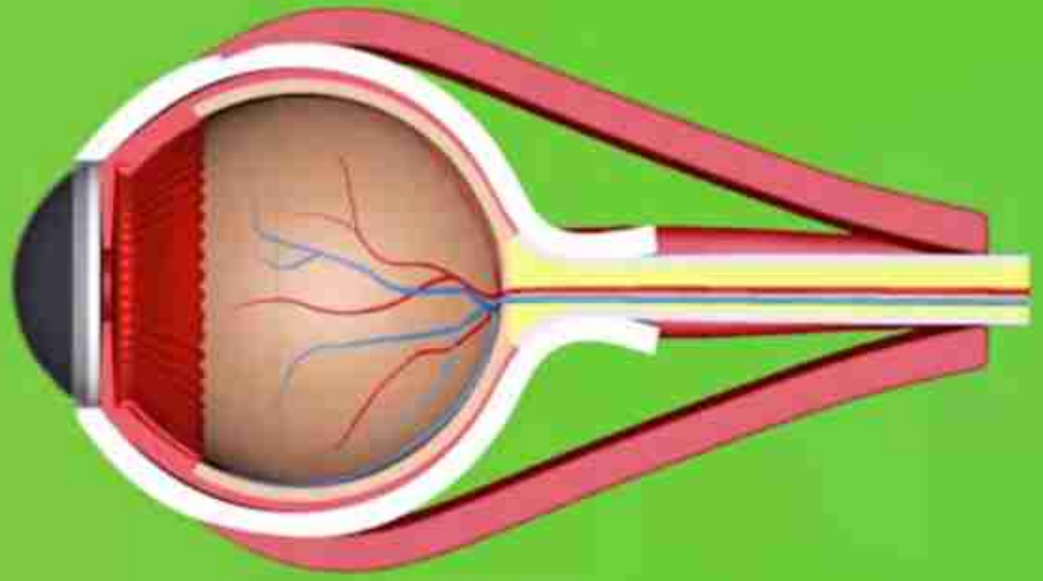
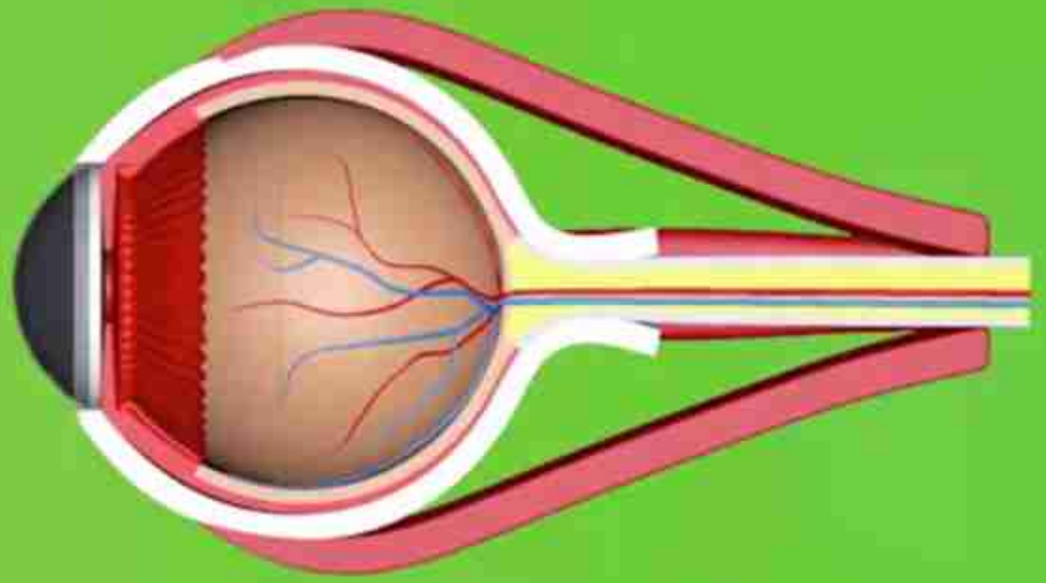
Light Rays



