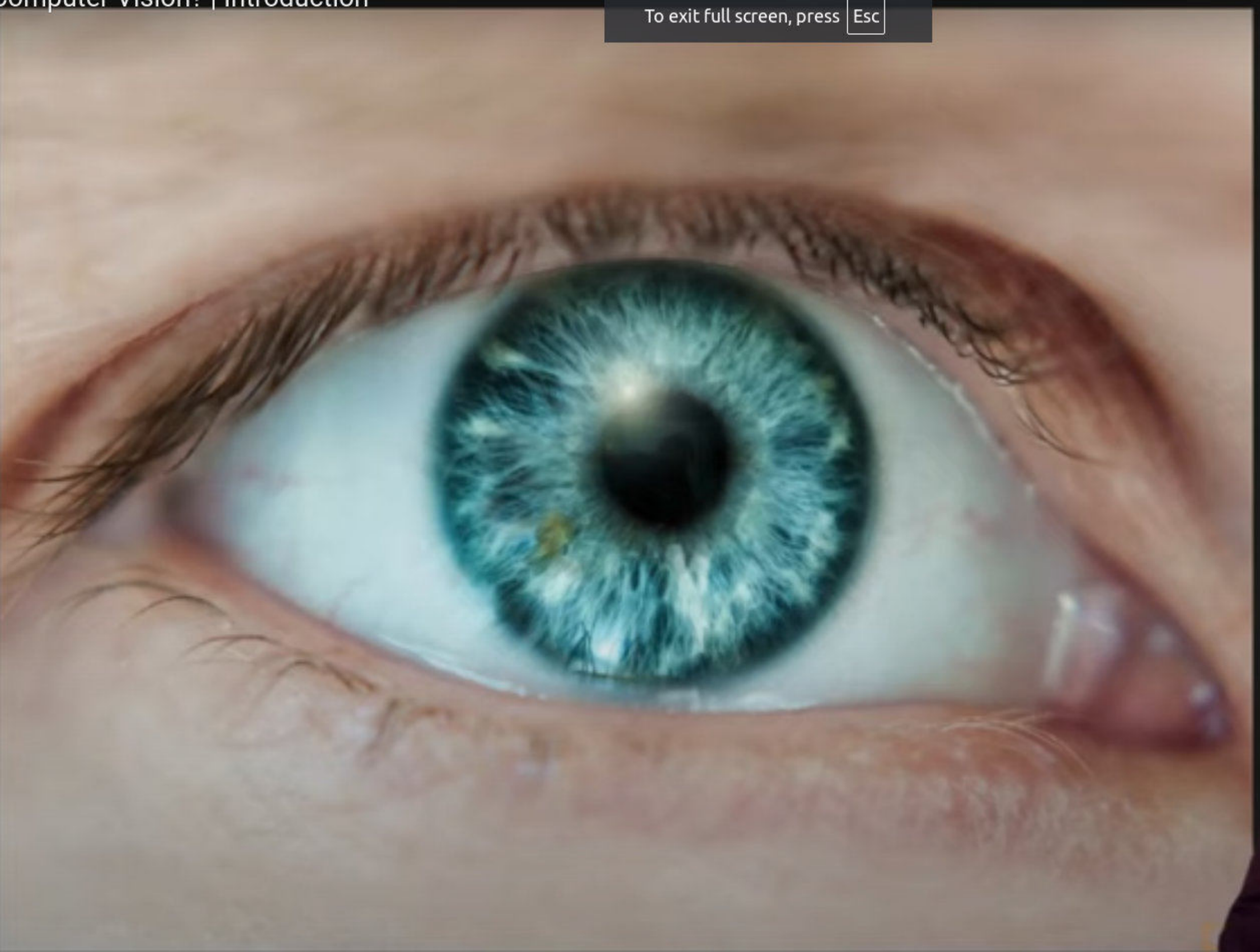
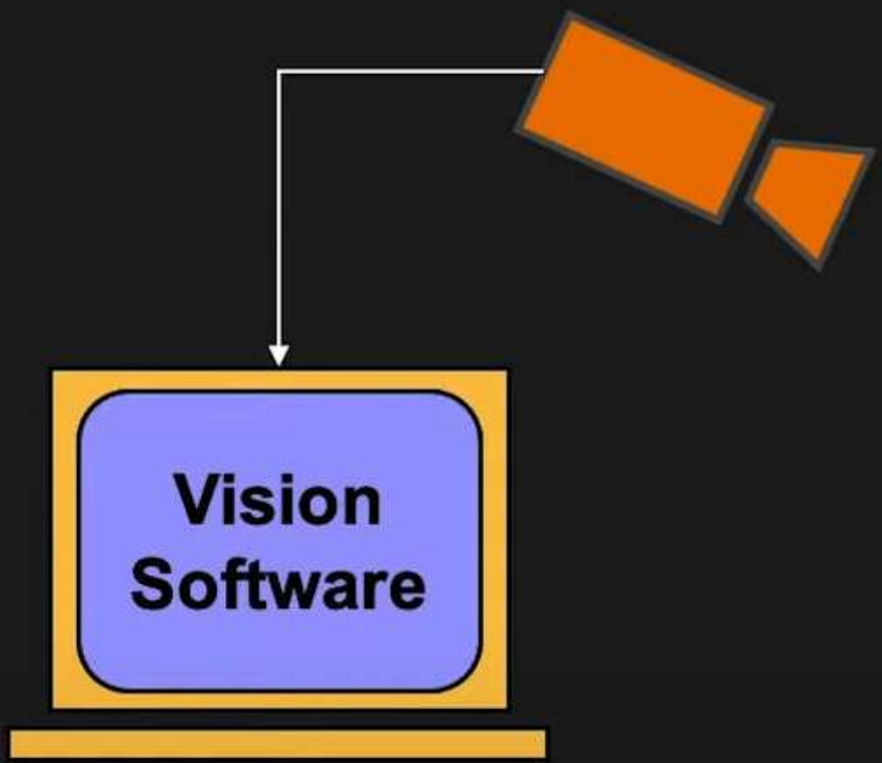

To exit full screen, press Esc

What is Computer Vision?

Camera



Scene Description

Lighting

Scene

But, What Really is Computer Vision?

Vision is

... automating human visual processes

... an information processing task

... inverting image formation

... inverse graphics

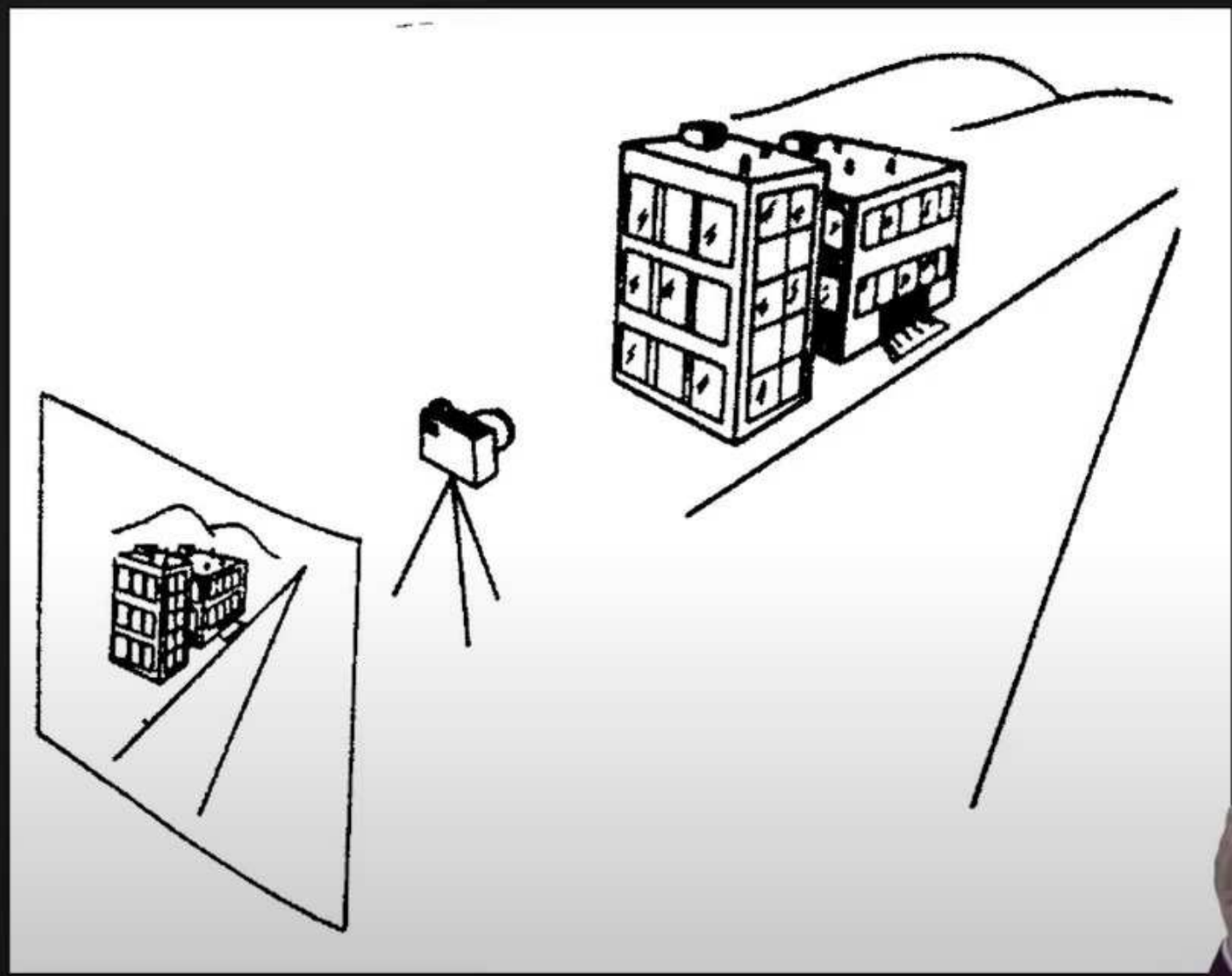
... really useful!

Vision Deals with Images

An Image is an Array of Pixels

A Pixel has Values:

- Brightness
- Color

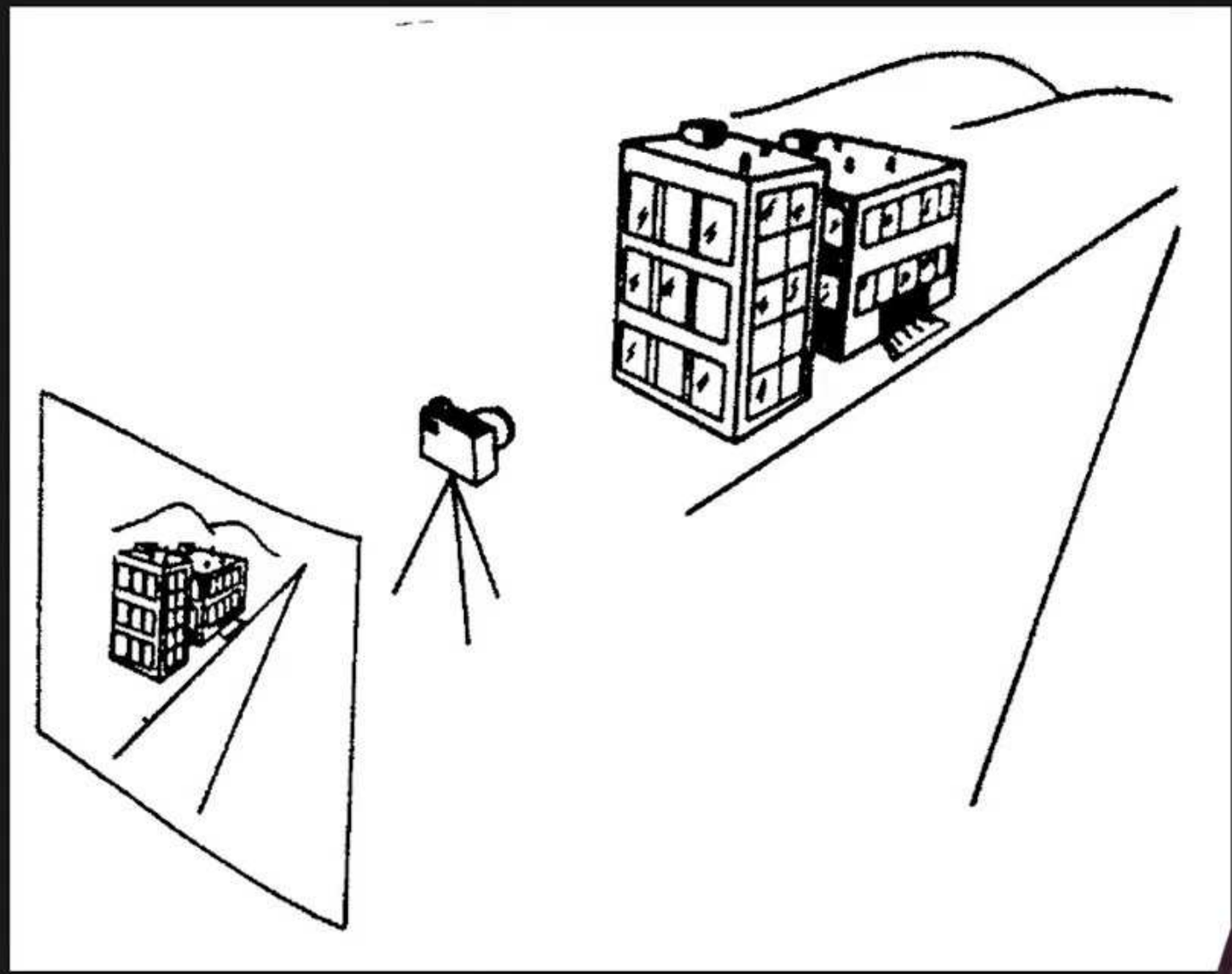


Vision Deals with Images

An Image is an Array of Pixels

A Pixel has Values:

- Brightness
- Color
- Distance
- Material
- ...



