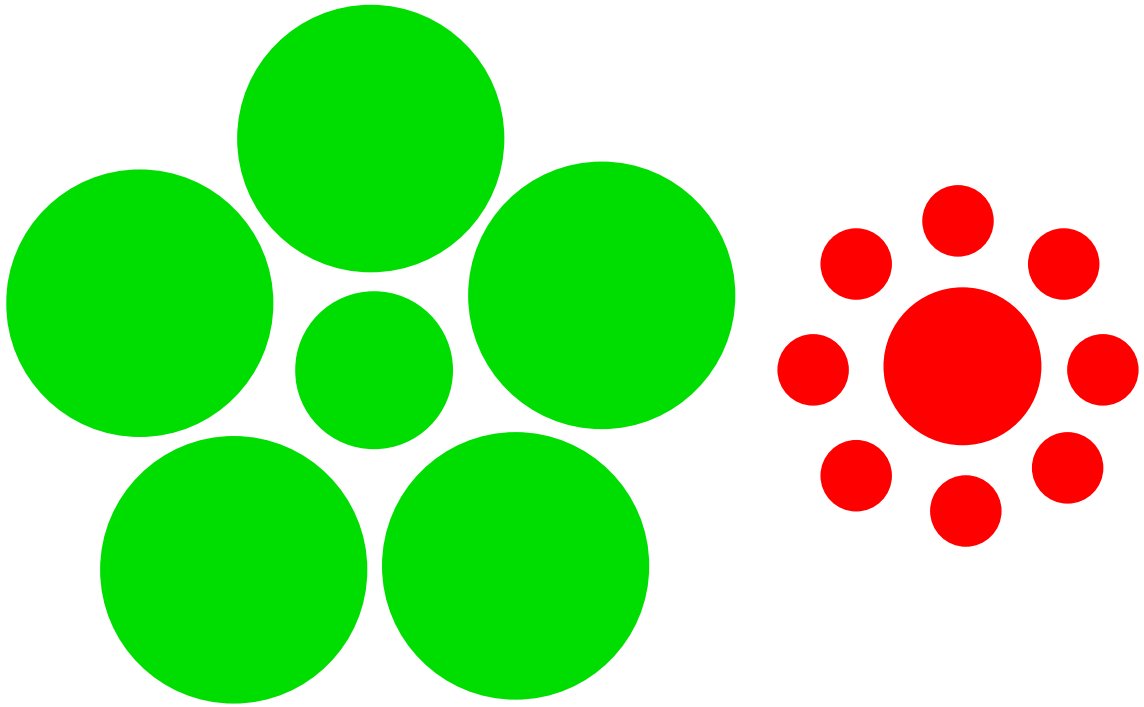
Visual Illusions:
Are the Horizontal Lines Parallel?





Visual Illusions: Size

Ebbinghaus Illusion:



Which central circle is larger?

What Do You See?

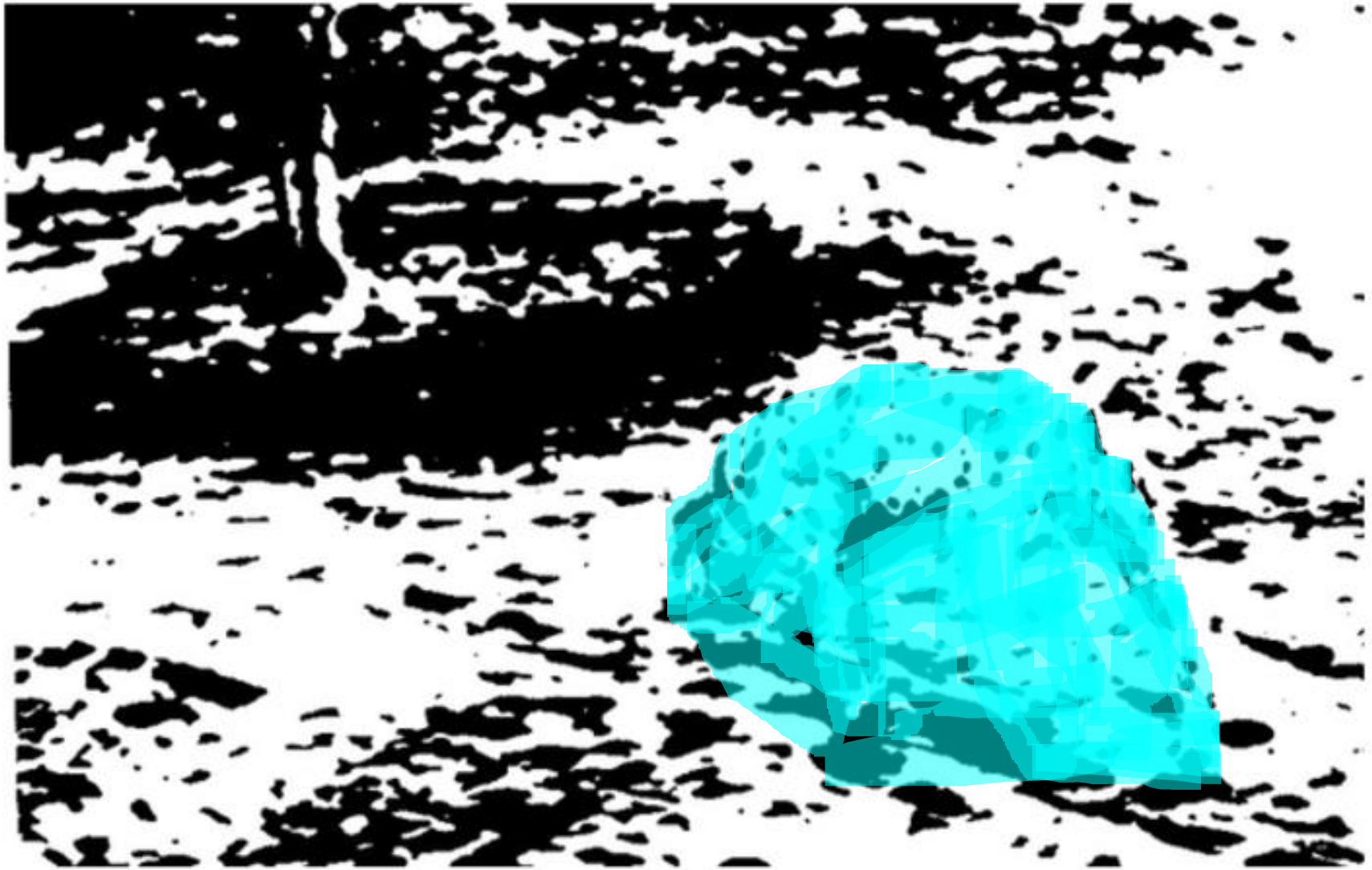
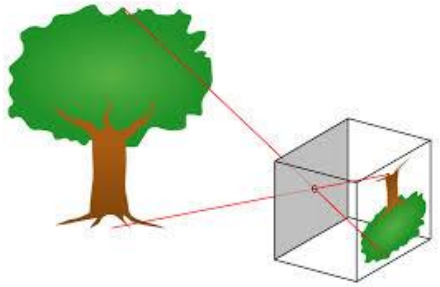