Pass through retina

Share

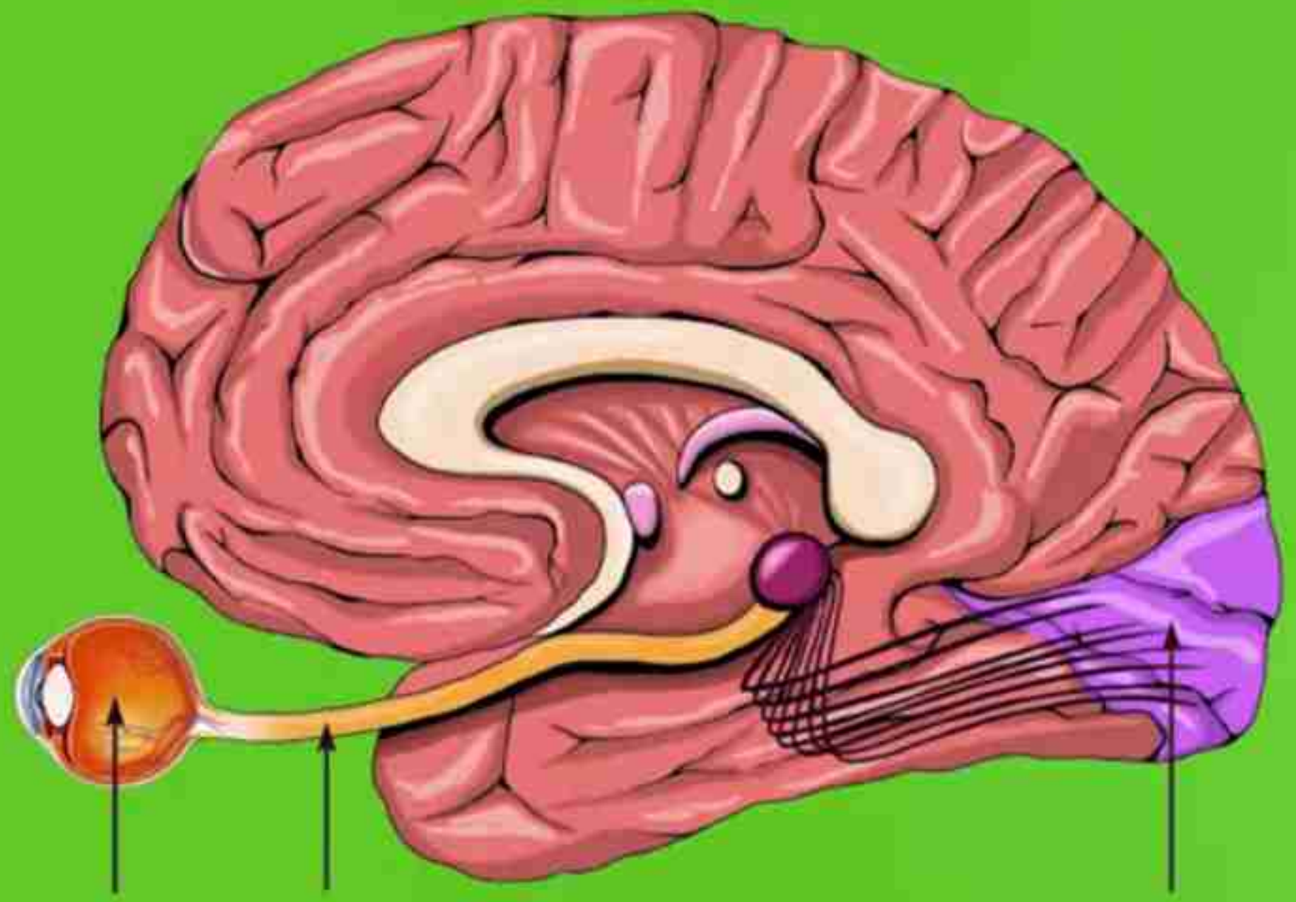


Pass through retina

Processing

1:14 / 6:38

Processing



# Computer Vision

theoretical and algorithmic basis to
achieve automatic visual understanding

# Computer Vision

high-level understanding from digital images or videos

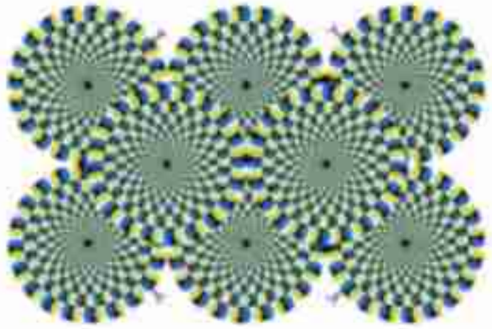
# Computer Vision

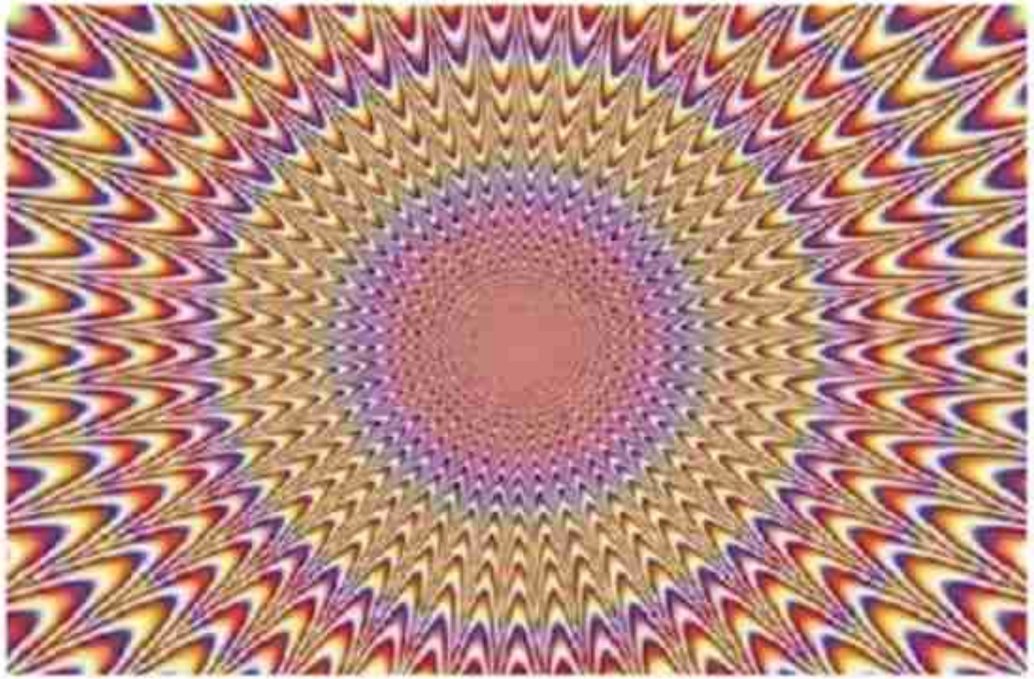
computer vision aims to come up with computational models of the human visual system