Images Are Interesting

But When You Look Close...

157	159	159	104	104	115	128	131	133	133	132	131	132	130	129	118	132	158	156	153	190	144	117	126	120	81
159	165	153	101	103	113	126	129	130	130	126	124	127	128	127	120	122	158	159	154	160	190	121	118	67	47
162	154	154	98	101	114	124	127	130	132	144	159	155	132	123	119	119	148	154	150	140	185	161	60	48	45
141	132	158	93	98	110	121	125	122	129	143	172	191	188	143	105	117	148	140	145	142	153	105	44	49	71
100	130	157	93	99	110	120	116	116	129	138	163	191	205	211	130	107	153	98	133	147	107	44	47	81	151
87	130	157	92	97	109	124	111	123	134	139	175	194	201	207	205	126	151	74	114	160	57	49	63	141	163
93	131	159	92	98	112	132	108	123	133	162	180	183	192	196	205	184	151	138	199	195	54	47	119	161	156
96	134	164	95	97	113	147	108	125	142	156	171	173	178	184	181	186	191	206	203	161	44	84	158	159	155
95	137	165	95	95	111	168	122	130	137	145	139	144	139	145	179	193	203	194	158	95	49	135	160	157	155
101	139	166	94	96	104	172	130	126	130	108	77	85	80	153	191	188	161	144	113	48	83	161	160	156	153
101	133	167	94	96	100	154	137	123	92	67	57	72	153	182	184	175	101	116	53	48	119	166	163	159	152
99	130	169	97	99	109	131	128	84	55	60	75	149	176	170	194	209	99	79	51	67	150	158	155	154	151
97	129	170	97	98	118	122	94	66	56	56	140	161	114	136	187	163	81	85	52	98	161	159	154	148	137
92	123	173	101	98	129	95	74	74	45	94	174	106	115	126	168	108	60	92	55	128	157	153	148	145	157
81	115	175	104	116	87	78	69	84	56	140	124	158	170	143	173	150	76	90	68	148	153	146	148	186	196
69	108	172	107	103	87	82	54	83	105	93	107	153	166	132	162	153	68	87	97	157	149	141	179	204	206
71	119	172	106	91	78	97	70	99	104	59	116	142	153	141	165	123	55	84	132	154	146	148	199	209	210
61	126	175	112	83	74	92	123	130	53	61	108	137	132	138	154	77	58	82	150	152	143	155	210	211	213
53	128	175	105	71	82	109	127	75	50	57	74	115	139	151	117	47	67	89	154	154	143	159	218	214	199
56	115	173	105	61	76	106	114	70	54	52	60	102	137	160	146	78	67	96	135	130	125	165	215	142	81
117	106	176	101	55	71	81	112	101	57	55	70	117	139	152	188	198	112	87	146	131	112	178	164	81	91
107	121	177	89	50	64	60	103	114	66	56	90	120	140	149	169	201	194	100	148	134	155	208	120	99	99

Vision Research

- Vision is a Hard Problem
- Vision is Multi-Disciplinary
- Considerable Progress Has Been Made
- Many Successful Real-World Applications

What is Vision Used For?

Factory Automation: Vision-Guided Robotics

What is Vision Used For?

Factory Automation: Visual Inspection

What is Vision Used For?

ATA 010

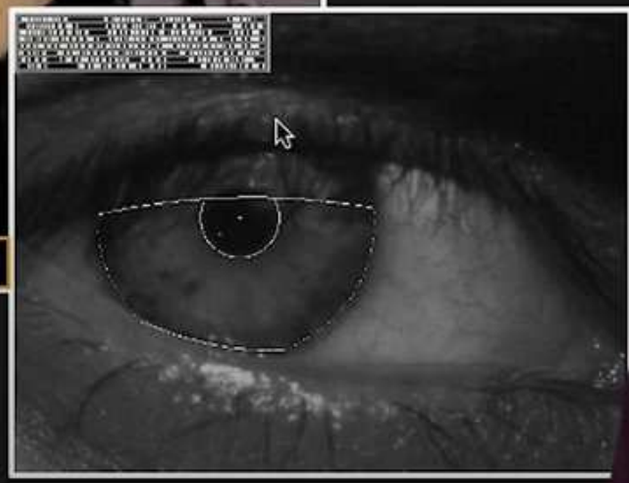
Optical Character Recognition (OCR): Reading License Plates

What is Vision Used For?

I.2

Optical Character Recognition (OCR): Book Digitization

What is Vision Used For?



Biometrics: Iris Recognition

What is Vision Used For?

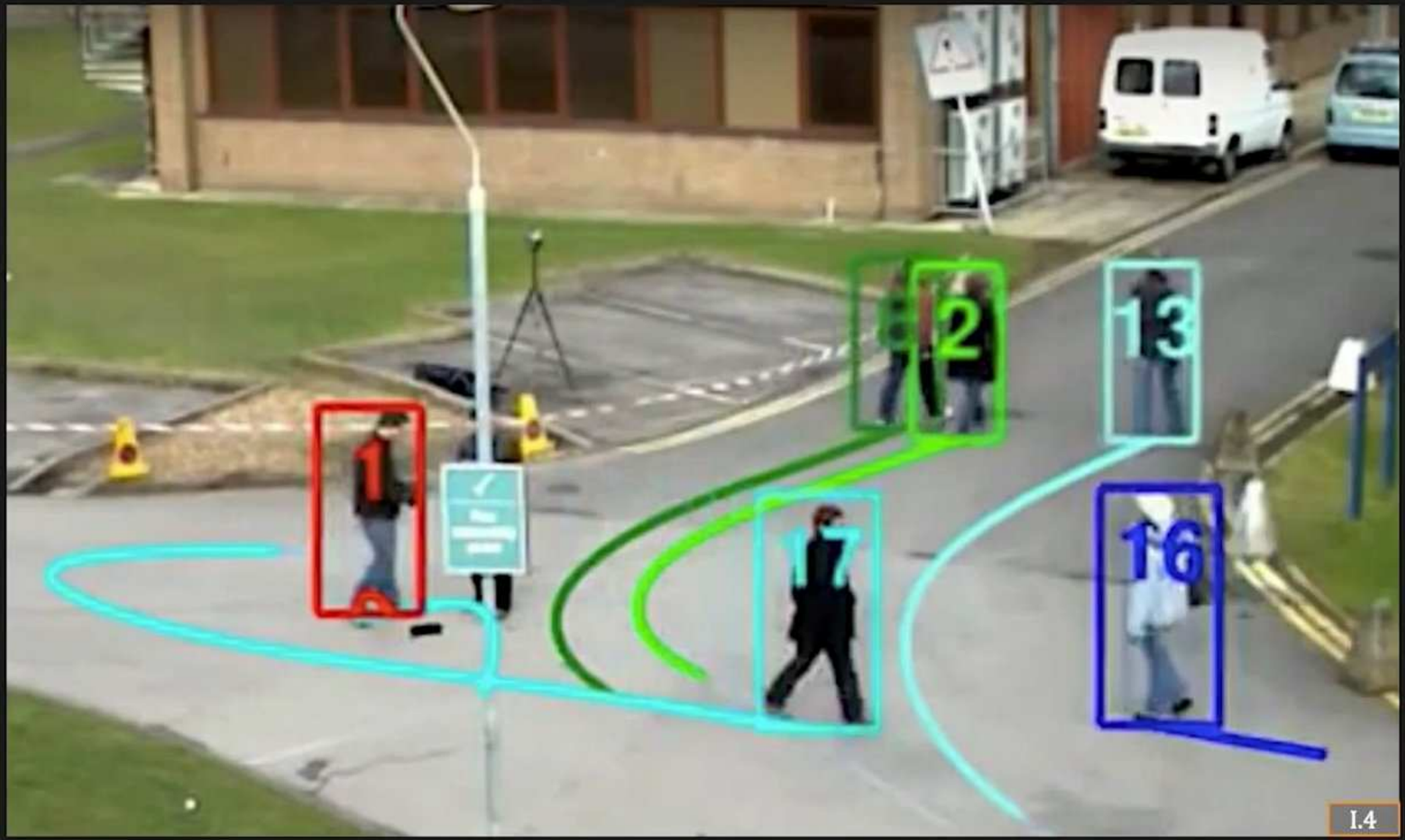
Biometrics: Face Detection and Recognition

What is Vision Used For?

I.5

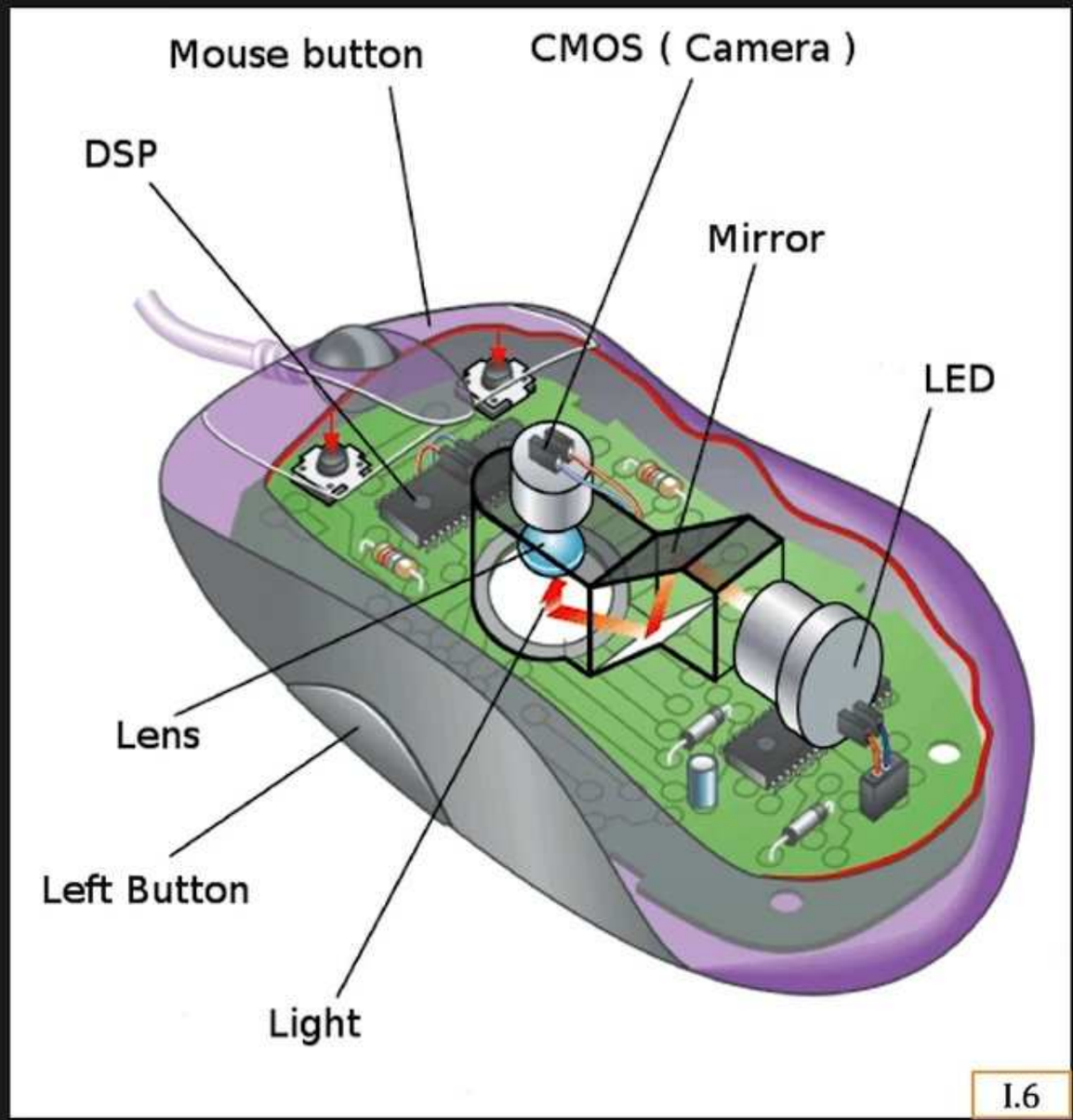
Intelligent Marketing: Vending Machine with Face Detection

What is Vision Used For?



Security: Object Detection and Tracking

What is Vision Used For?