Image Digitization: 3 Stages

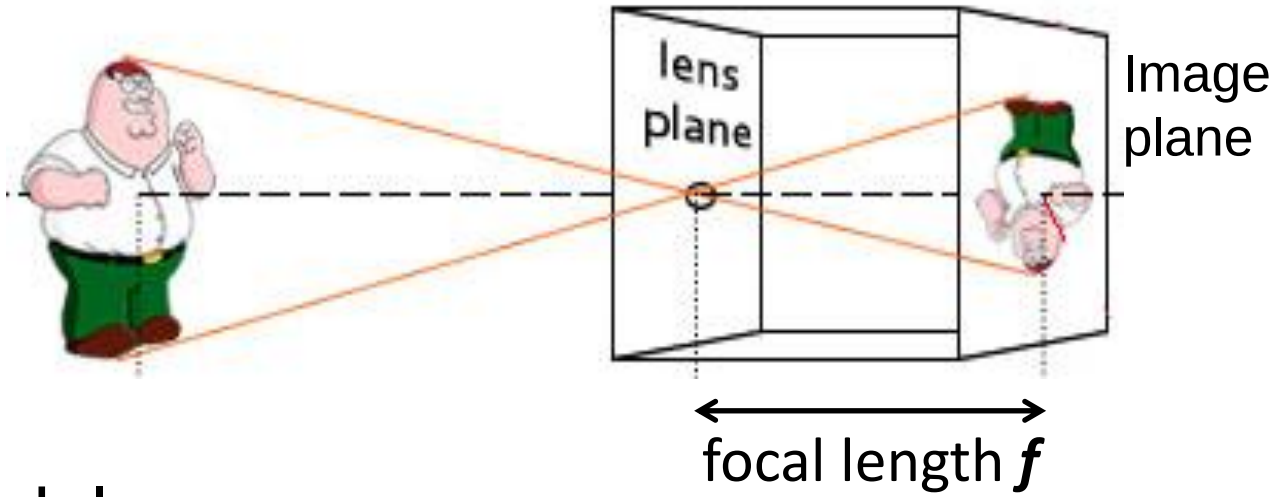
1) Transforming the **3D** world into **2D** image
   - Perspective Projection (Optics, Continuous)



2) Sampling the Image Plane
   - Finite number of **Pixels**

3) Quantizing the color/gray-level
   - Finite number of colors (e.g. 8 bits per color)

Pinhole Camera (Camera Obscura (Latin) = Dark Room)

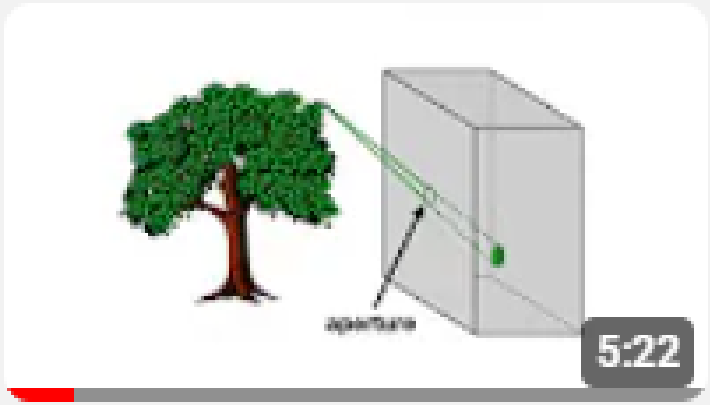


- Pinhole model:
  - Captures pencil of rays – all rays through a single point
  - The point (pinhole) is called Center of Projection (COP)
  - The image is formed on the Image Plane
  - Focal length ***f*** is distance from COP to Image Plane

5 Minute Video Clips

- 12) Perspective projection
- 13) Perspective projection: Part 2 – the math!