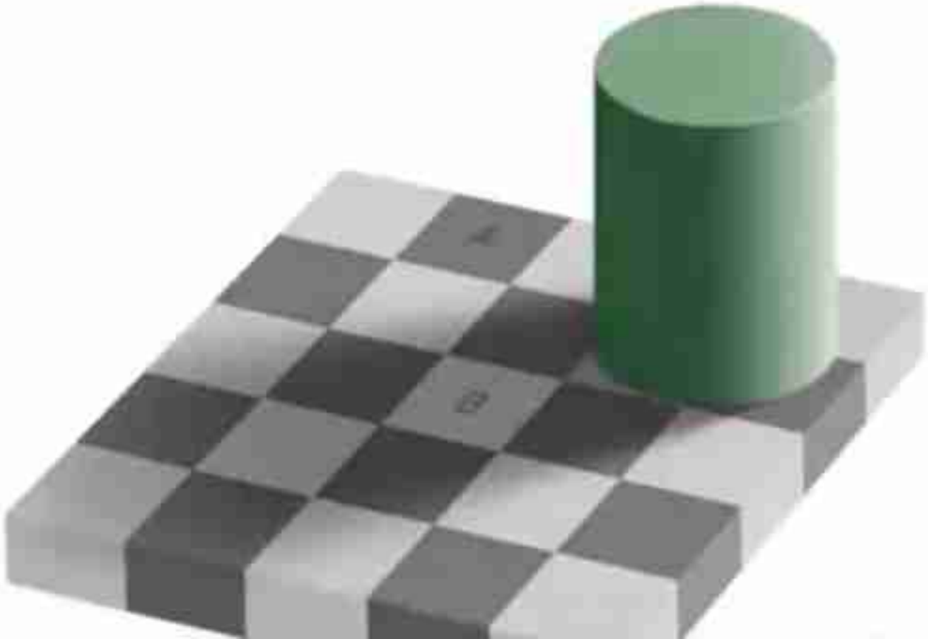
# Computer Vision

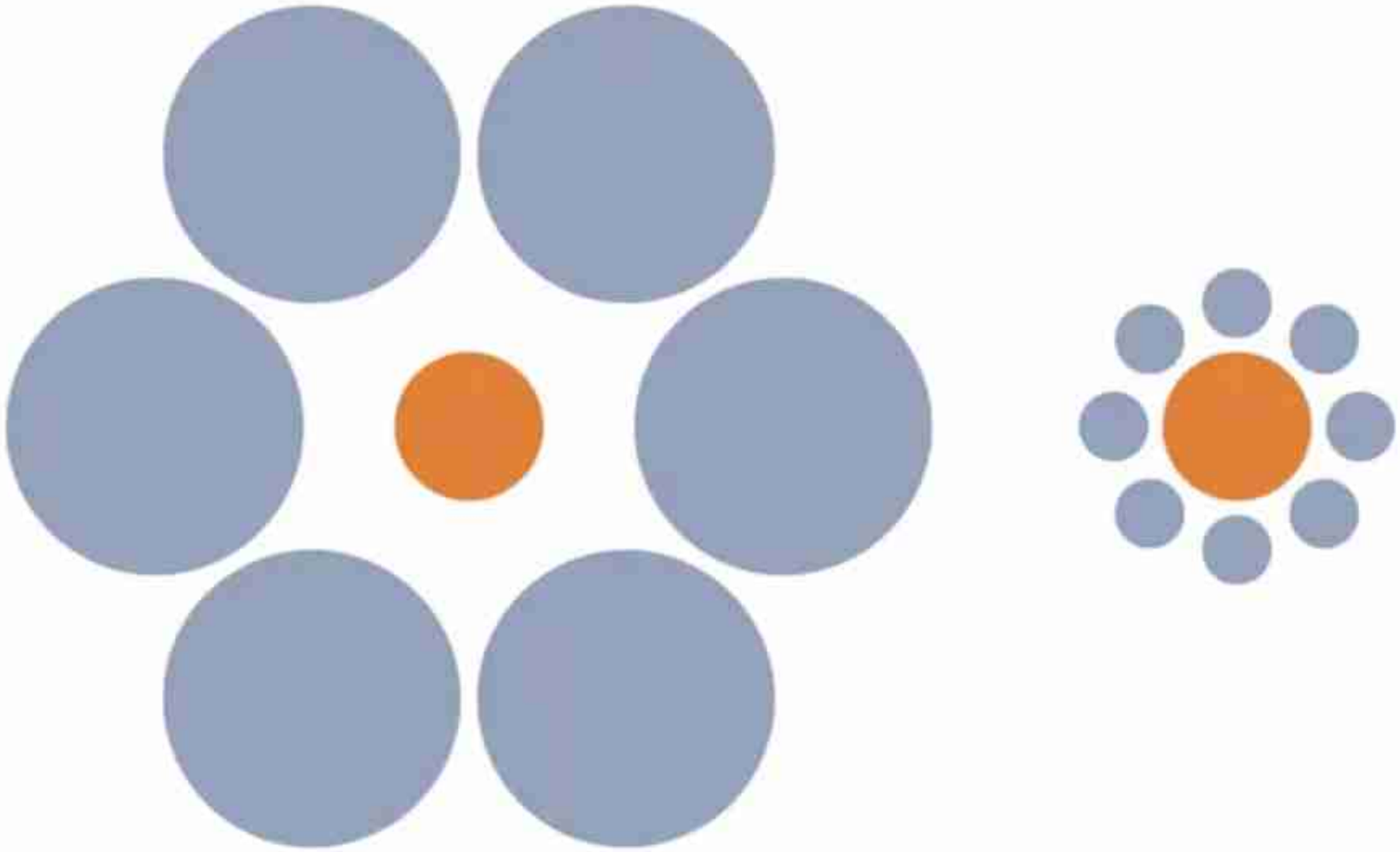
computer vision aims to build autonomous systems to perform some of the tasks which the human visual system can perform and even surpass it in many cases