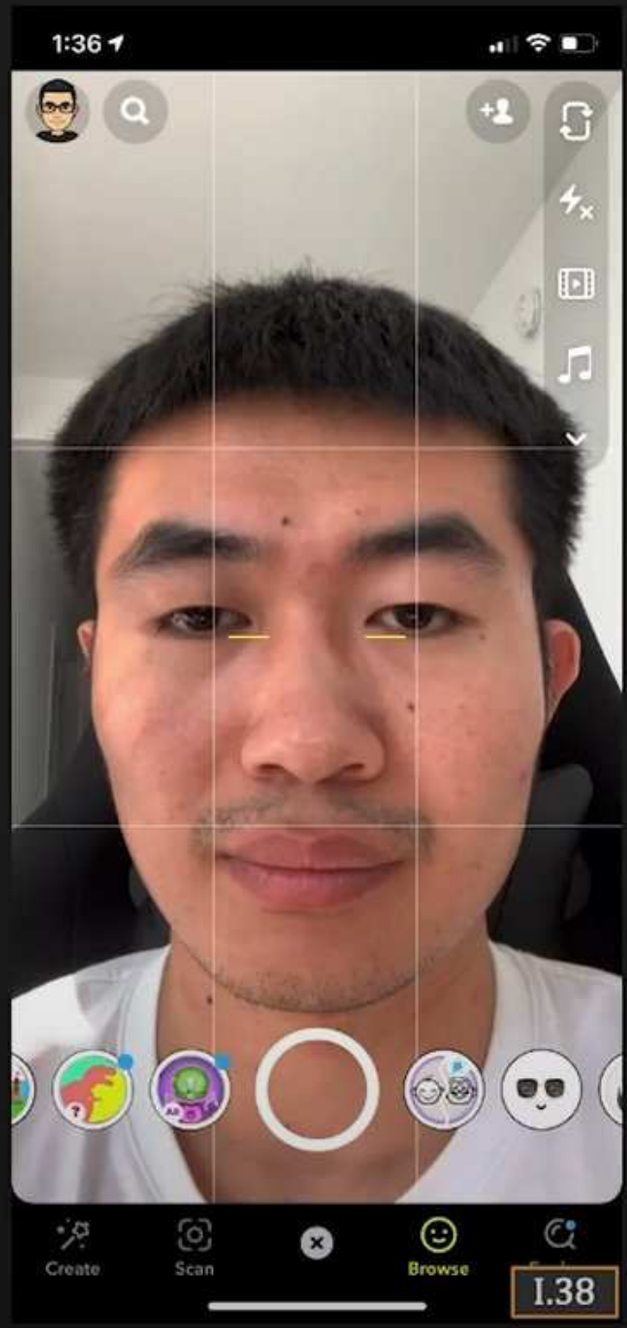
Human Computer Interaction: Optical Mouse

What is Vision Used For?

Doug

Elbor

What is Vision Used For?

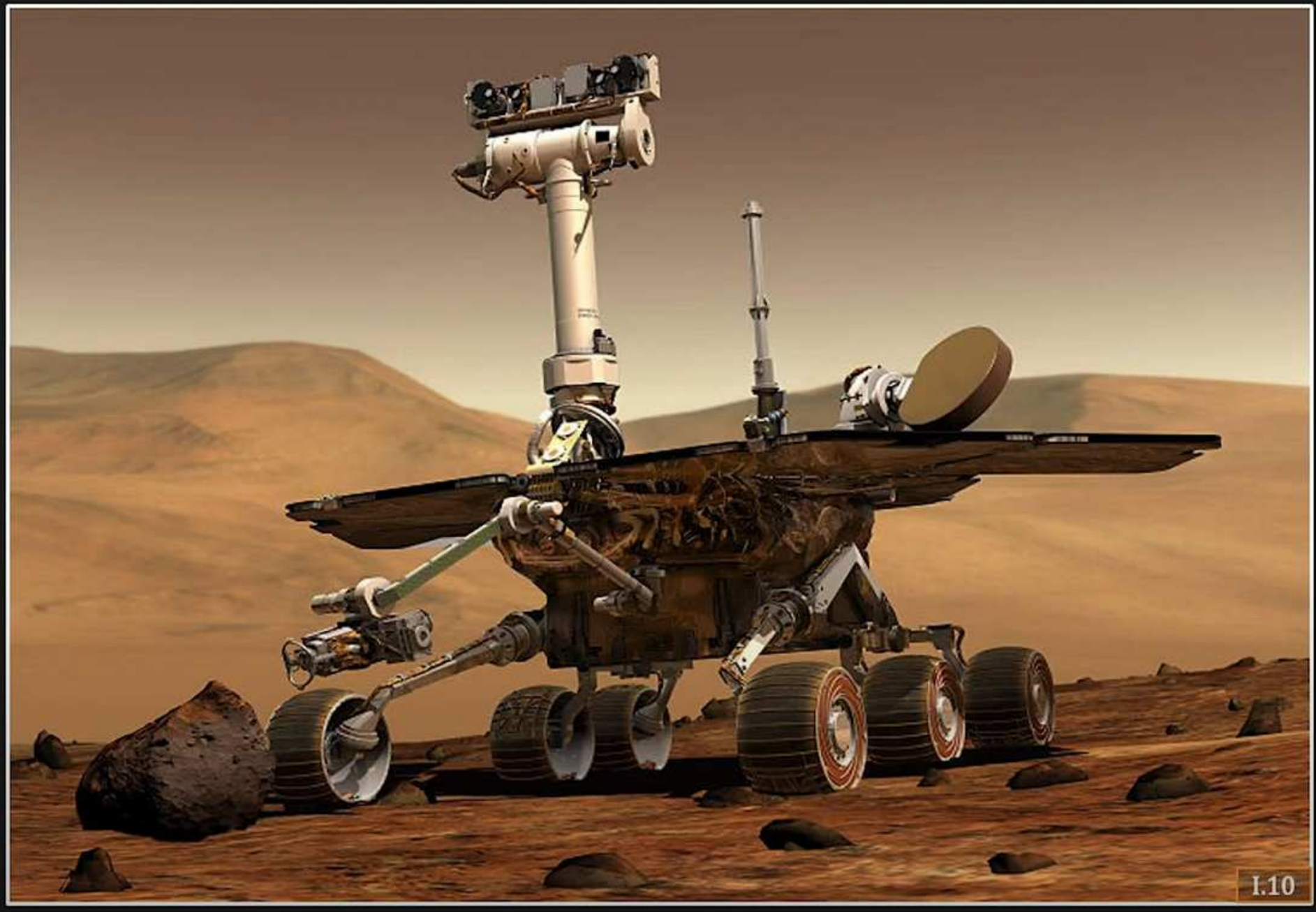


Augmented Reality: Face Manipulation

What is Vision Used For?

Visual Search: Landmark Recognition

What is Vision Used For?

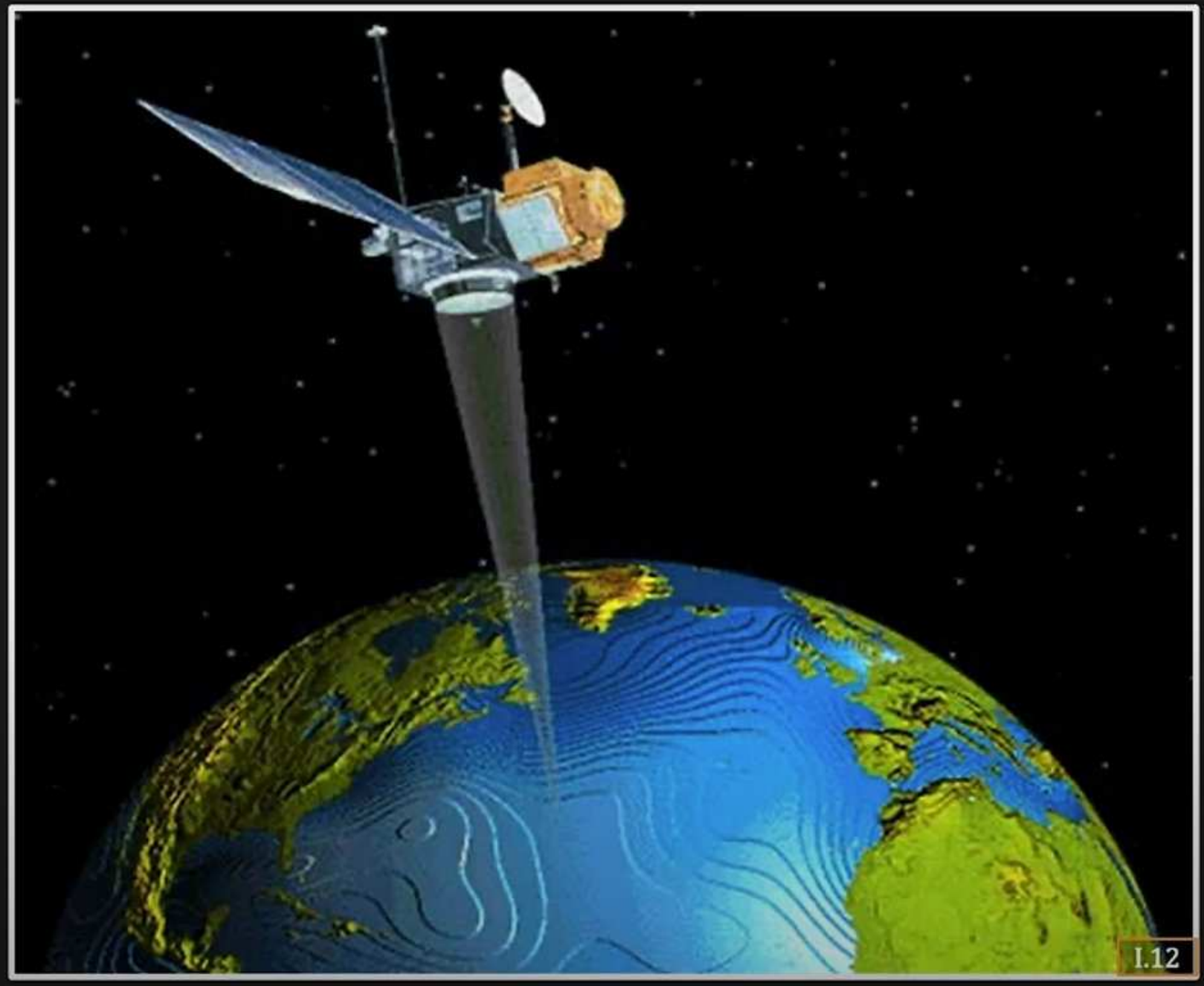


Autonomous Navigation: Space Exploration

What is Vision Used For?

Autonomous Navigation: Driverless Car

What is Vision Used For?



Remote Sensing

What is Vision Used For?