12



Perspective projection in 5 minutes

Graphics in 5 Minutes • 24K views • 2 years ago

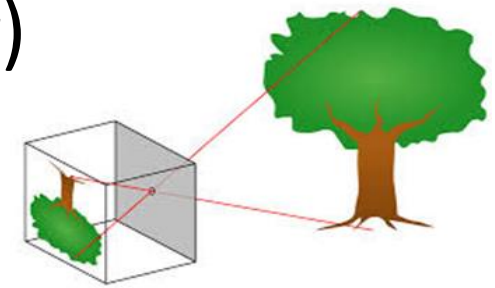
13

Perspective projection in 5 minutes: Part 2 -- the math!

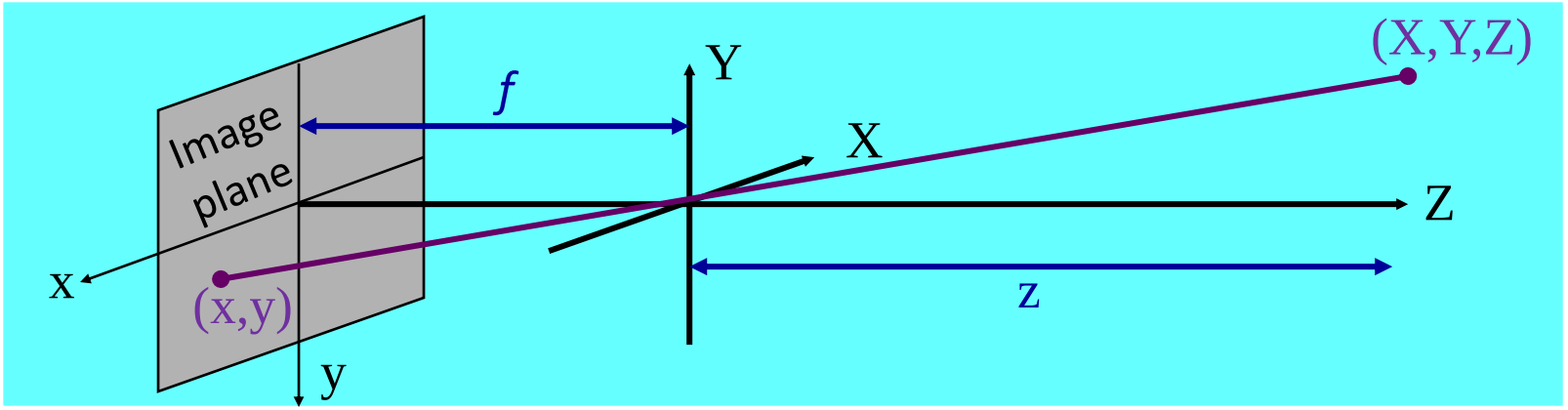
Graphics in 5 Minutes • 8.3K views • 2 years ago

Perspective Projection

- Transforming the 3D world (X, Y, Z) into 2D image (x, y)
  - Continuous Perspective Projection (optics)
  - All rays pass through one point (f = focal length)



Simple case: Aligned World axis (X, Y, Z) and Image axis (x, y)