# Optical Illusions





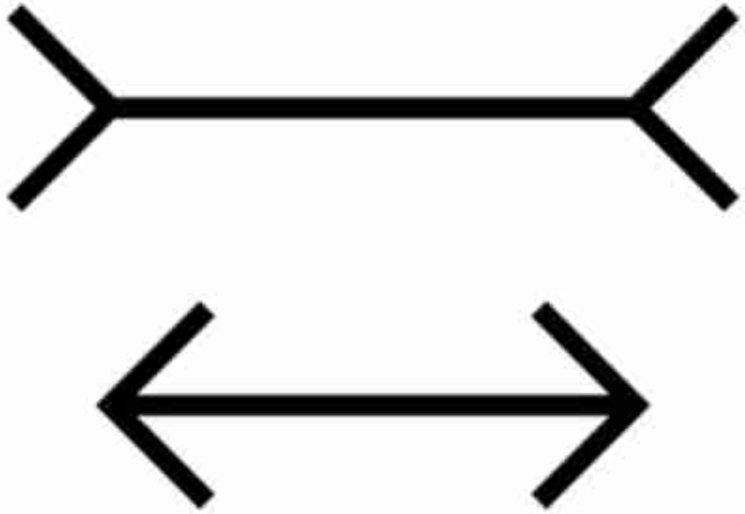

A
B



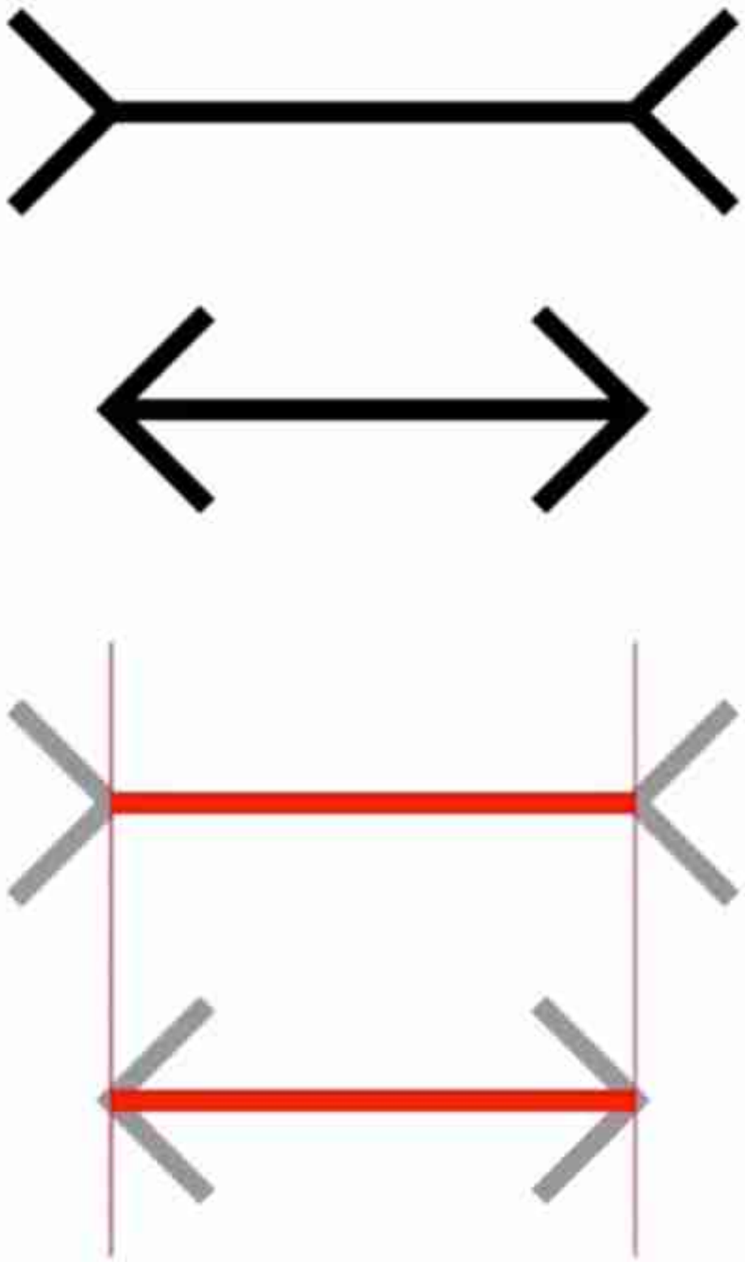
The Ebbinghaus illusion

Share

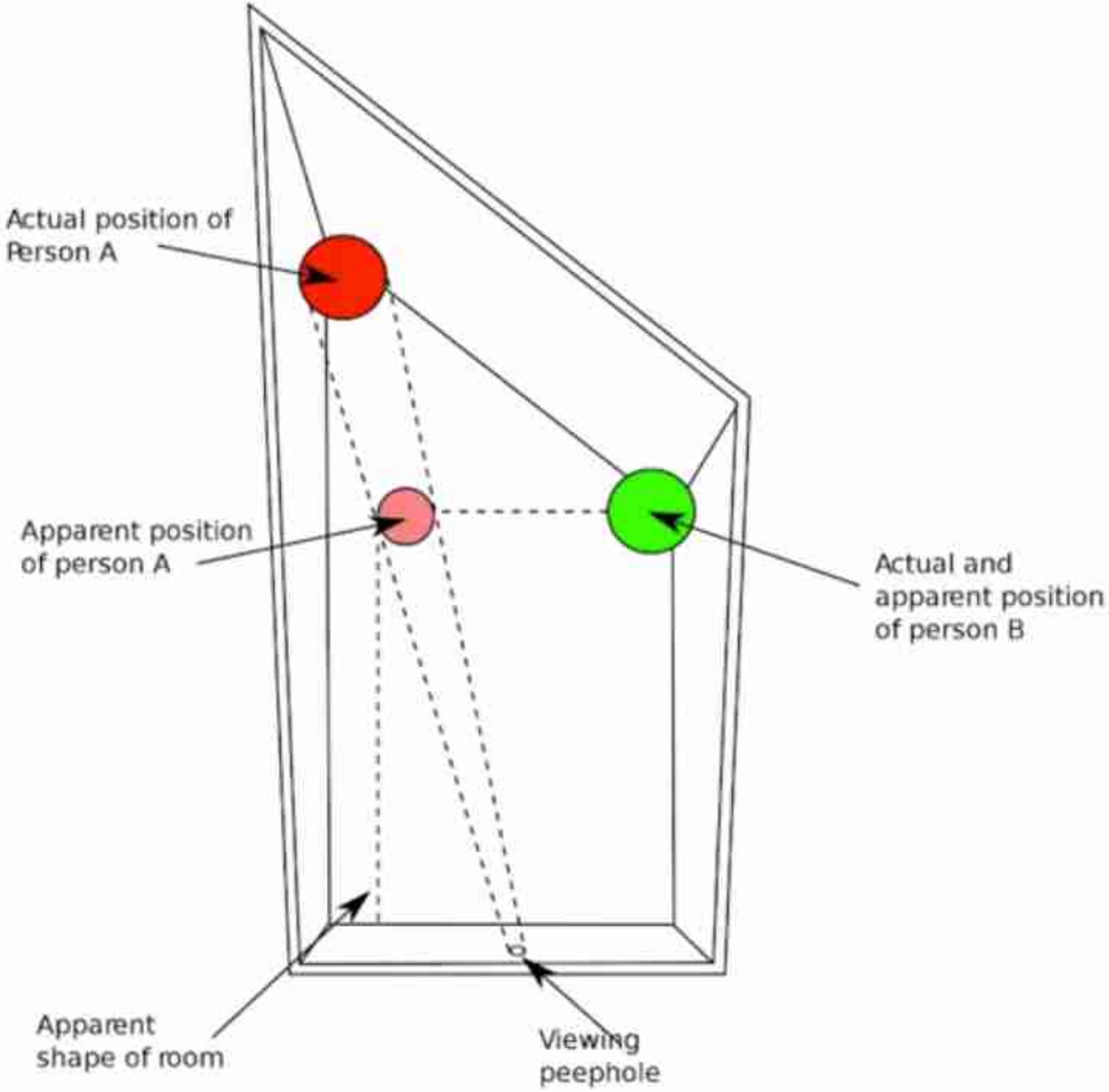
4:57 / 6:38