Medical Image Analysis

How do Humans do it

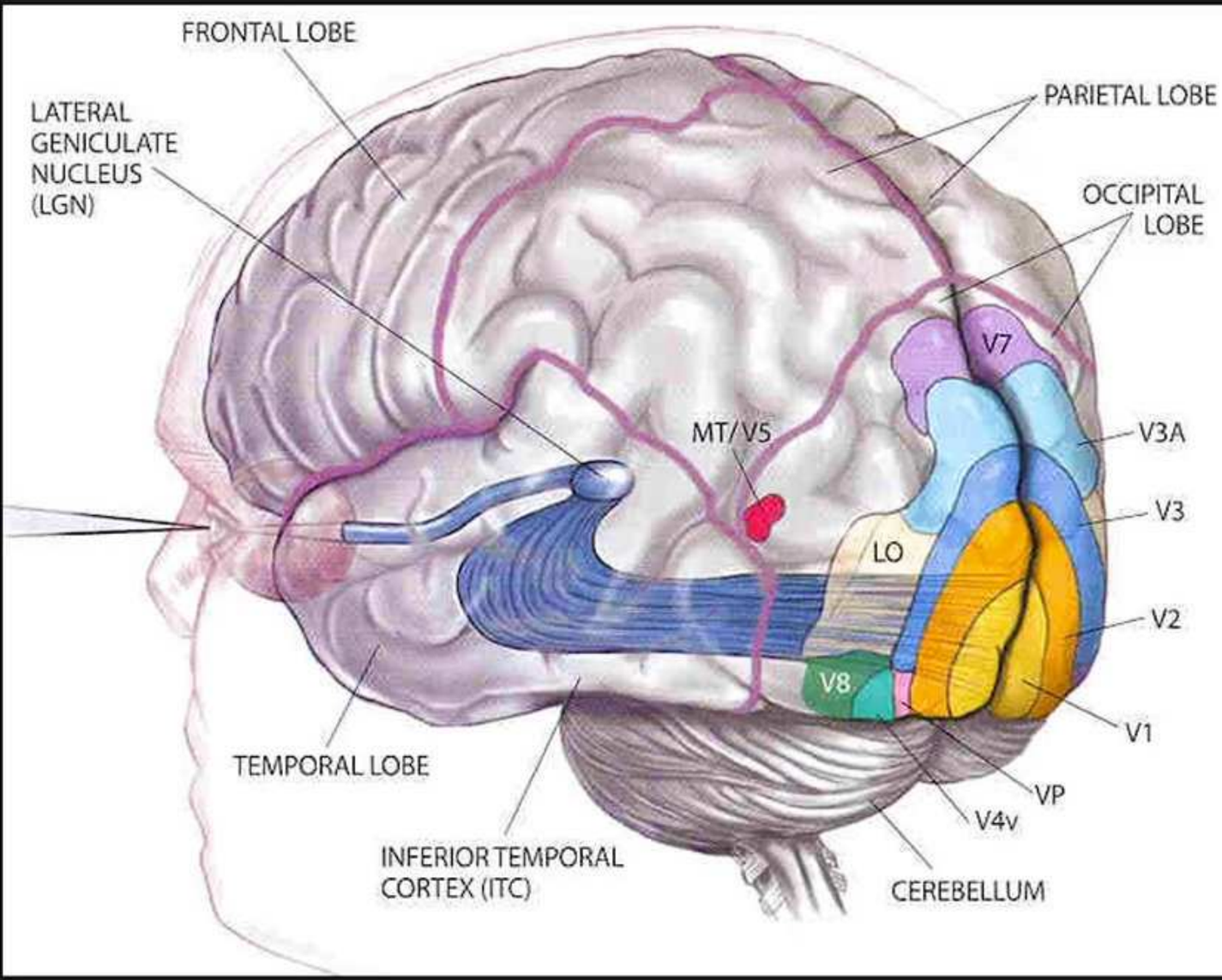
Shree K. Nayar

Columbia University

Topic: Introduction, Module: Introduction

First Principles of Computer Vision

Human Eye and Visual Cortex



KEY TO FUNCTION

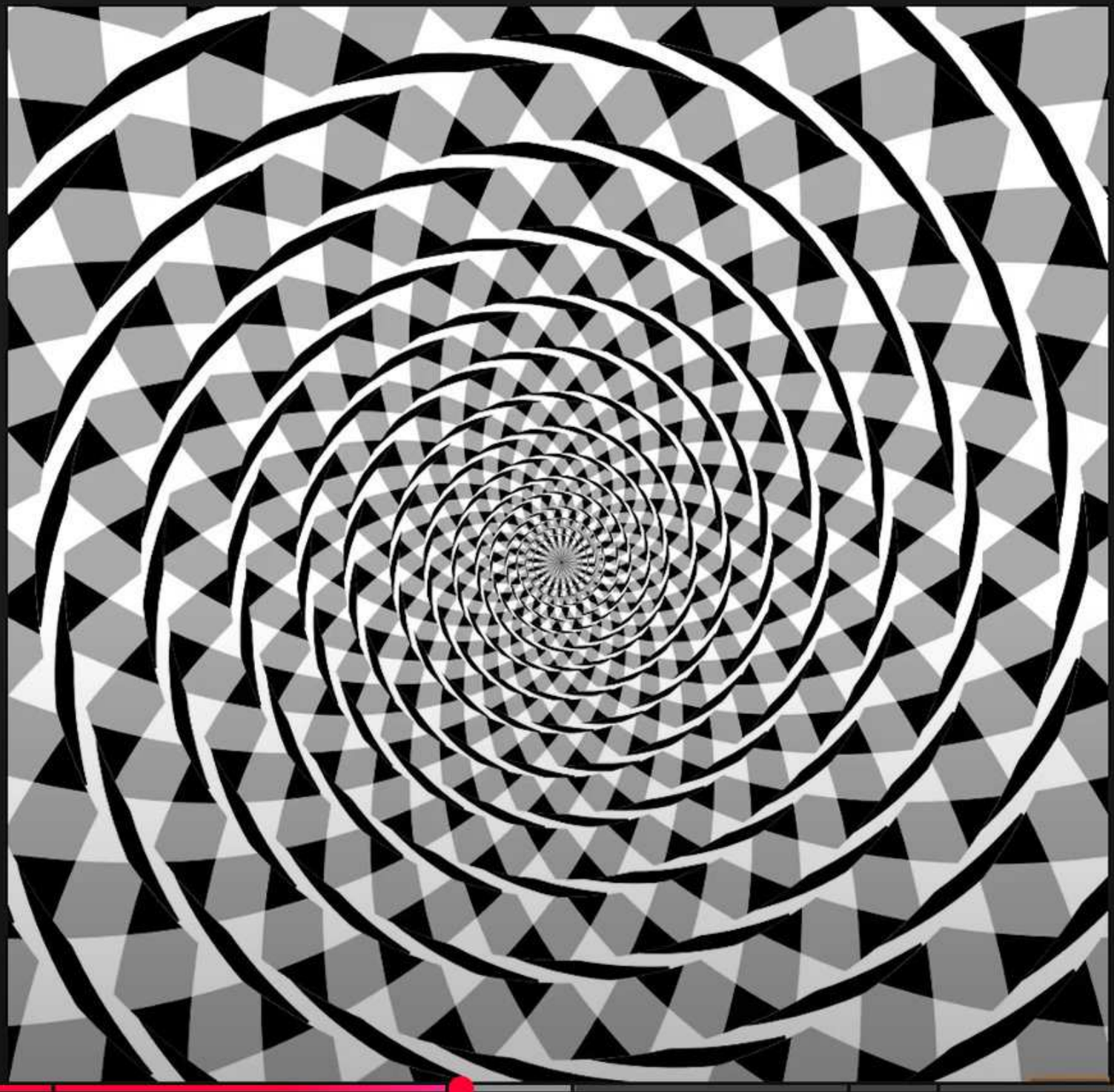
- **V1:** Primary visual cortex; receives all visual input. Begins processing of color, motion and shape. Cells in this area have the smallest receptive fields.
- **V2, V3** and **VP:** Continue processing; cells of each area have progressively larger receptive fields.
- **V3A:** Biased for perceiving motion.
- **V4v:** Function unknown.
- **MT/V5:** Detects motion.
- **V7:** Function unknown.
- **V8:** Processes color vision.
- **LO:** Plays a role in recognizing large-scale objects.

Note: A V6 region has been identified only in monkeys.

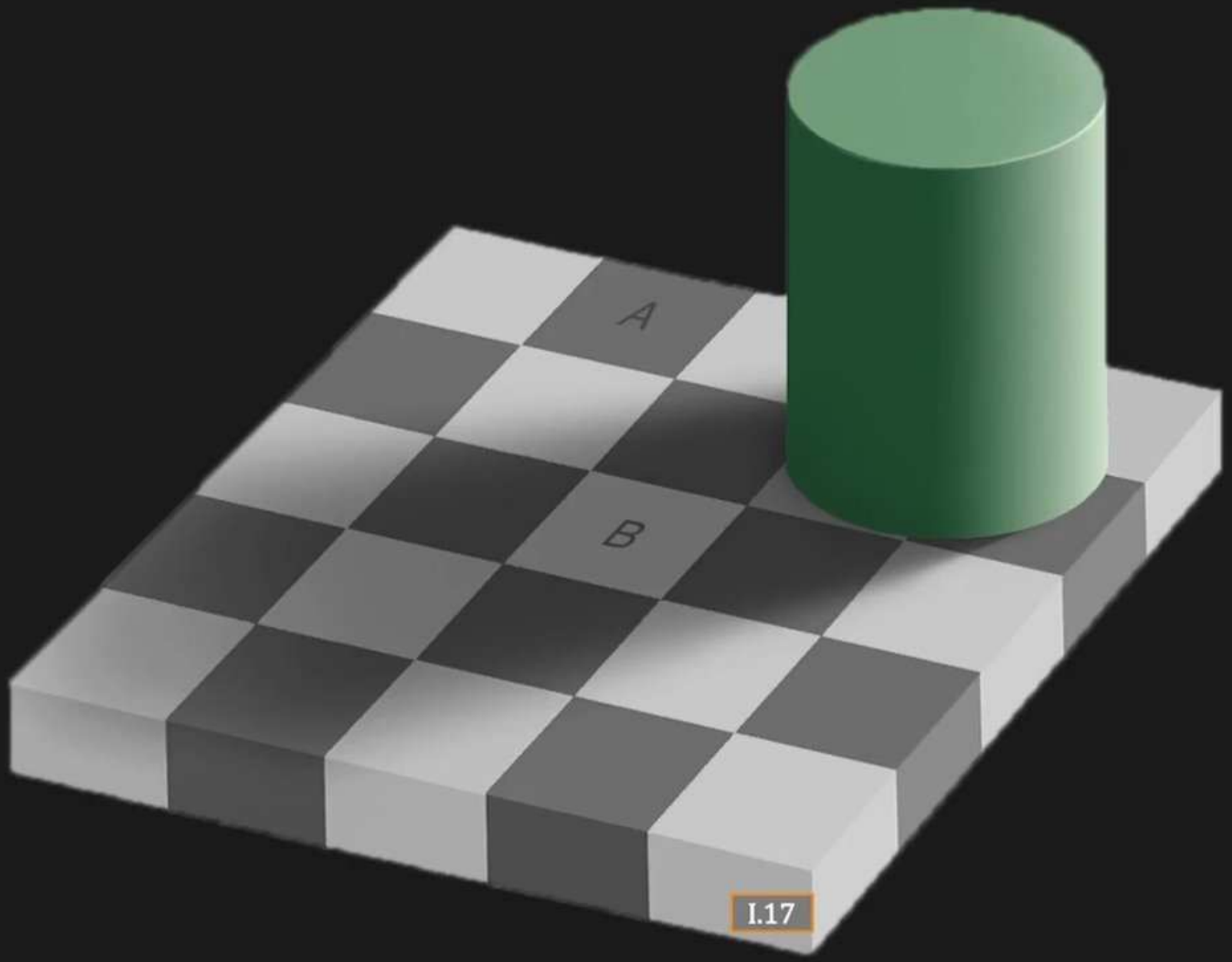
Vision is easy for us

But, we don't fully understand how we do it!

Illusions: Fraser's Spiral

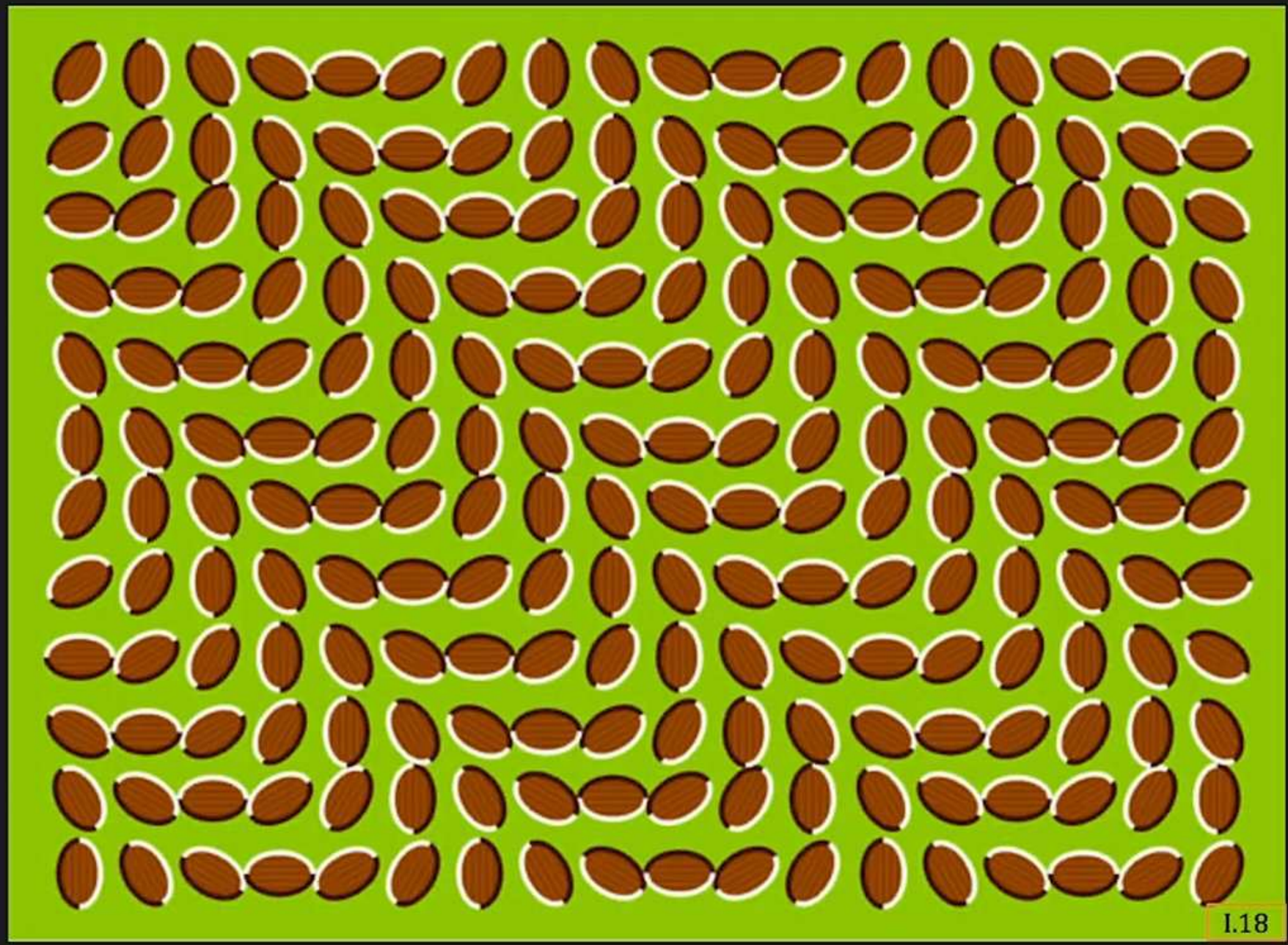


Illusions: Checker Shadow



B seems Brighter than A

Illusions: Donguri Wave

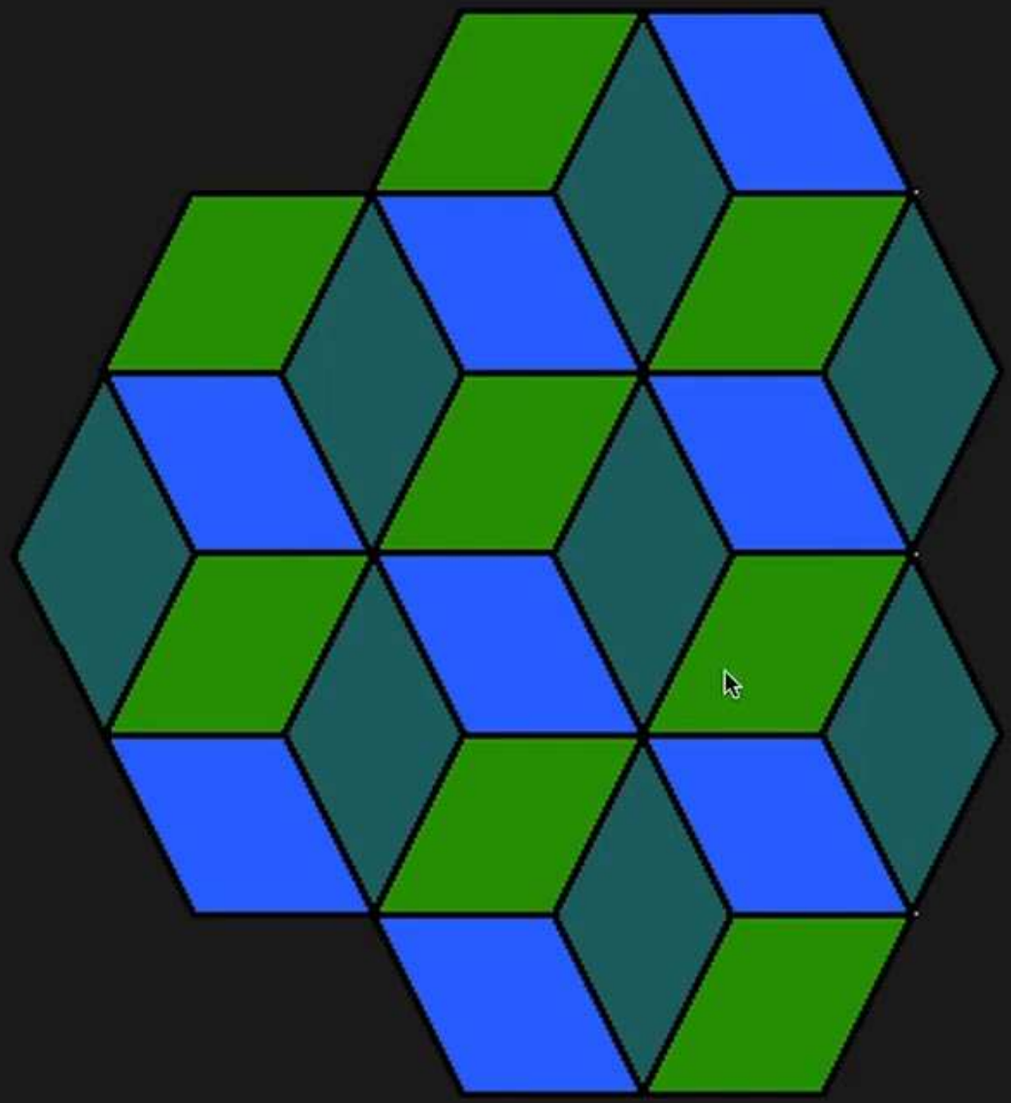


Perceived Motion Without Motion

Illusions: Forced Perspective

These two people are of the same height!