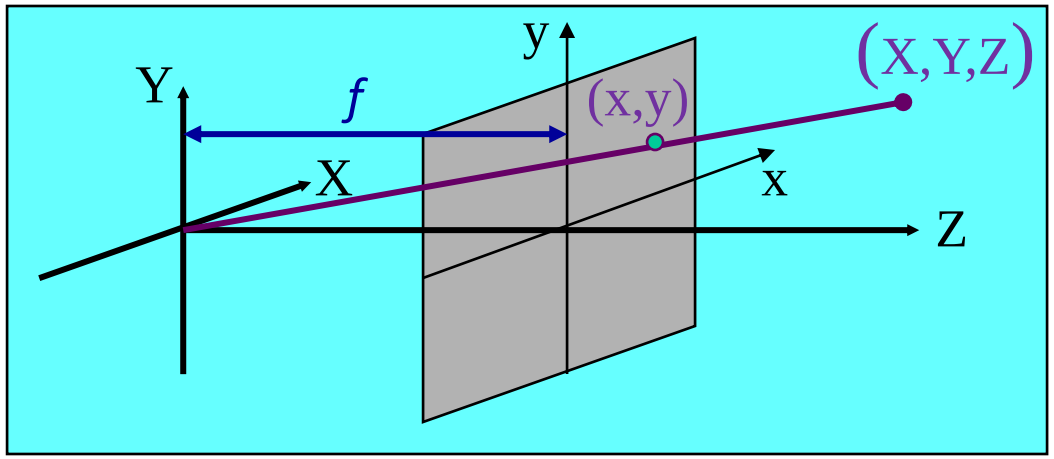


$$\frac{Y}{Z} = \frac{y}{f}$$

Similar Triangles

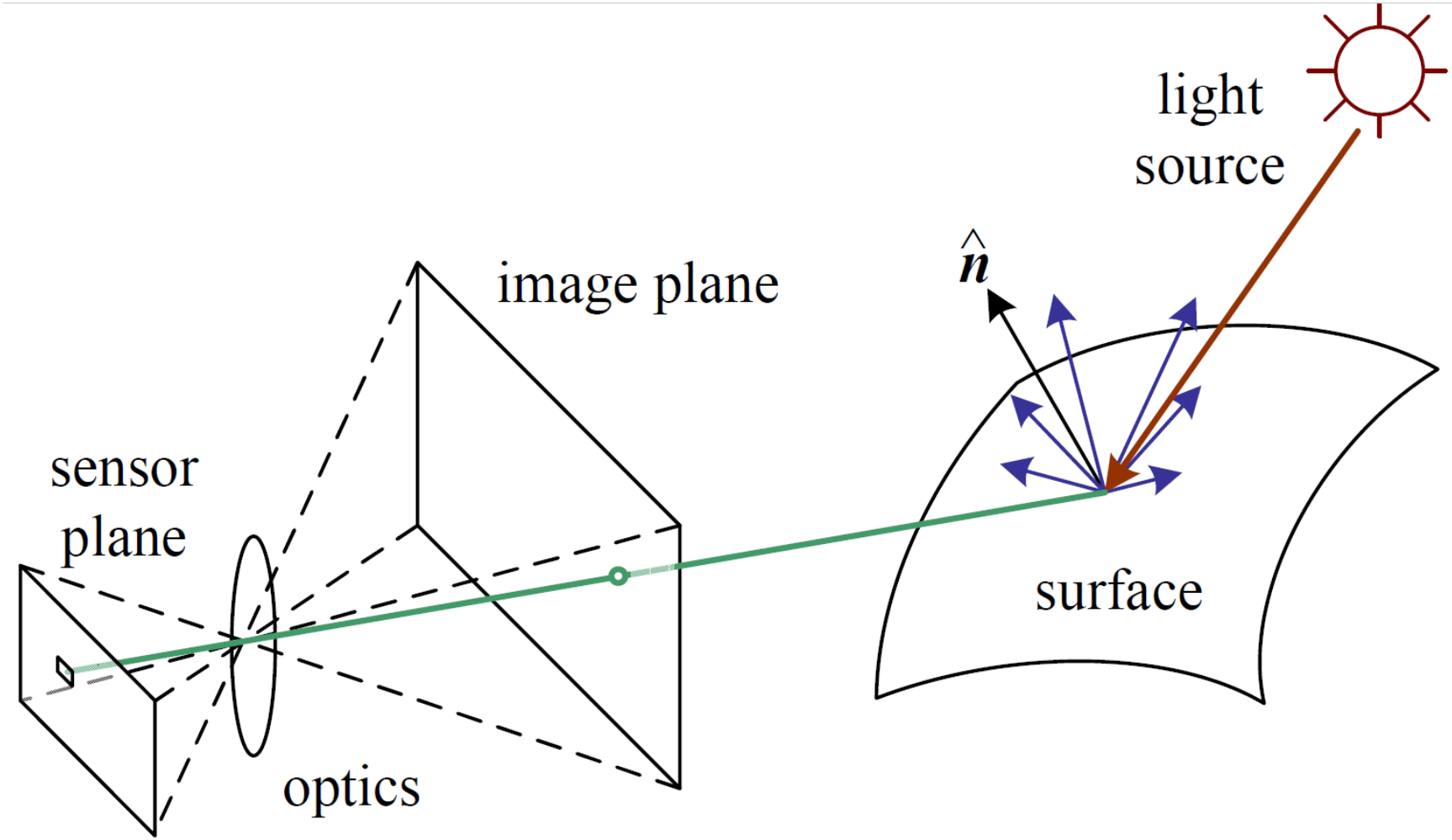
$$x = \frac{f}{Z} X$$

$$y = \frac{f}{Z} Y$$



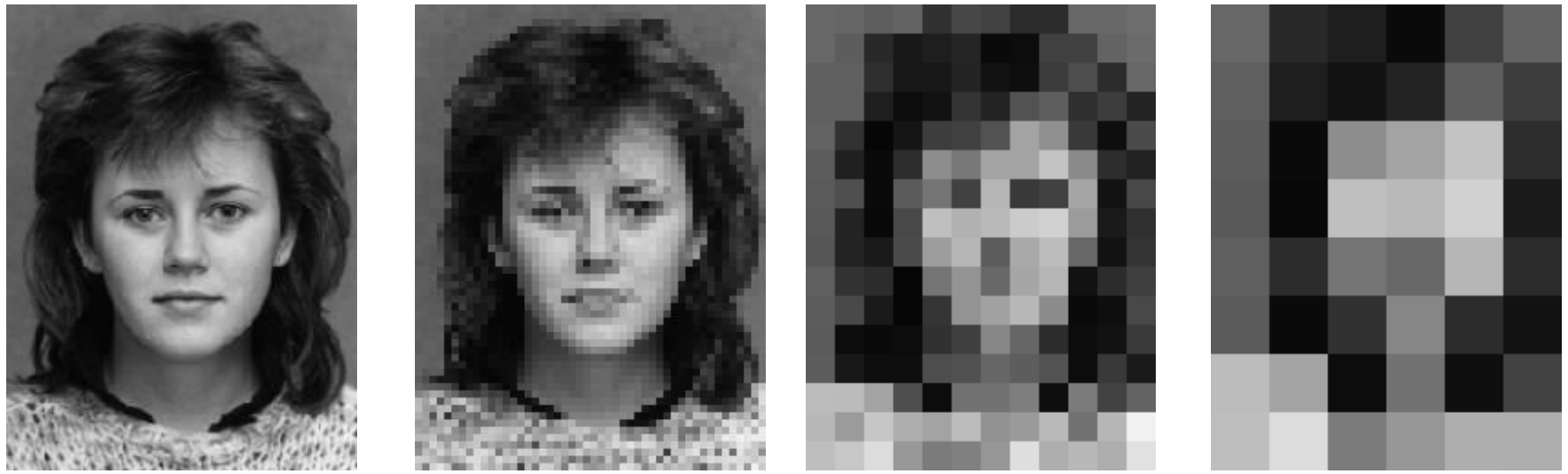
Summary: First Stages of Image Acquisition

Analog Light



Spatial Sampling to Pixels

- Sampling the Image Plane
  - Finite number of **Pixels**
  - Do we always want maximum number of pixels?