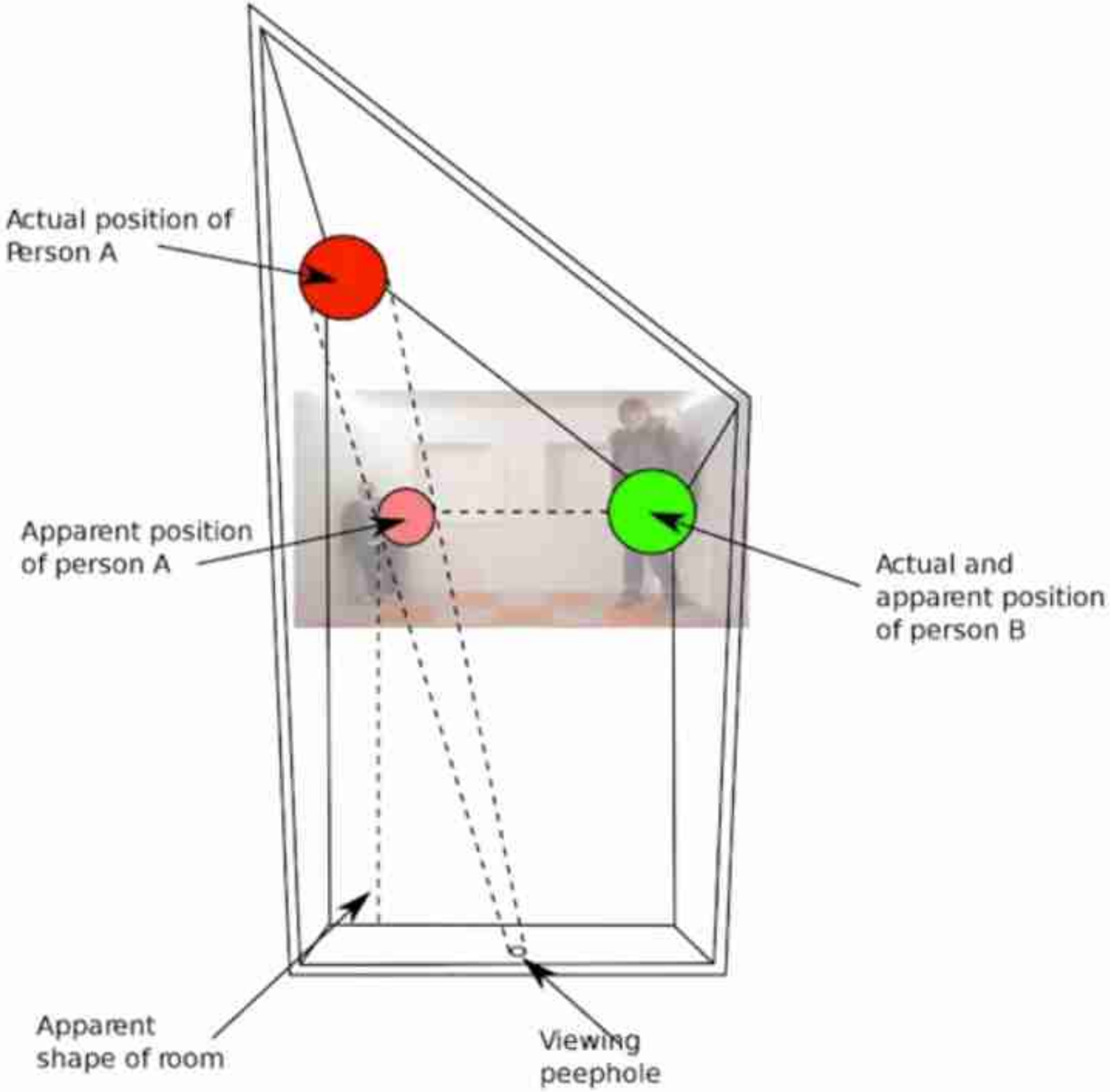


The Muller-Lyer illusion



Ames Room

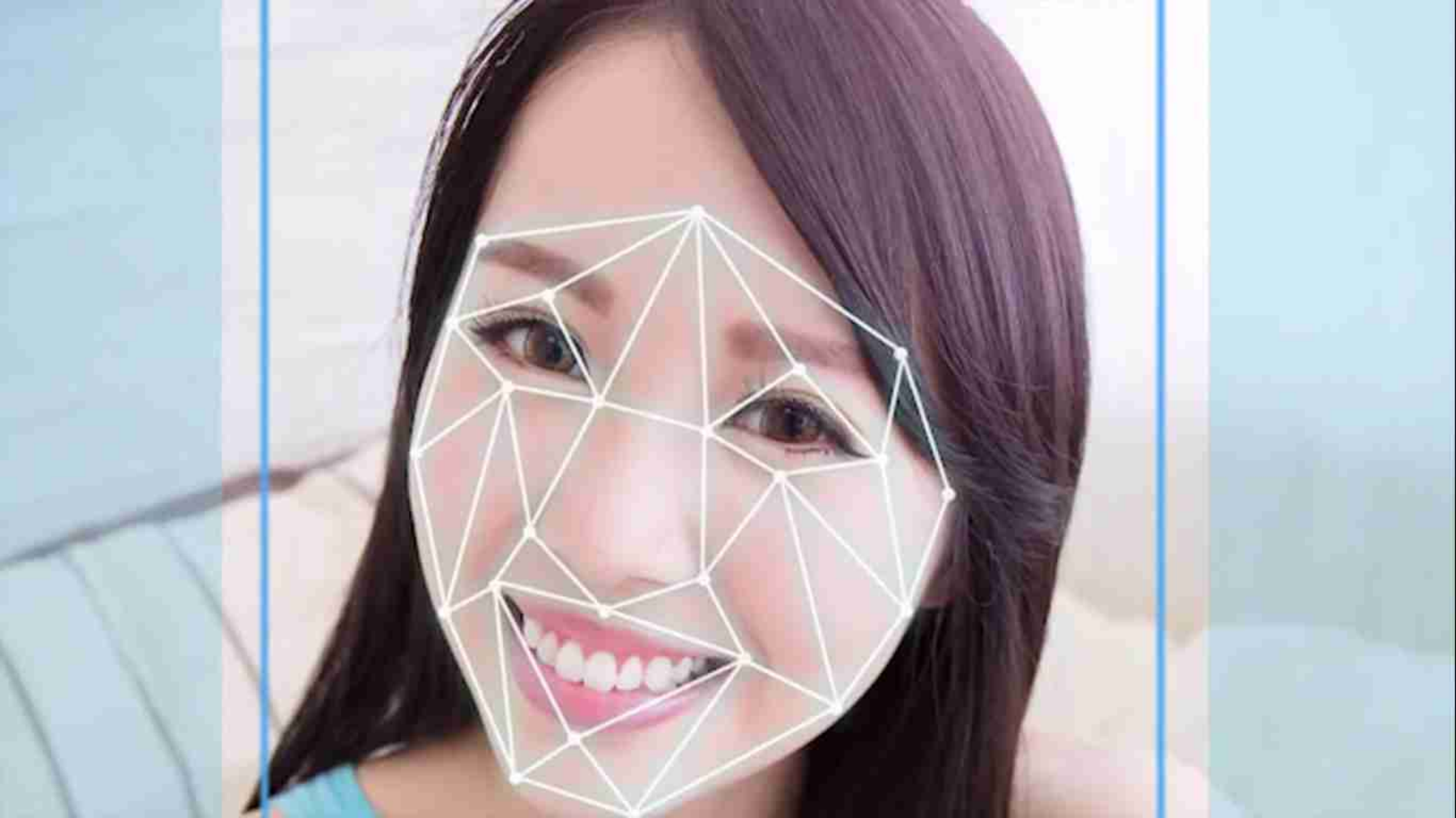

Actual position of
Person A
Apparent position
of person A
Actual and
apparent position
of person B
Apparent