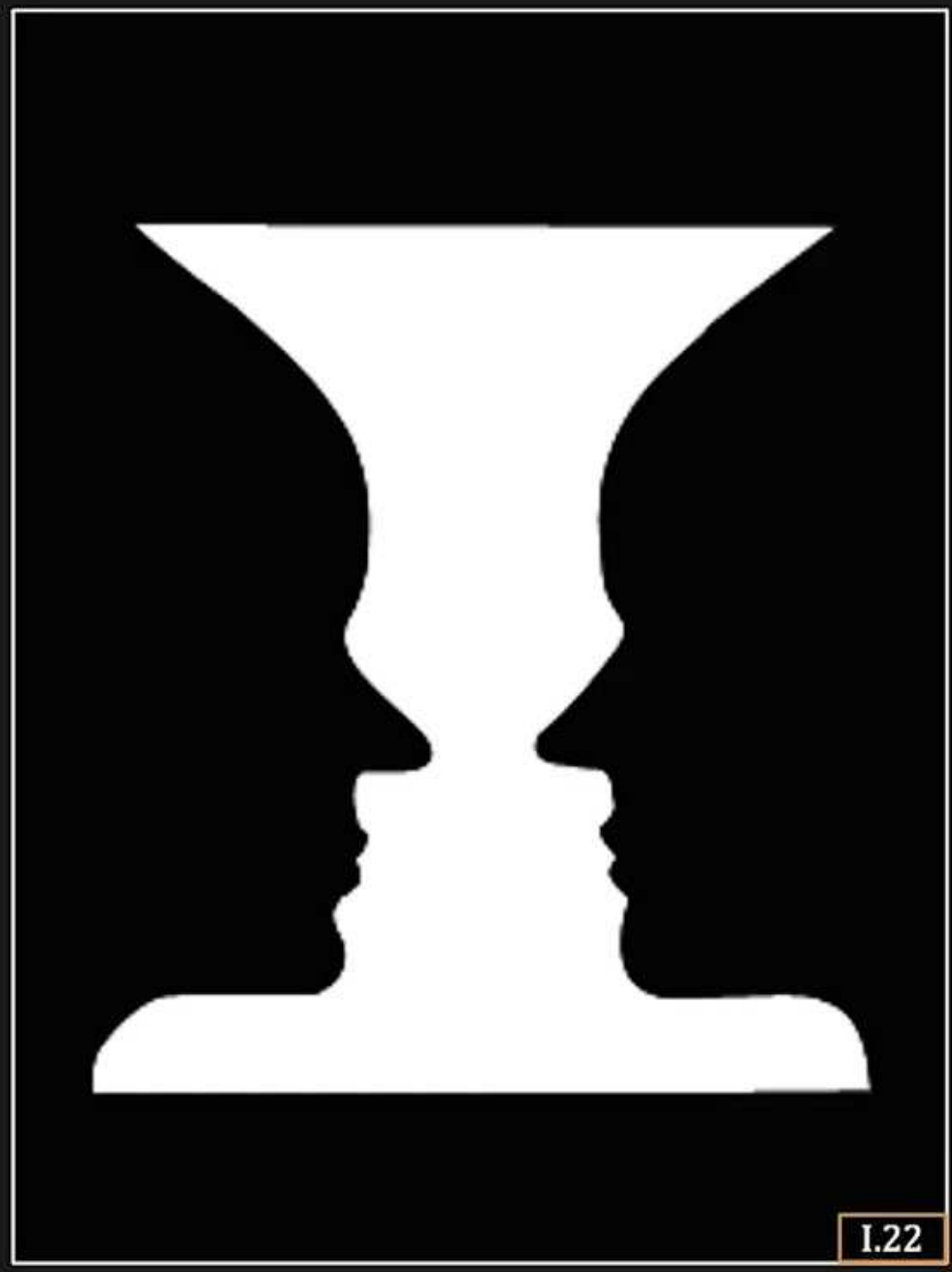
Visual Ambiguities



Six Cubes or Seven Cubes?