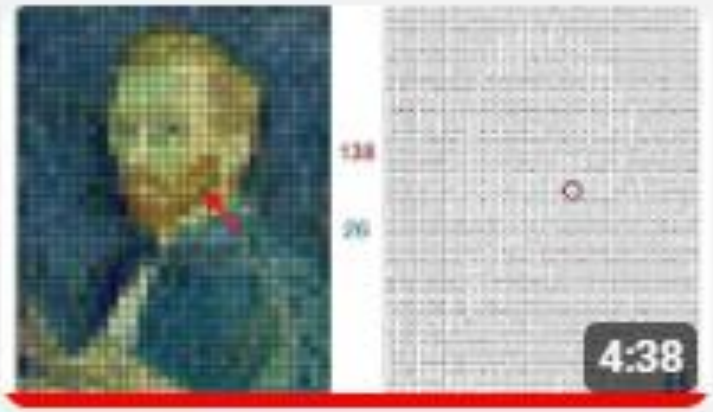


256 lines | 64 lines | 16 lines | 8 lines

5 Minute Video Clips

- 4) Images in 5 minutes

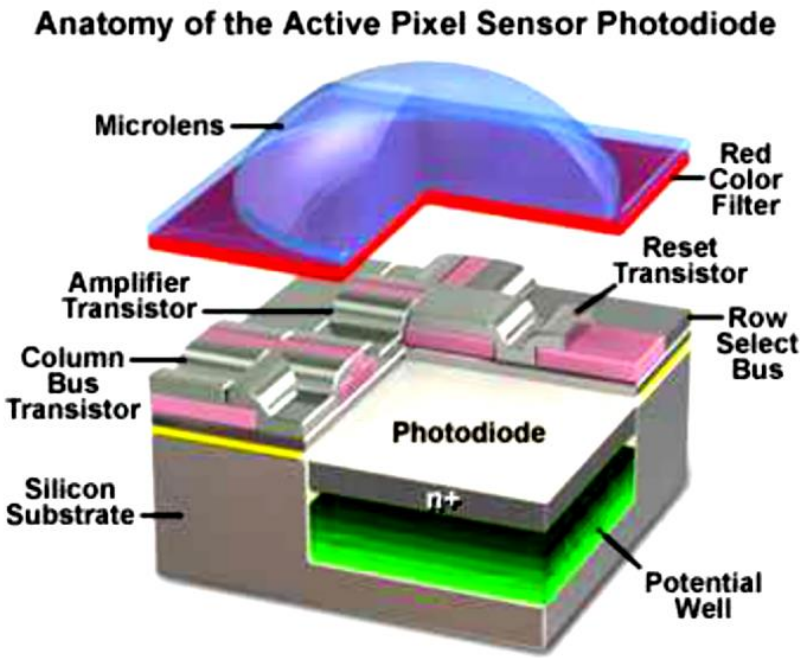
4

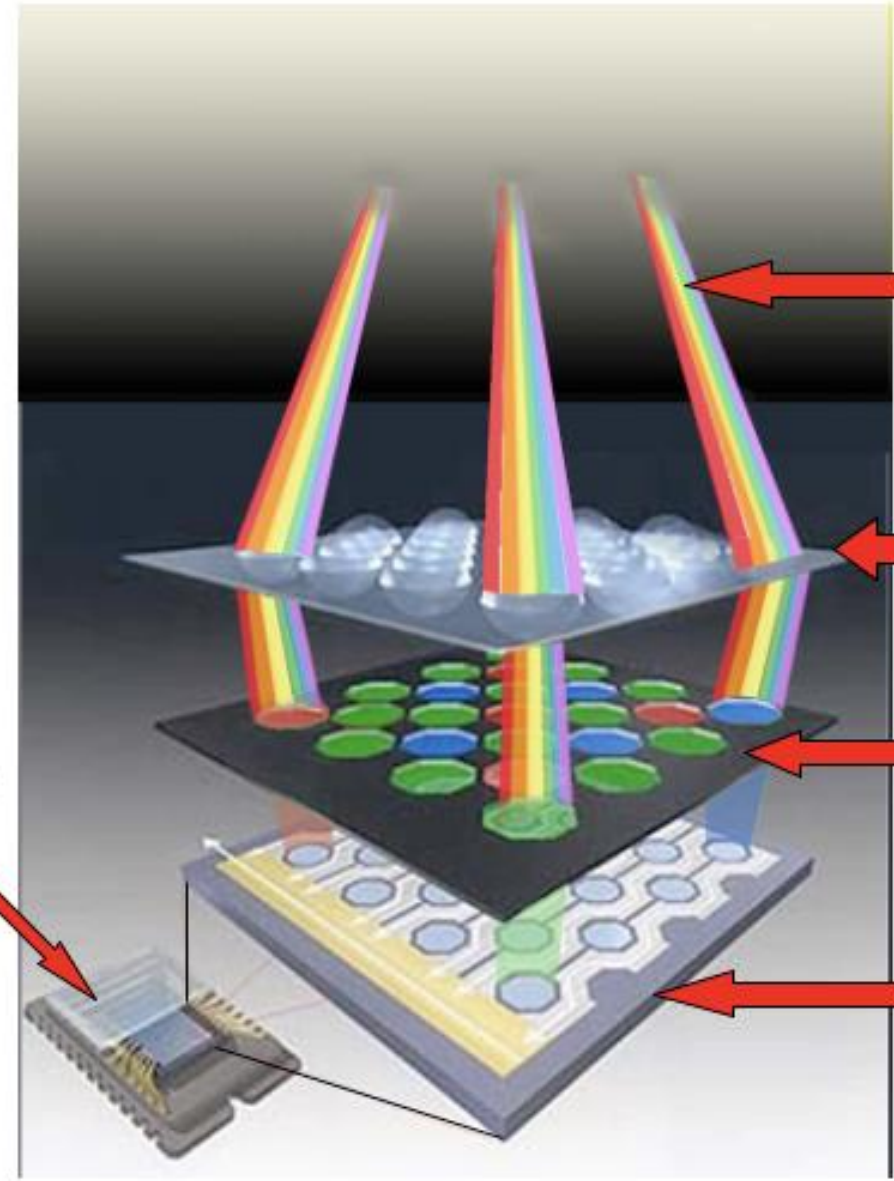


Images in 5 minutes: The Case of the Splotched Van Gogh, Part 1

Graphics in 5 Minutes • 4.4K views • 2 years ago

RGB Inside the Camera





Millions of light sensors

Incoming Visible light