shape of room
Viewing
peephole

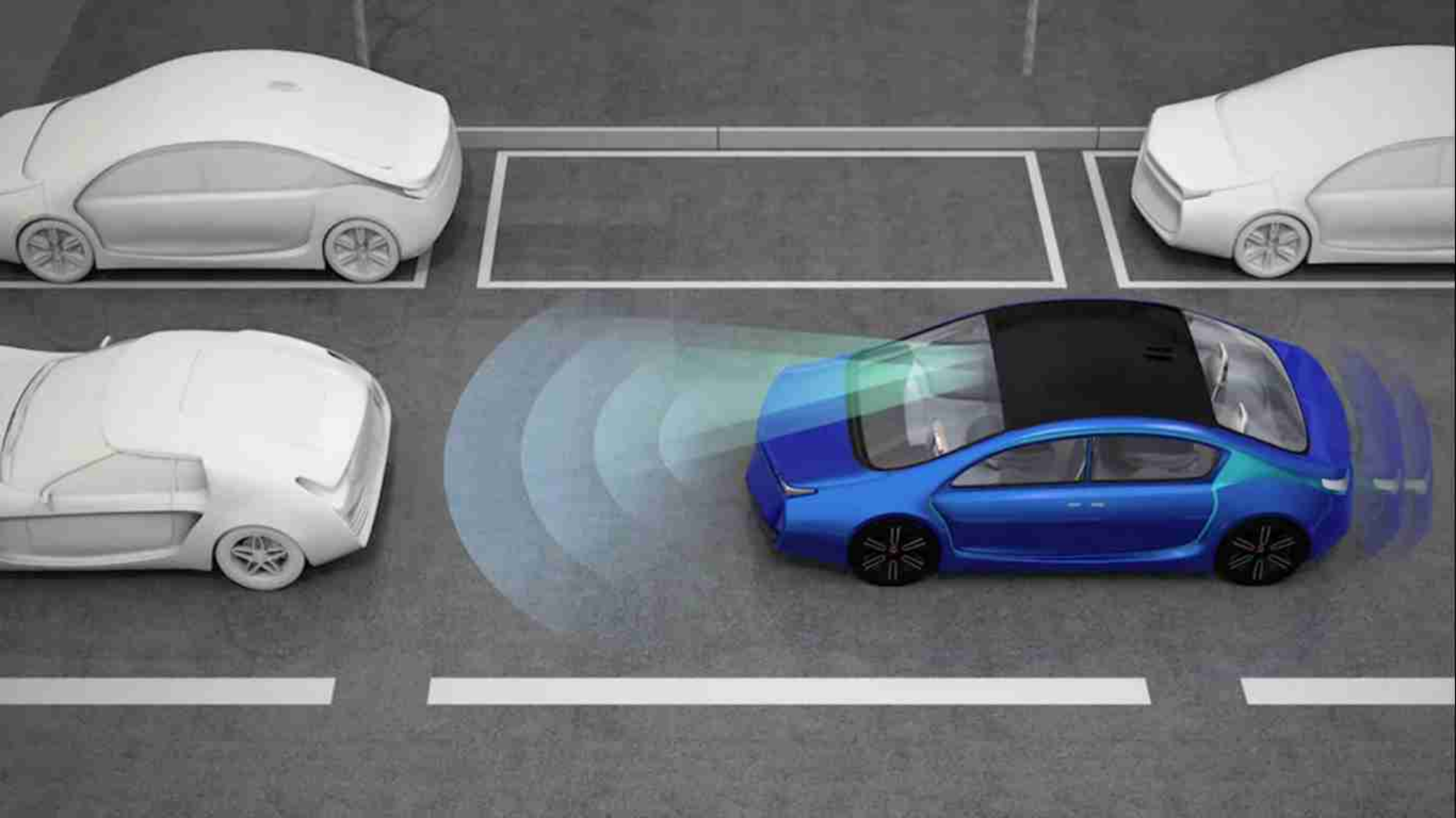

Actual position of
Person A
Apparent position
of person A
Actual and
apparent position
of person B
Apparent
shape of room
Viewing
peephole