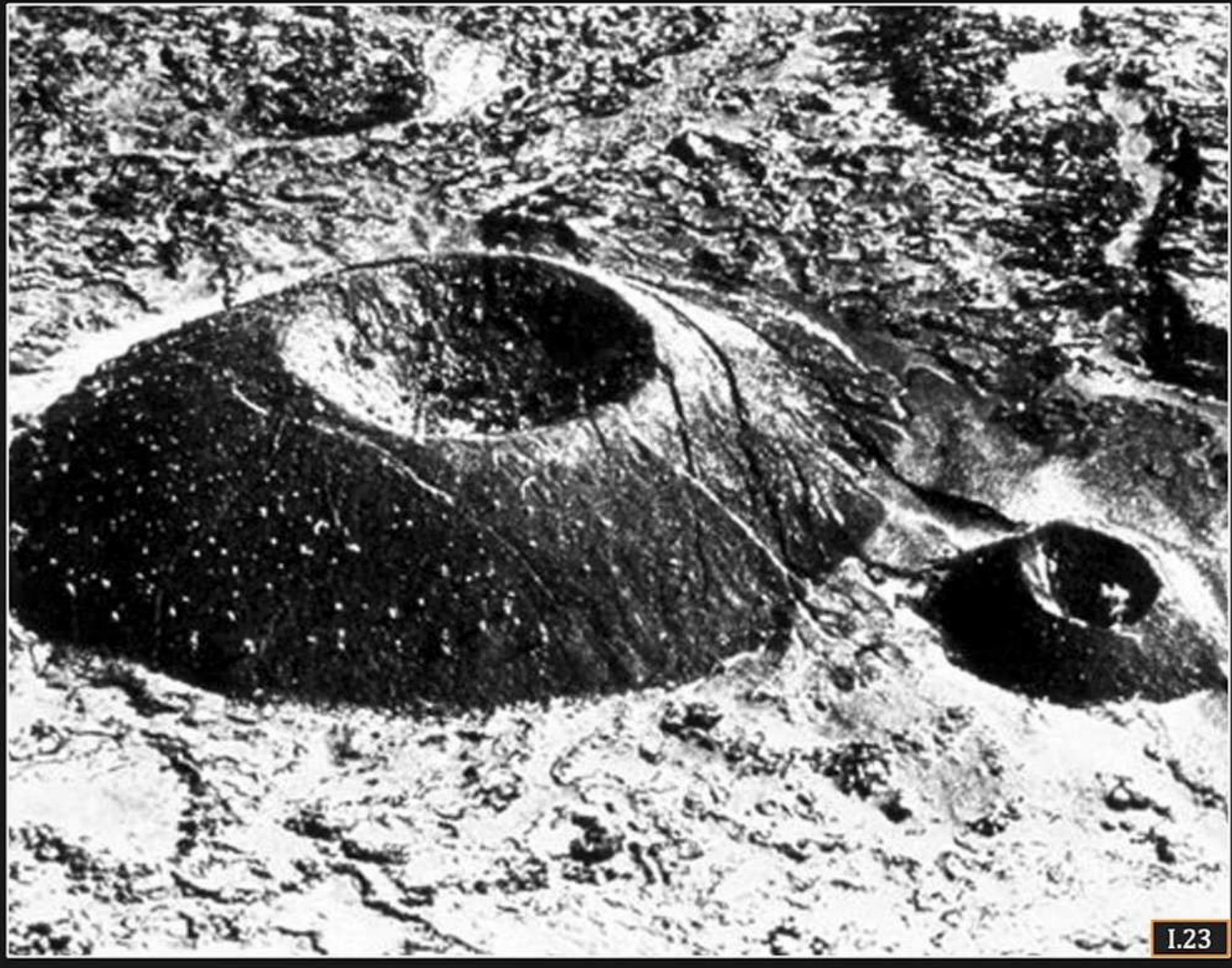
Visual Ambiguities

Young-Girl/Old-Woman



Face/Vase

Visual Ambiguities

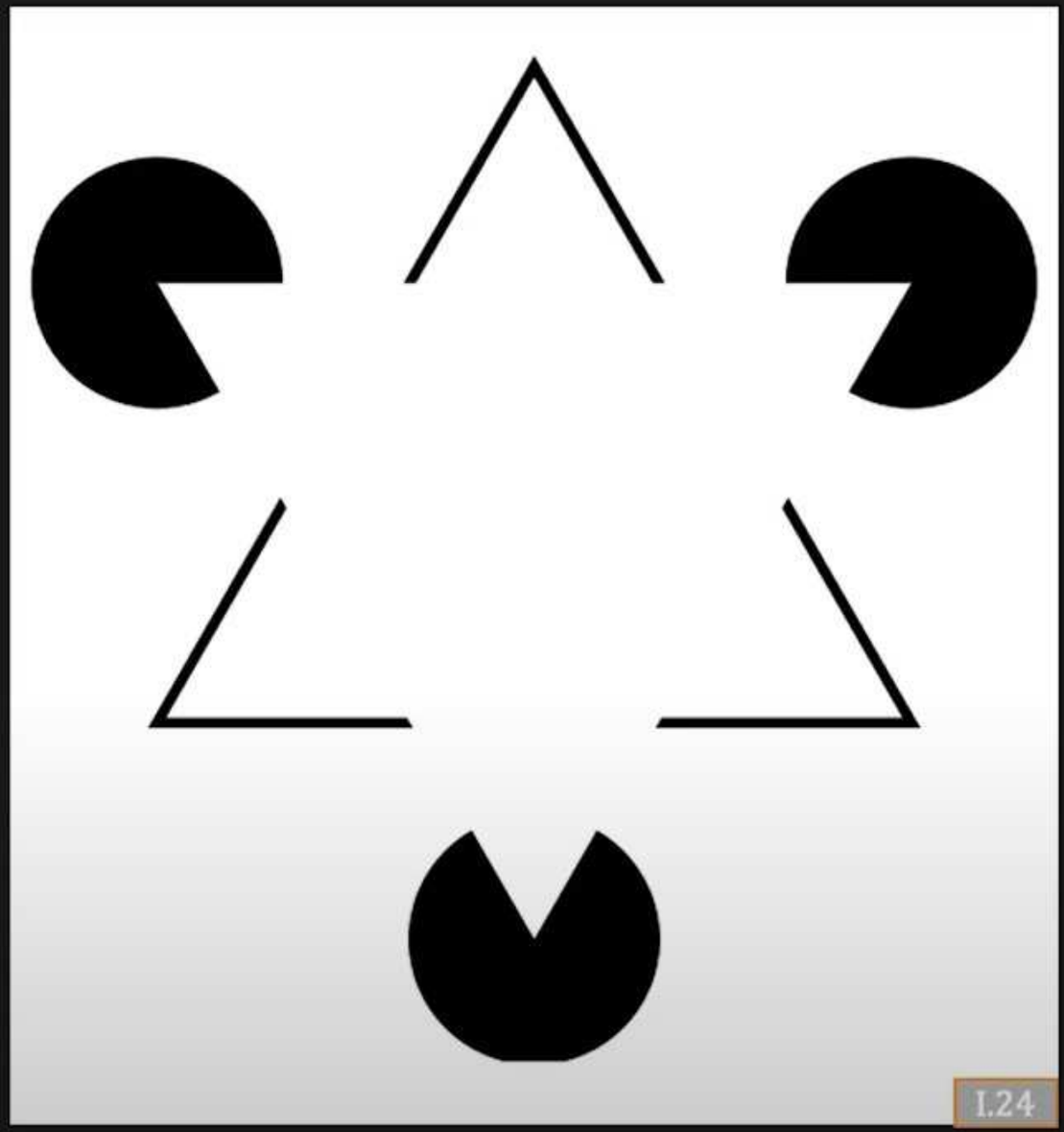


Crater on a Mound

Visual Ambiguities

Mound in a Crater

Seeing vs. Thinking



Kanizsa Triangle

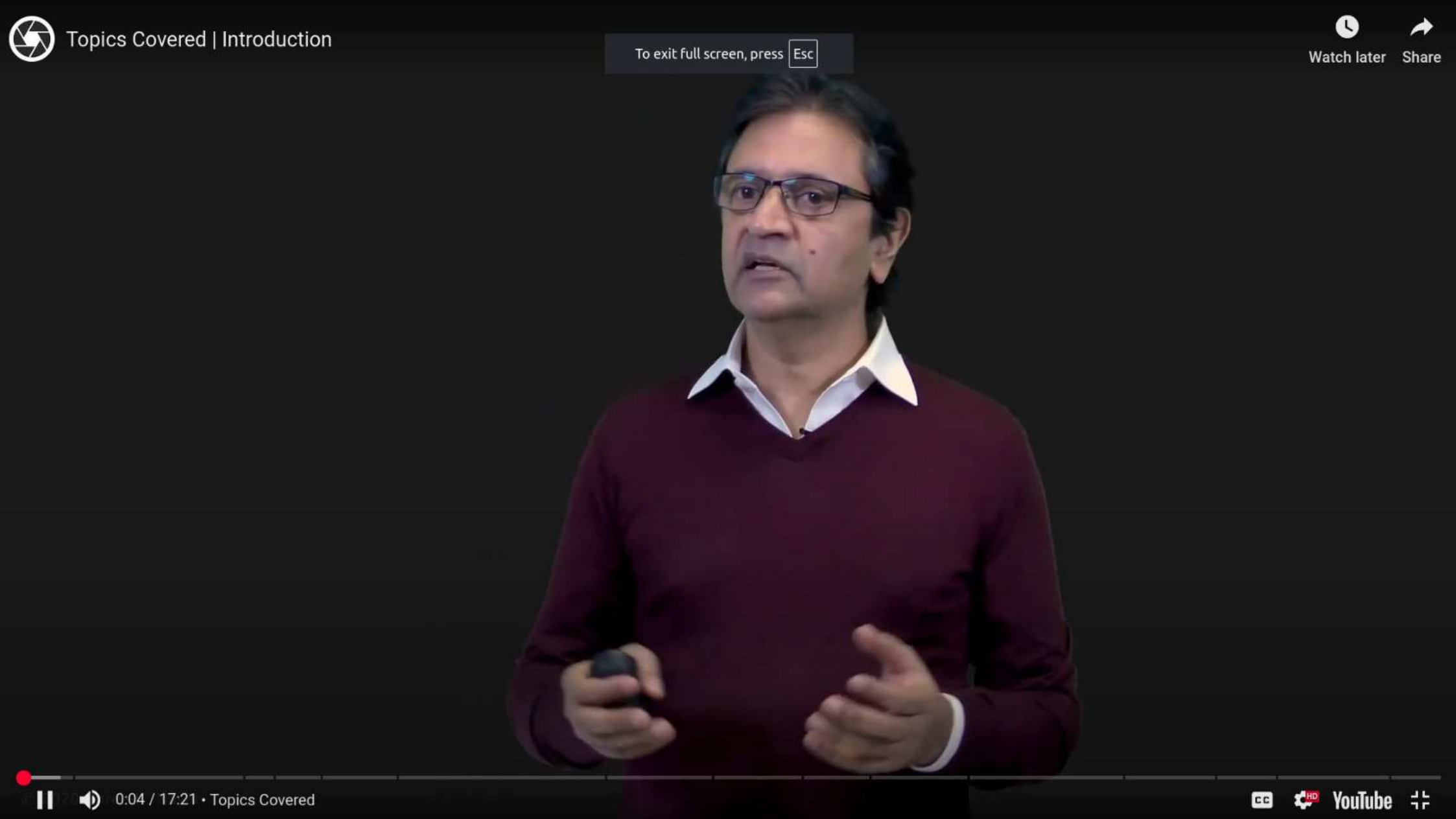
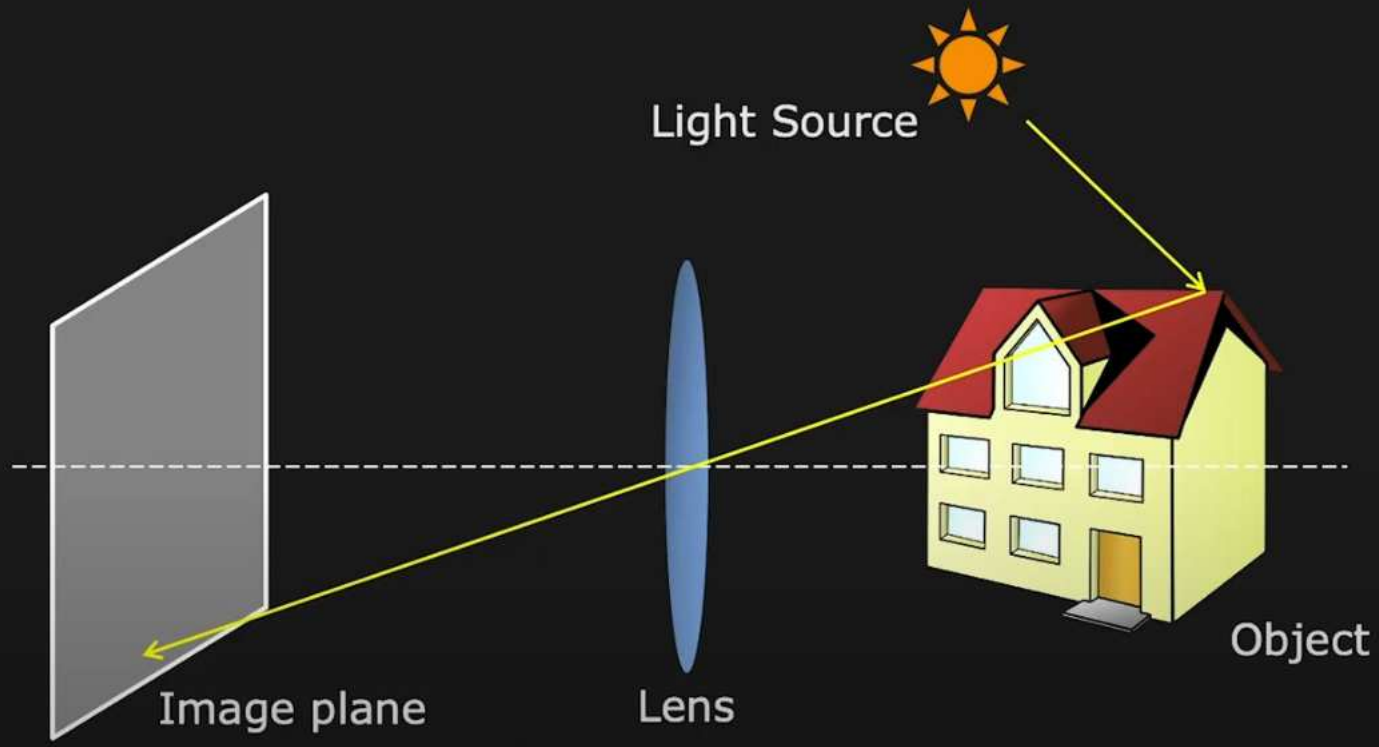

Topics Covered | Introduction
To exit full screen, press Esc
Watch later
Share
0:04 / 17:21 • Topics Covered
CC
HD
YouTube

Image Formation and Optics

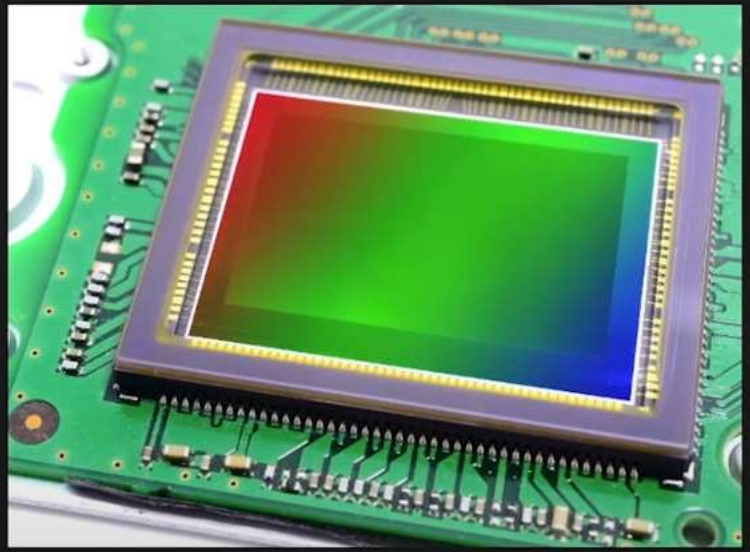
Where do Images Come From?



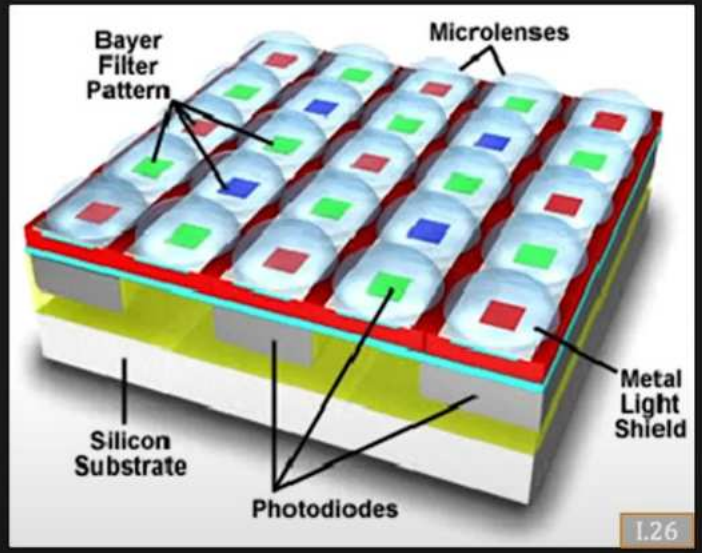
Projection of 3D world on a 2D Plane

Image Sensors

Convert Optical Images to Electrical Signals



Consumer Image Sensor



Typical Structure of Image Sensor

Binary Images

Two-Valued Images: Easy to Store and Process

Grayscale Image

Binary Image

Image Processing

Transform Image to New One that is More Useful

Input Image

Edge-Preserved Smoothing

Edge and Corner Detection

Detecting Intensity Changes in the Image

Input Image

Edges

Boundaries from Edges

Finding Continuous Lines from Edge Segments

2D Recognition using Features

Matching using "Interesting Points"