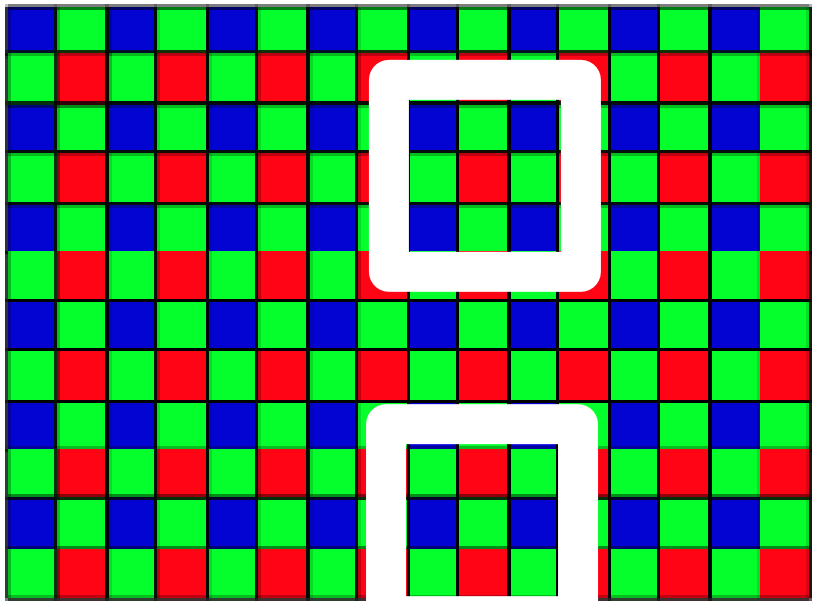
Visible Light passes through IR-Blocking Filter

Color Filters control the color light reaching a sensor

Color blind sensors convert light reaching each sensor into electricity

Bayer Filter

- In 1975, Bruce Bayer invents the color filter array, used in most digital camera.
- ¼ pixels detect red; ¼ pixels blue; ½ pixels green;
- The camera invents 2 missing colors in pixels. How?