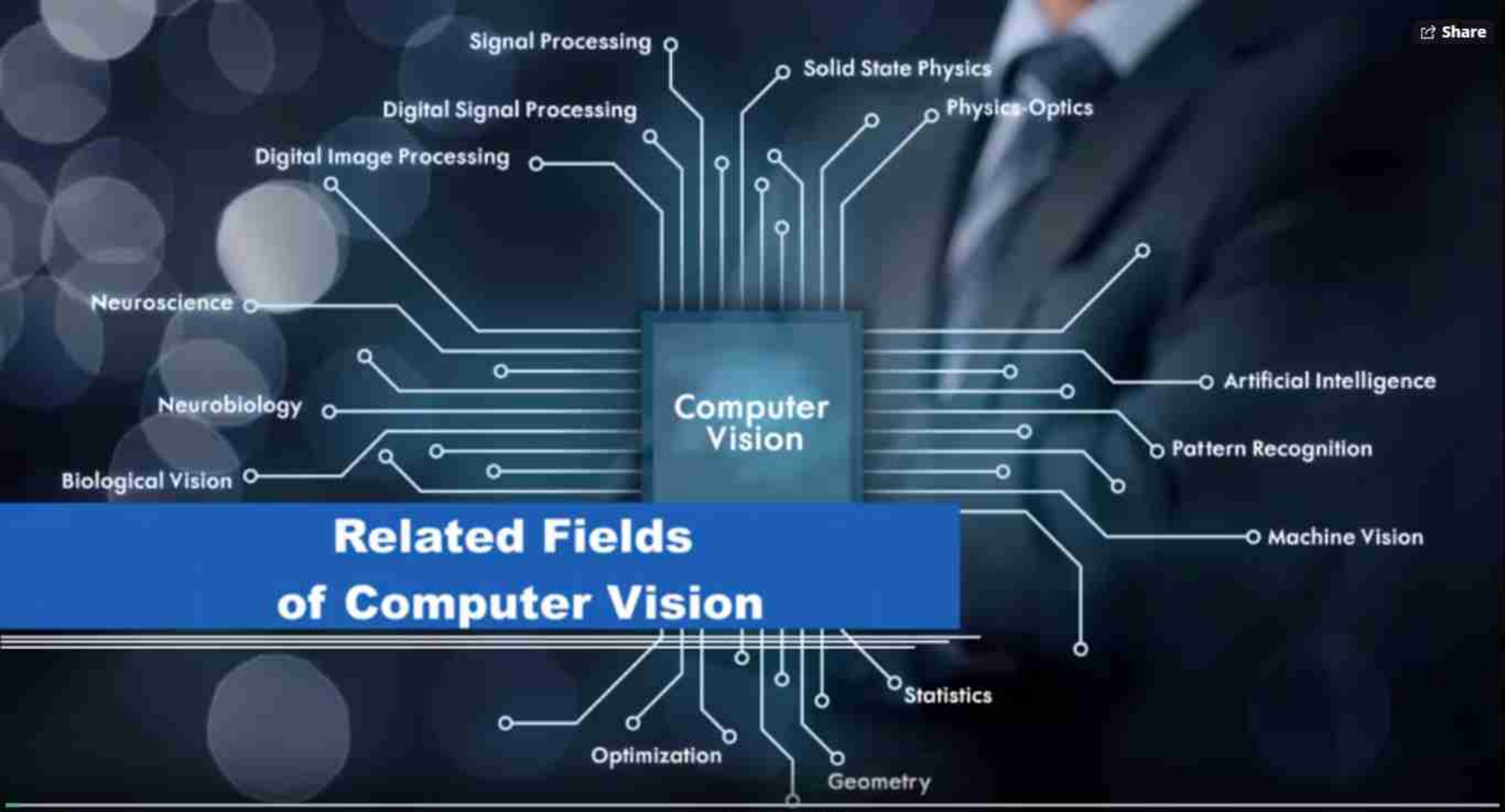
# #TheDress

Black
Blue

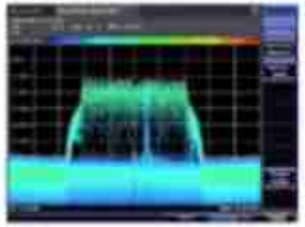
White
Gold

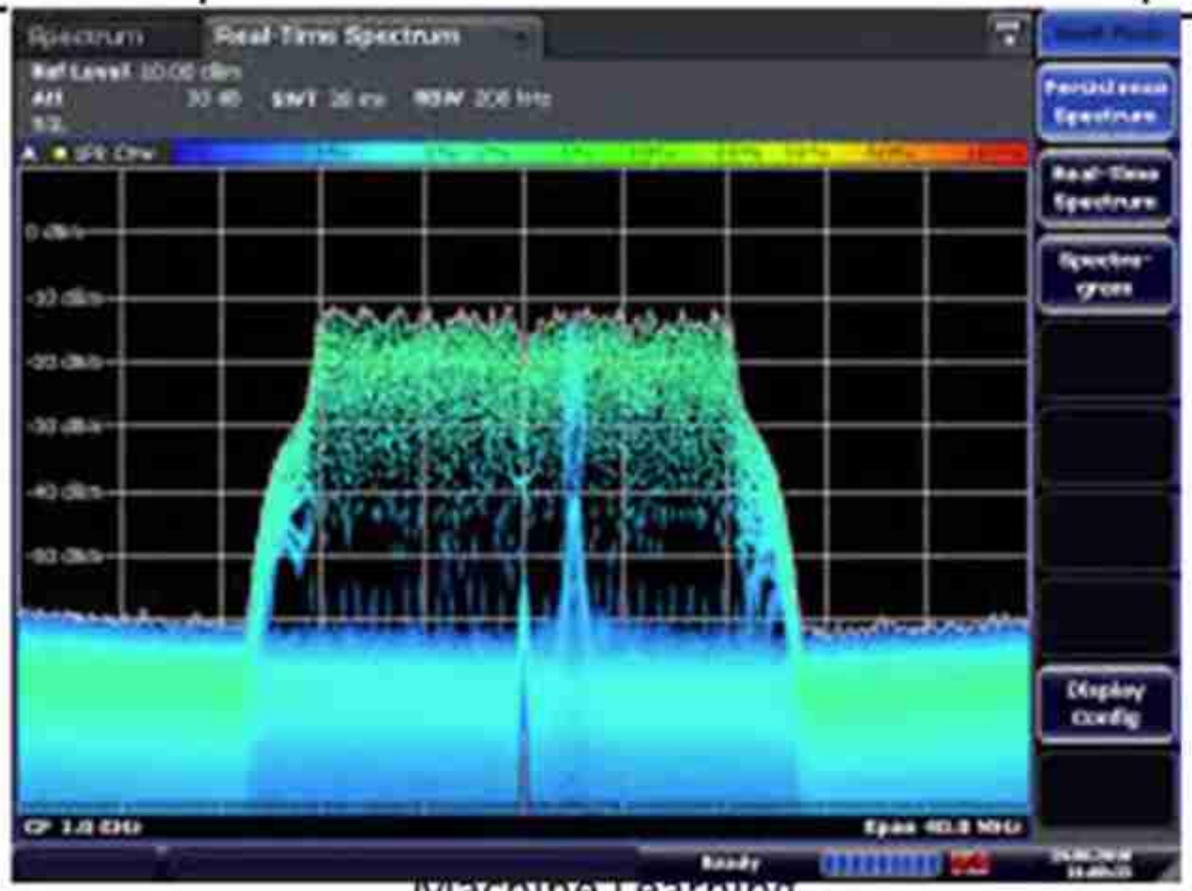





Share
Signal Processing
Solid State Physics
Digital Signal Processing
Physics-Optics
Digital Image Processing
Neuroscience
Computer Vision
Artificial Intelligence
Neurobiology
Pattern Recognition