Object in Database

Input Image and Detected Object

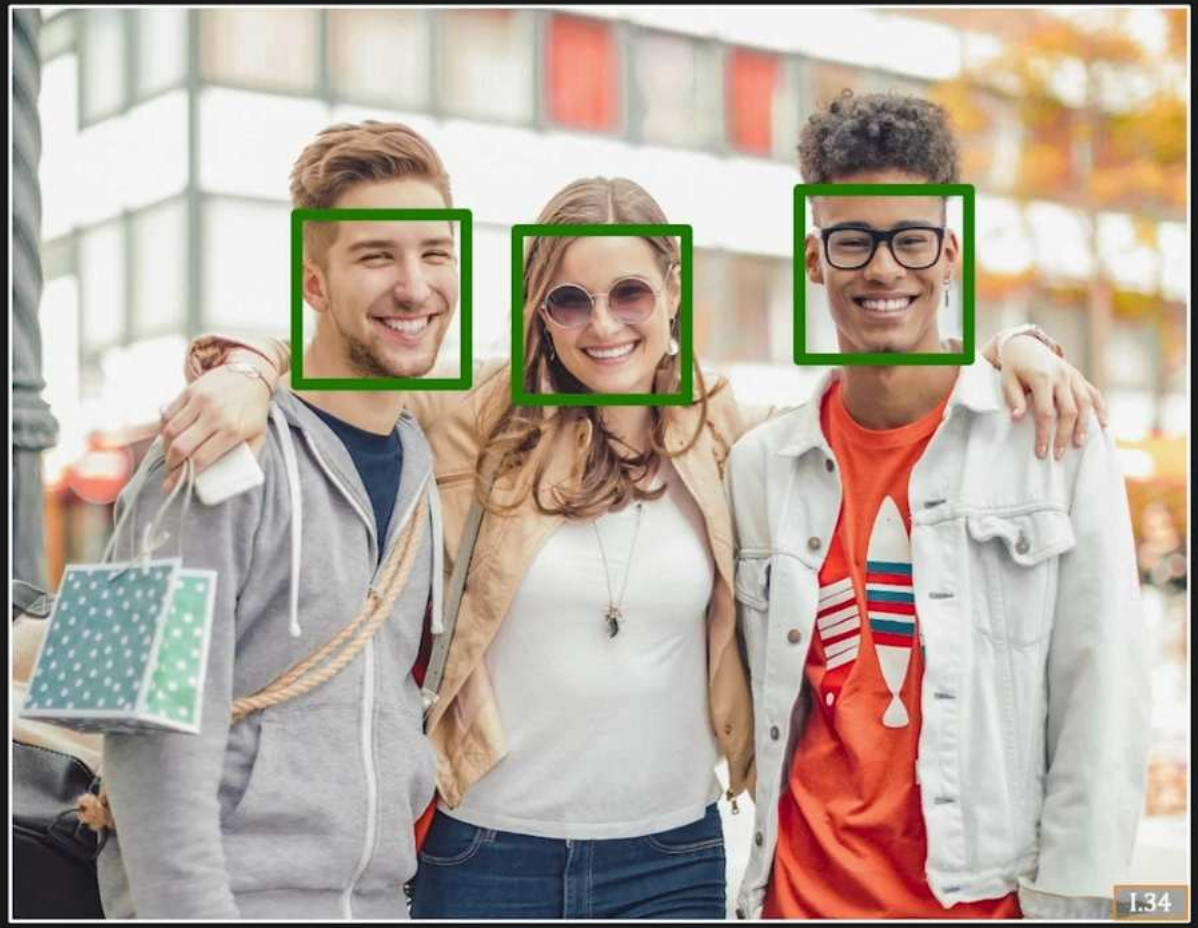
Image Alignment and Stitching

Combine multiple photos to create a larger photo

Source Image 1 Source Image 2 Source Image 3

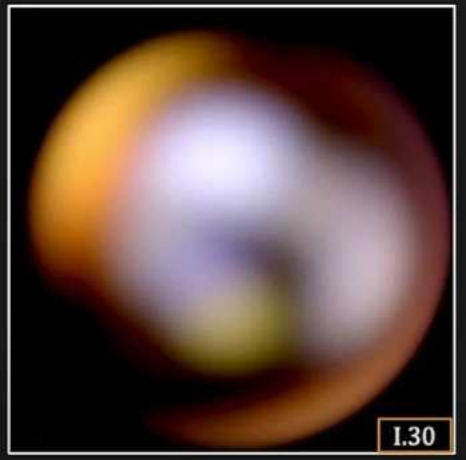
Stitched Image

Face Detection



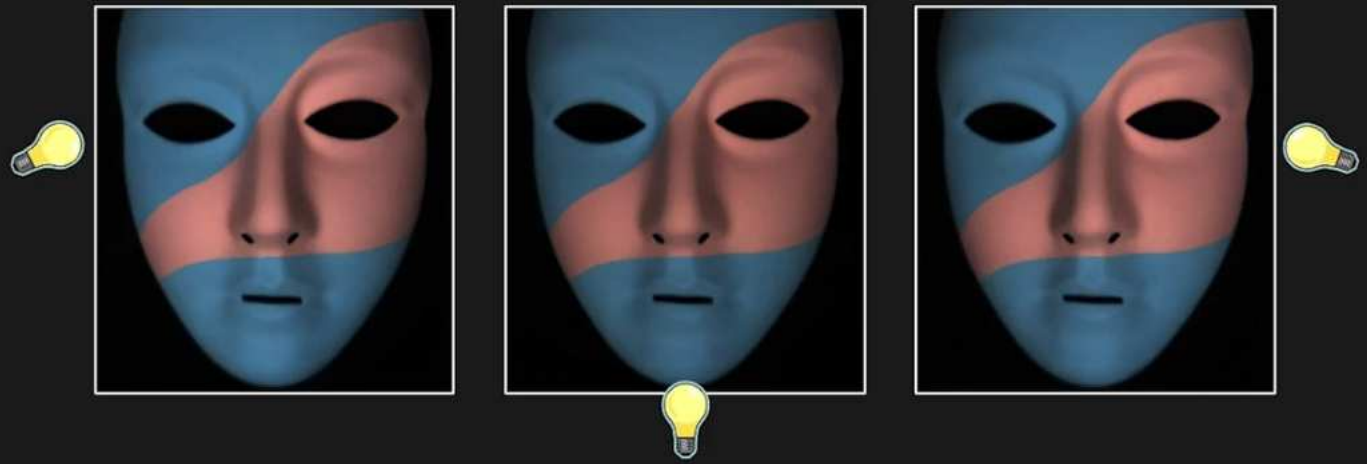
Radiometry and Reflectance

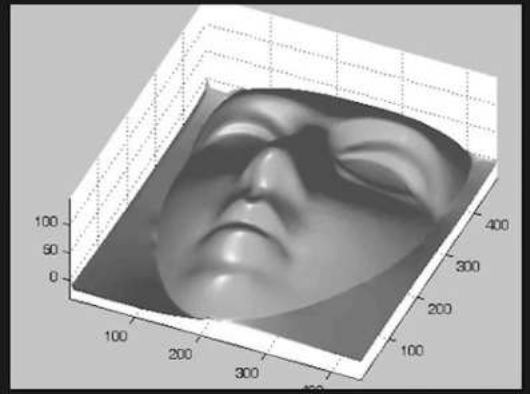
Why do these Spheres Look Different?

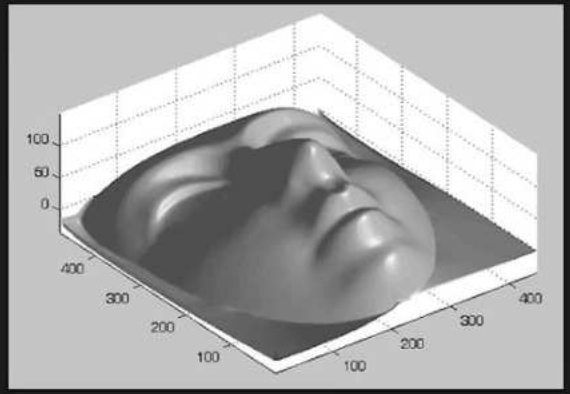


Photometric Stereo

3D Shape from Images under Different Lighting







Computed Shape

Shape from Shading

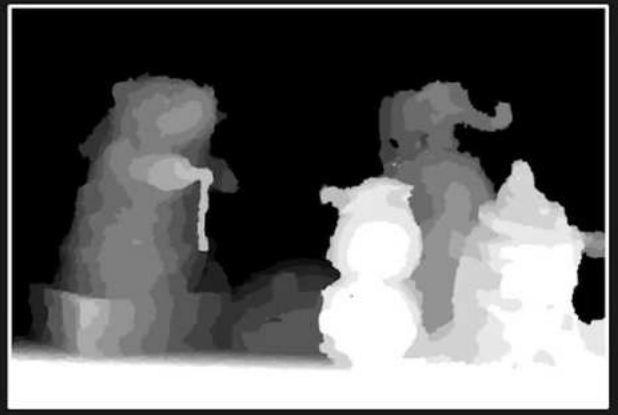
3D Shape from a Single Image

Input Image

Computed Shape

Depth from Focus/Defocus

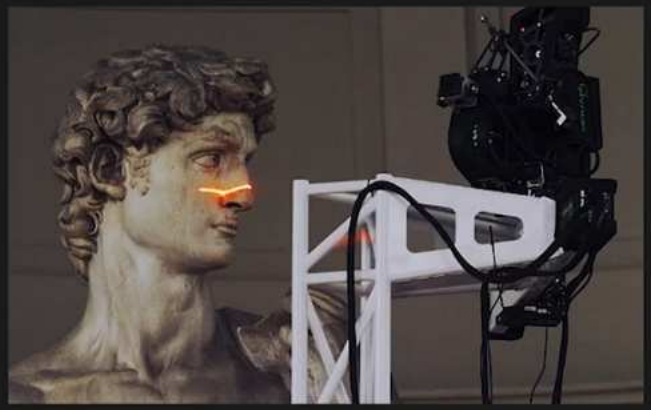
Near-Focus Image



Estimated Depth Map

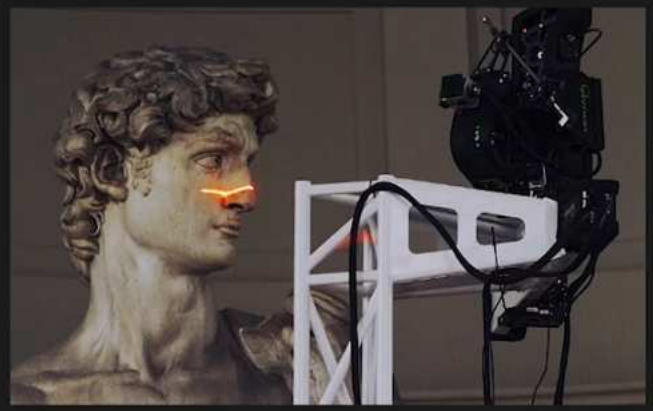
Active Illumination Methods

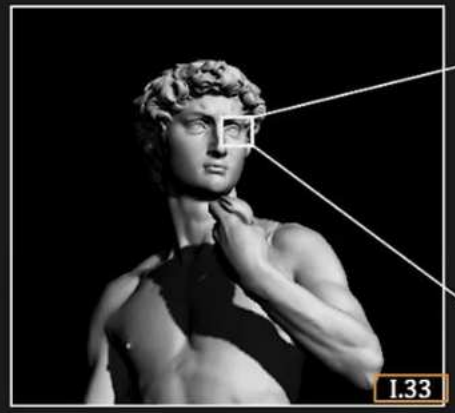
Using Patterned Lighting to Recover Shape

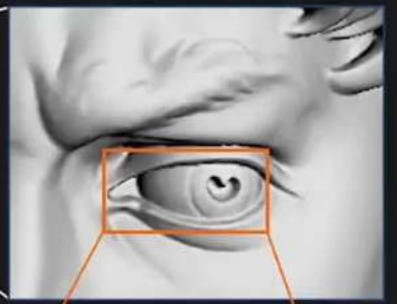


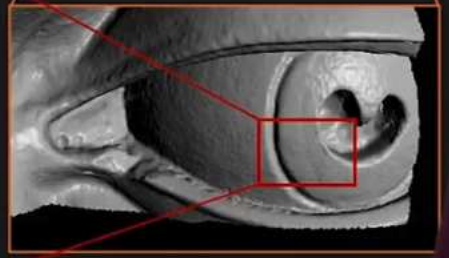
Active Illumination Methods

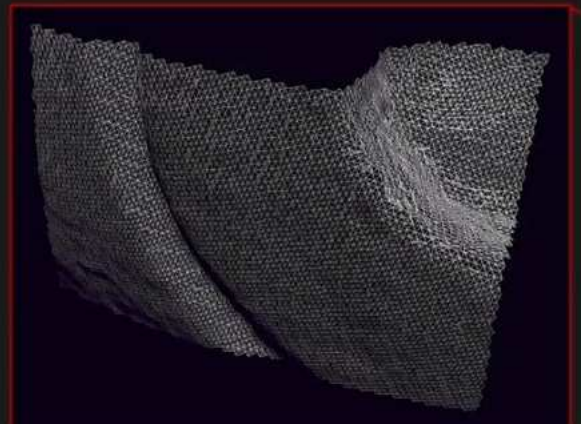
Using Patterned Lighting to Recover Shape











Camera Calibration

Estimating Camera Parameters

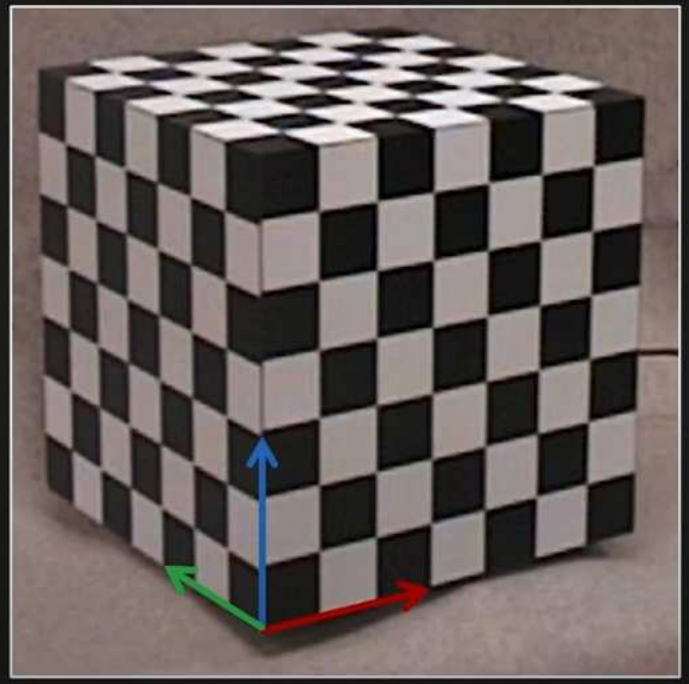


Image of object with known geometry

Binocular Stereo

Computing Depth using Two Views

Right View

Left View

Estimated Depth Map

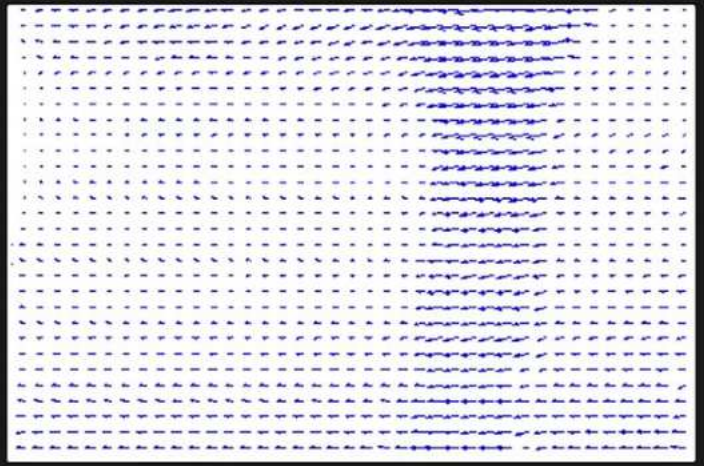
Motion and Optical Flow

Determining the Movement of Scene Points

Frame 1

Frame 2

Frame 3



Estimated Motion

Structure From Motion

Casual Video

Structure From Motion

Casual Video

Reconstructed 3D Structure

Image Segmentation

Group pixels with similar visual characteristics.