- Demosaicing: Invents missing colors
- Many methods, mostly proprietary
- A possible (bad) method:
  - Average 2 or 4 neighbors

Color/Gray-level Quantization

- Quantizing the color/gray-level
  - Finite number of colors

256 Levels 8 Levels 4 Levels 2 Levels

Digital Pictures

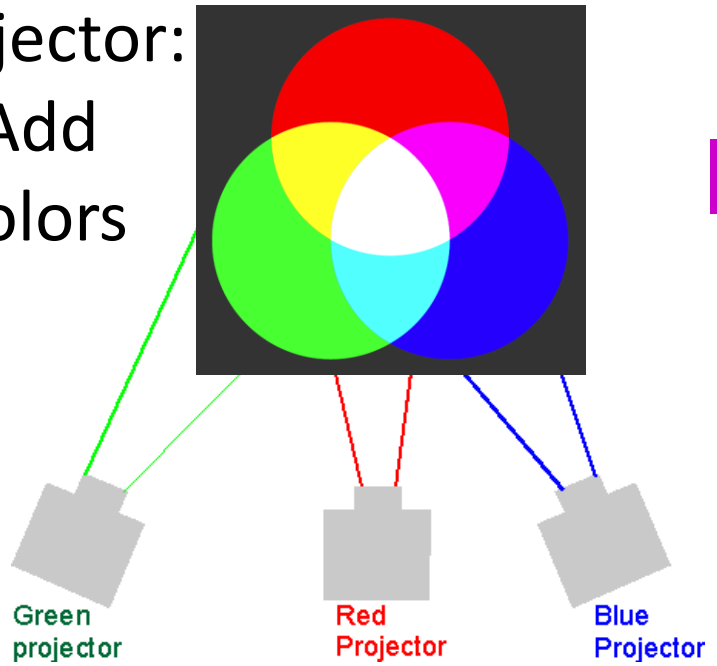
- A Matrix of numbers (Greylevel image)
- A Matrix of triplets (RGB Color, etc.)

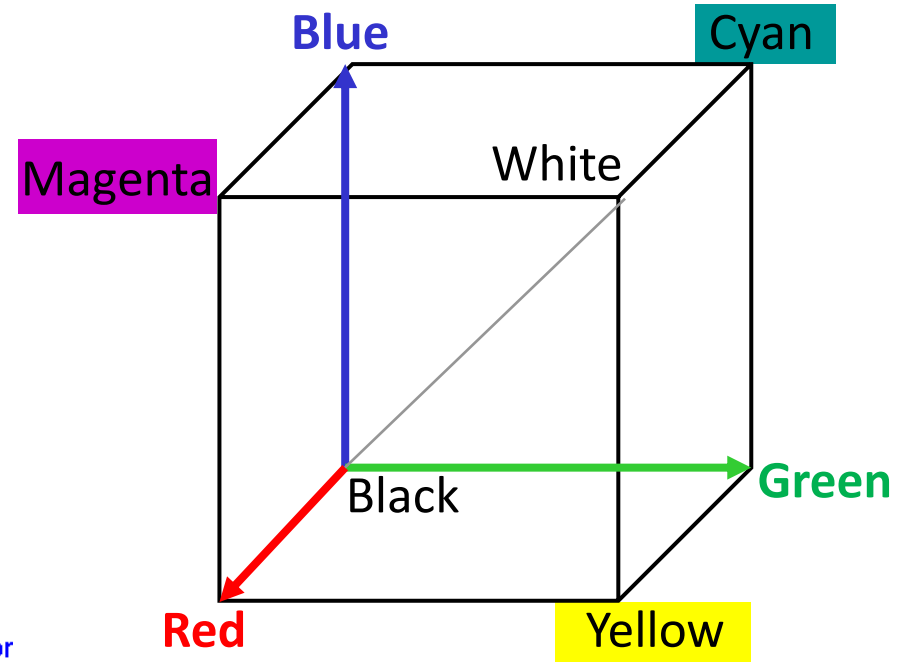
2	4	5	6	7	8	9	10	11	12	11	10	9	8	7
3	5	6	7	8	9	10	11	12	13	12	11	10	9	8
4	6	7	8	9	10	11	12	13	14	13	12	11	10	9
5	7	8	9	10	11	12	13	14	15	14	13	12	11	10
6	8	9	10	11	12	13	14	15	16	15	14	13	12	11
7	9	10	11	12	13	14	15	16	17	16	15	14	13	12
8	10	11	12	13	14	15	16	17	18	17	16	15	14	13
9	11	12	13	14	15	16	17	18	19	18	17	16	15	14
10	12	13	14	15	16	17	18	19	20	19	18	17	16	15
9	11	12	13	14	15	16	17	18	19	18	17	16	15	14
8	10	11	12	13	14	15	16	17	18	17	16	15	14	13
7	9	10	11	12	13	14	15	16	17	16	15	14	13	12
6	8	9	10	11	12	13	14	15	16	15	14	13	12	11
5	7	8	9	10	11	12	13	14	15	14	13	12	11	10
4	6	7	8	9	10	11	12	13	14	13	12	11	10	9
3	5	6	7	8	9	10	11	12	13	12	11	10	9	8
2	4	5	6	7	8	9	10	11	12	11	10	9	8	7
1	3	4	5	6	7	8	9	10	11	10	9	8	7	6