Input Image

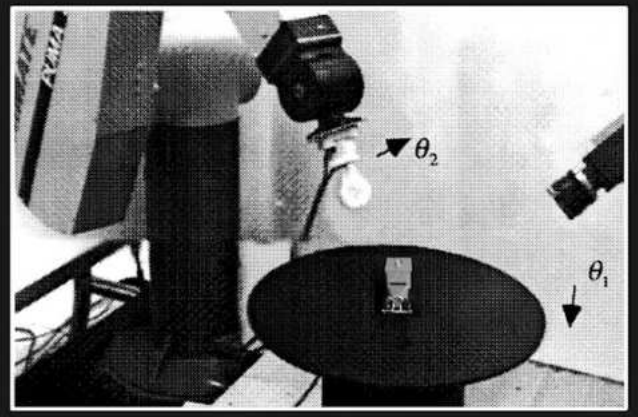
Segmented Image

Object Tracking

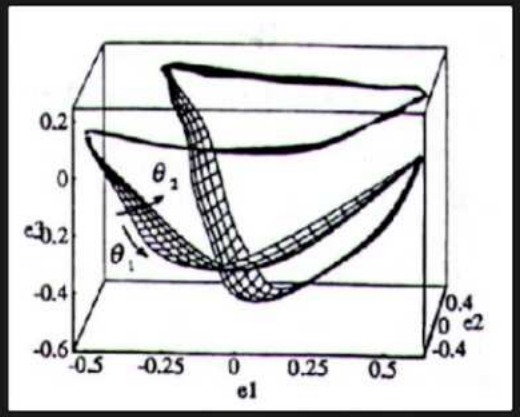
Determining the Movement of Objects in Videos

Appearance Matching

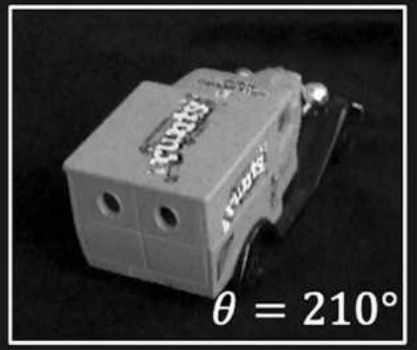
Object Recognition using Principle Component Analysis



Learning Object Appearance



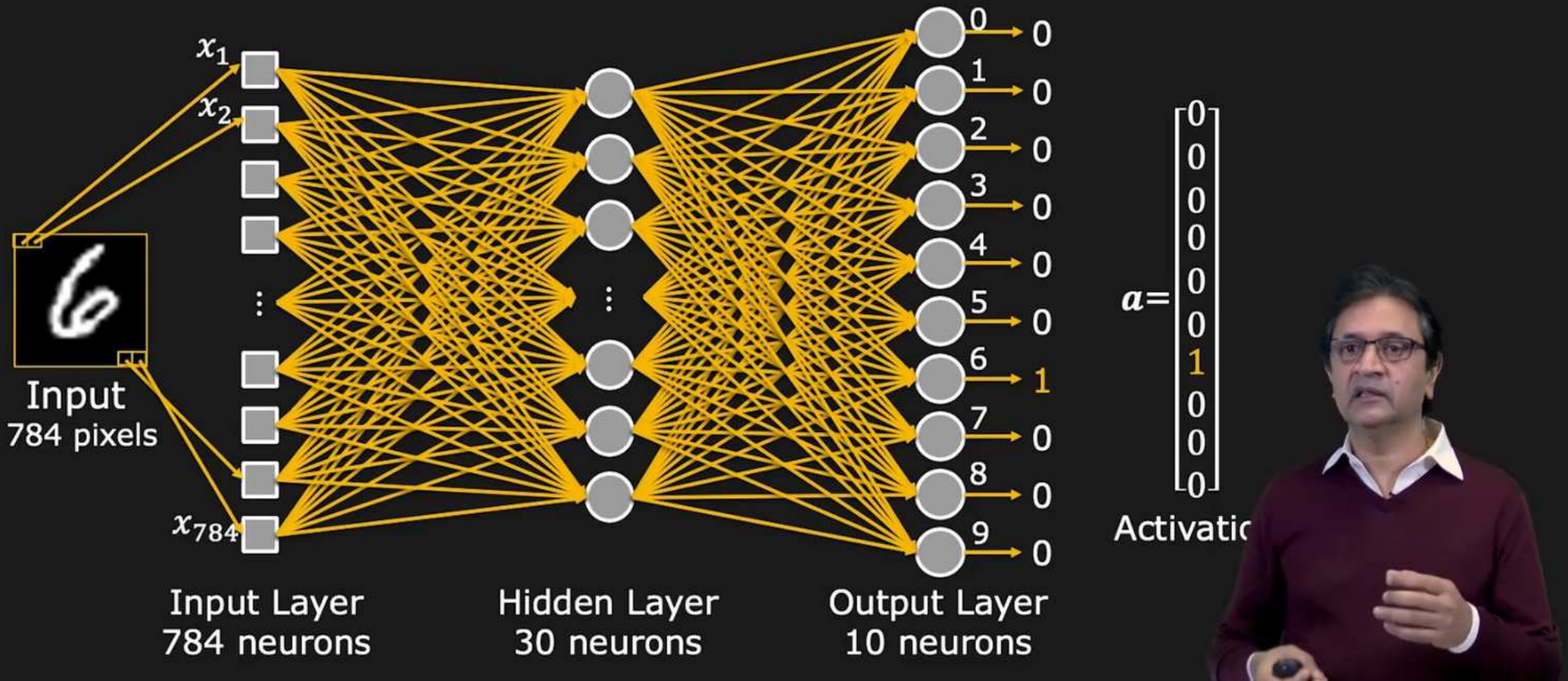
Appearance Manifold (Model)



Recognition by Matching Appearance

Artificial Neural Networks

Using Network of Neurons to Solve Complex Problems



Modules and Prerequisites

Modules:

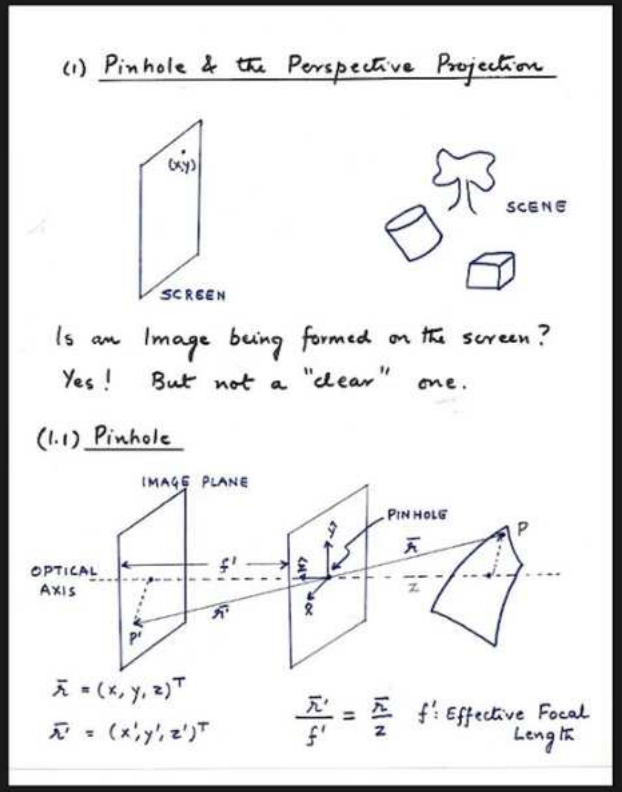
0. Introduction
1. Imaging: Image Formation, Sensing, Processing
2. Features: Edges, Boundaries, SIFT, Applications
3. Reconstruction 1: Shading, Focus, Active Illumination
4. Reconstruction 2: Stereo, Optical Flow, SFM
5. Perception: Segmentation, Tracking, Recognition

Prerequisites:

- Fundamental of Linear Algebra
- Fundamentals of Calculus
- One Programming Language

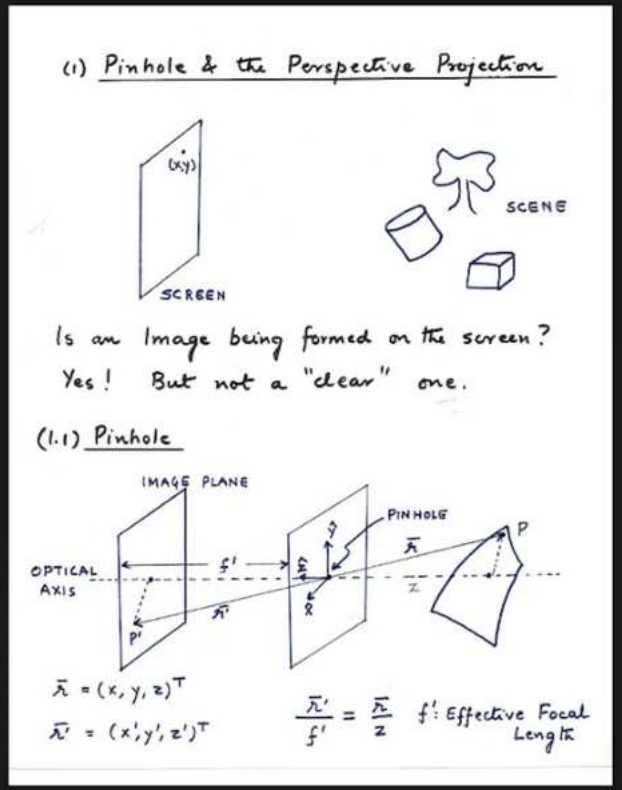
About the Slides

Once Upon a Time:



About the Slides

Once Upon a Time:



Slides Thanks to:

Jinwei Gu, Neeraj Kumar, Changyin Zhou, Oliver Cossairt, Guru Krishnan, Mohit Gupta, Daniel Miau, Manushree Gangwar, Avinash Nair, Parita Pooj, Henry Xu, Robert Colgan, Anne Fleming

Example: Math Primer Slide

$$\boxed{e^{i\theta} = \cos\theta + i\sin\theta} \qquad i = \sqrt{-1}$$

Expand $e^{i\theta}$ using Taylor Series:

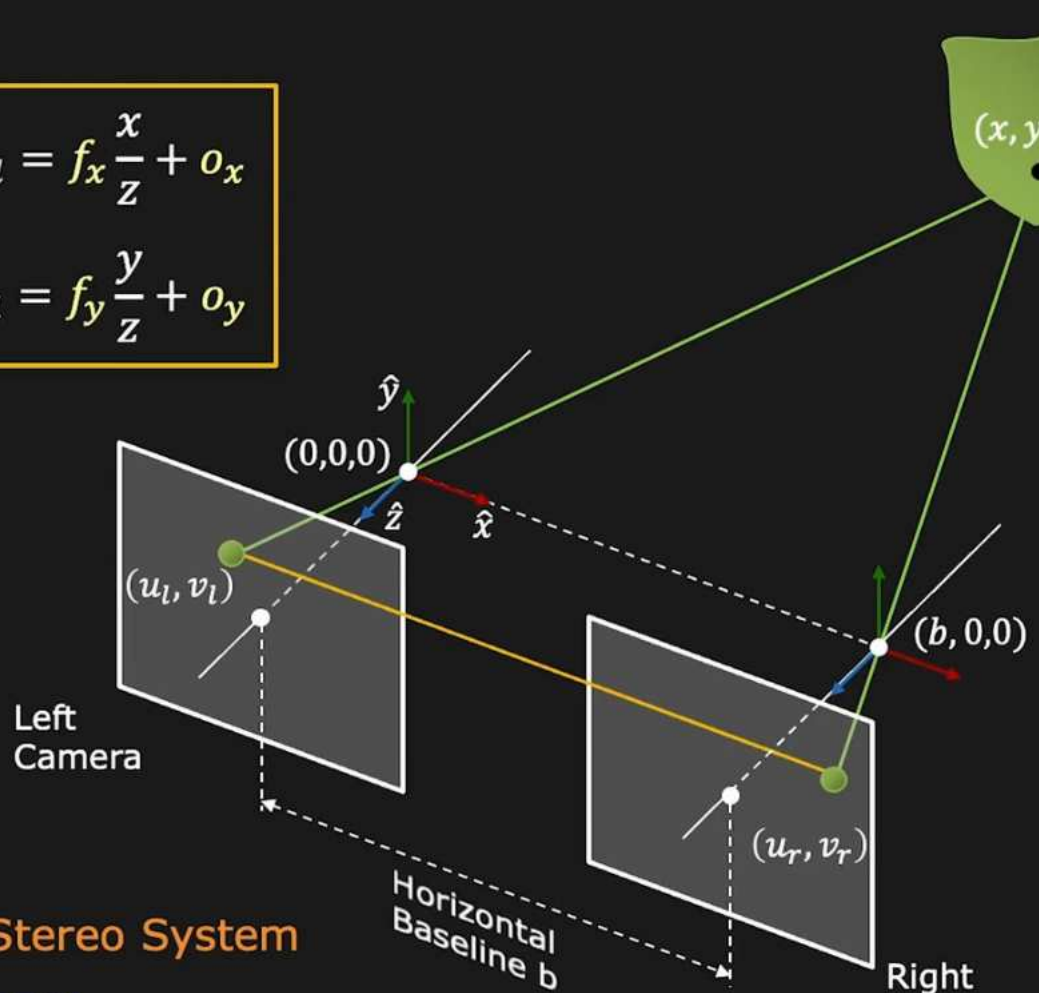
$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \frac{(i\theta)^4}{4!} + \frac{(i\theta)^5}{5!} + \frac{(i\theta)^6}{6!} + \ldots$$

$$e^{i\theta} = \underbrace{\left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \cdots\right)}_{\cos\theta} + i\underbrace{\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \frac{\theta^7}{7!} + \cdots\right)}_{\sin\theta}$$

Example: Review Slide

$$u_l = f_x \frac{x}{z} + o_x$$

$$v_l = f_y \frac{y}{z} + o_y$$



$$u_r = f_x \frac{x - b}{z} + o_x$$

$$v_r = f_y \frac{y}{z} + o_y$$

f_x, f_y, b, o_x, o_y are

Stereo System

REVIEW

Recommended Texts

Computer Vision: Algorithms and Applications (Vision)
Szeliski, R., Springer

Computer Vision: A Modern Approach (Vision)
Forsyth, D and Ponce, J., Prentice Hall

Robot Vision (Vision)
Horn, B. K. P., MIT Press

A Guided Tour of Computer Vision (Vision)
Nalwa, V., Addison-Wesley

Digital Image Processing (Image Processing)
González, R and Woods, R., Prentice Hall

Optics (Optics)
Hecht, E., Addison-Wesley

Eye and Brain (Human Vision)
Gregory, R., Princeton University Press

Animal Eyes (Biological Vision)
Land, M. and Nilsson, D., Oxford University Press