Color Spaces

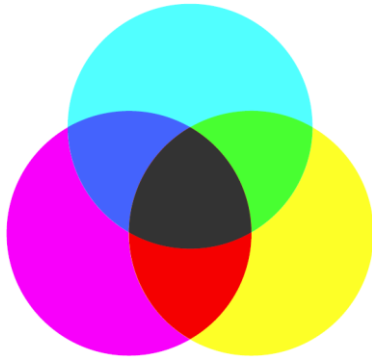
RGB (Camera, Projector - Add), CMYK (Print -Subtract), YIQ (TV)

Projector: Add colors





Ink: Absorb colors



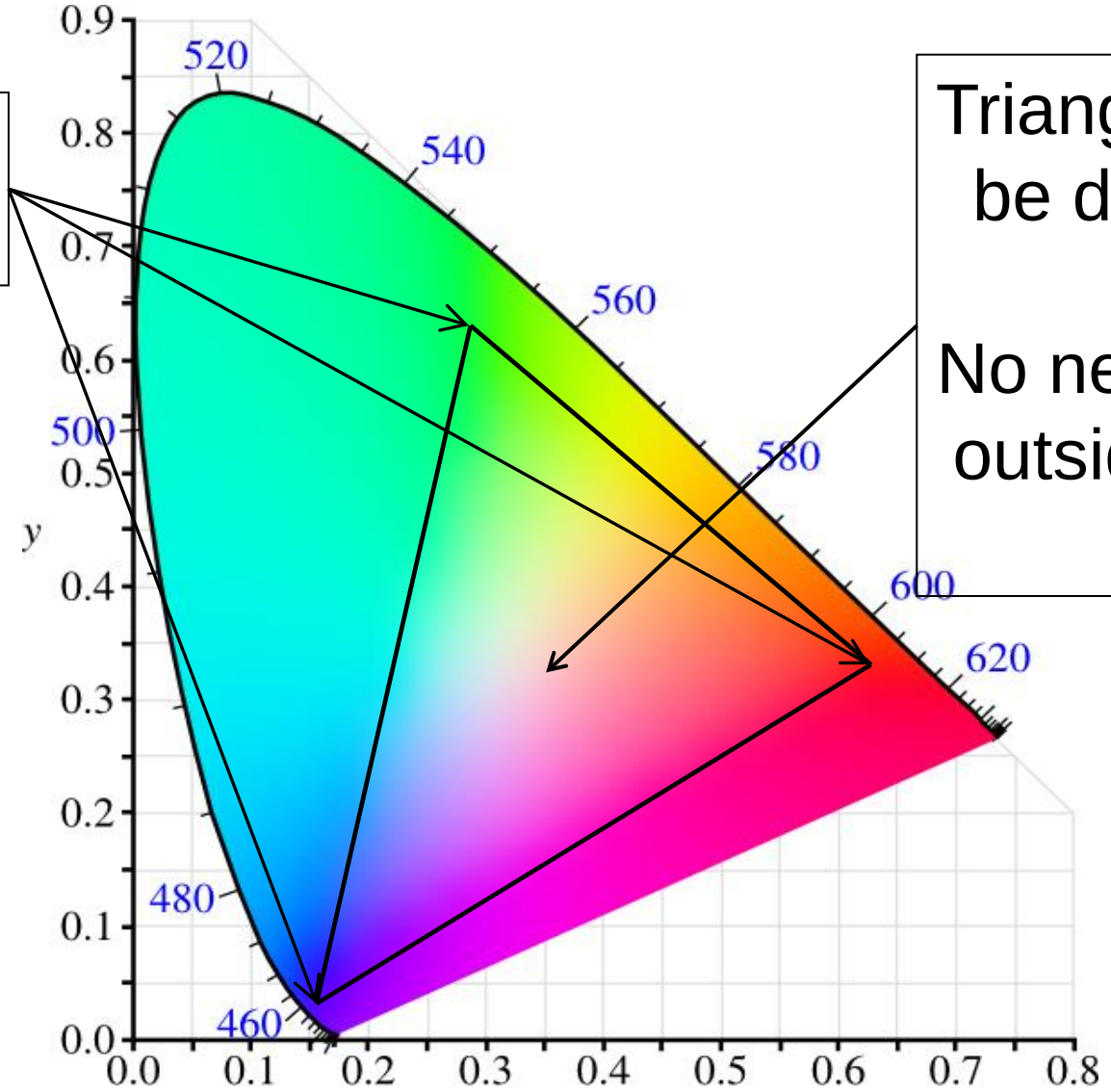
For Color to B/W TV

Y - Luminance

$$\begin{vmatrix} Y \\ I \\ Q \end{vmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

CIE Chromaticity Diagram (1931)
Boundary: Spectral Colors (Single Wavelength)

Colors of LED TV



Triangle: Colors that can be displayed (Gamut)

No new colors displayed outside Gamut (Display Limitation)