Biological Vision
Machine Vision
Related Fields of Computer Vision
Statistics
Optimization
Geometry
Mathematics
0:02 / 4:53

Digital Signal Processing



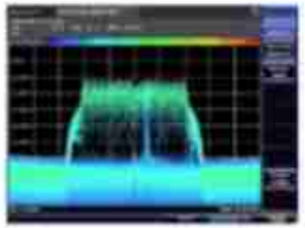



Neuroscience

Machine Learning

Artificial Intelligence

Digital Signal Processing

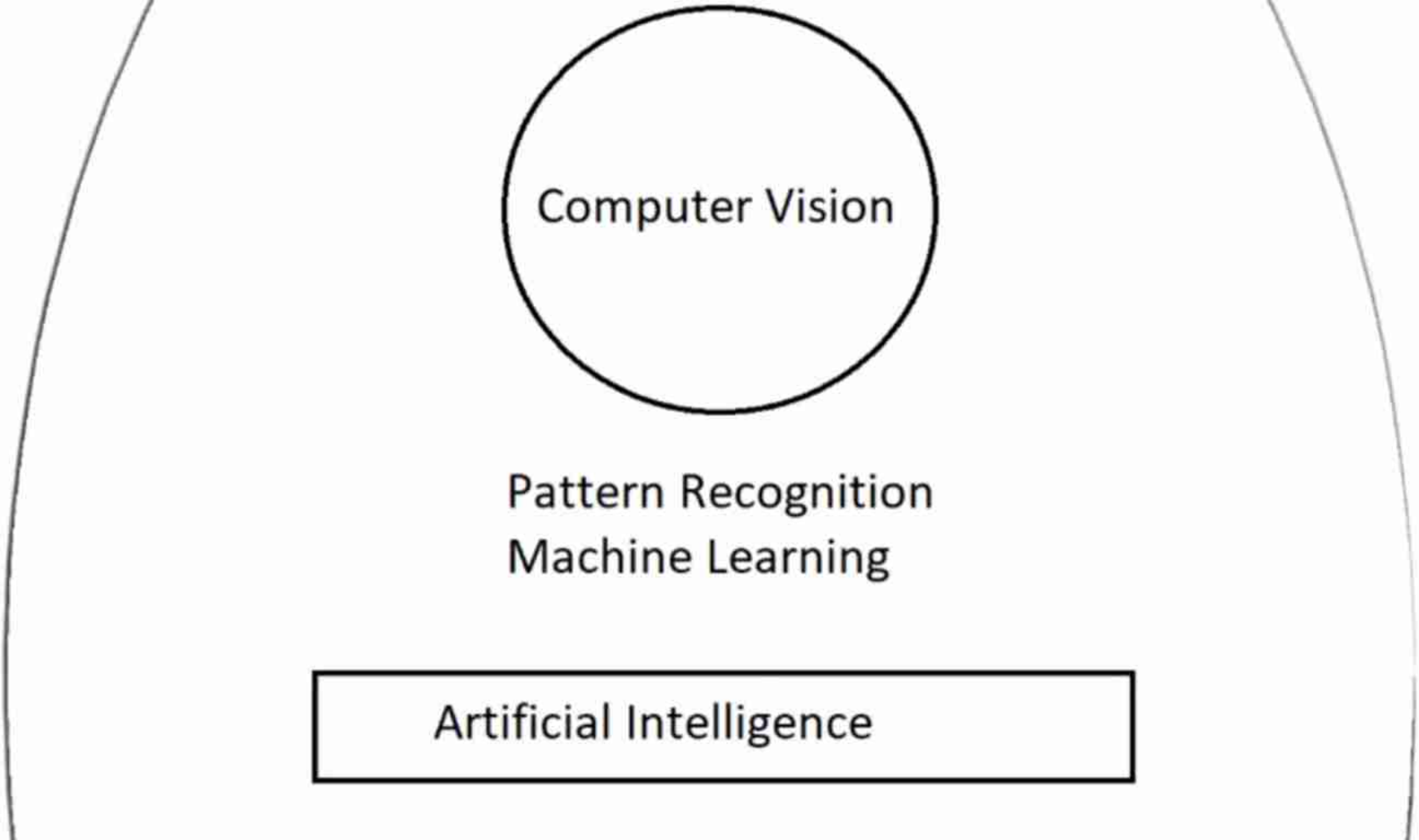


Neuroscience

Computer Vision

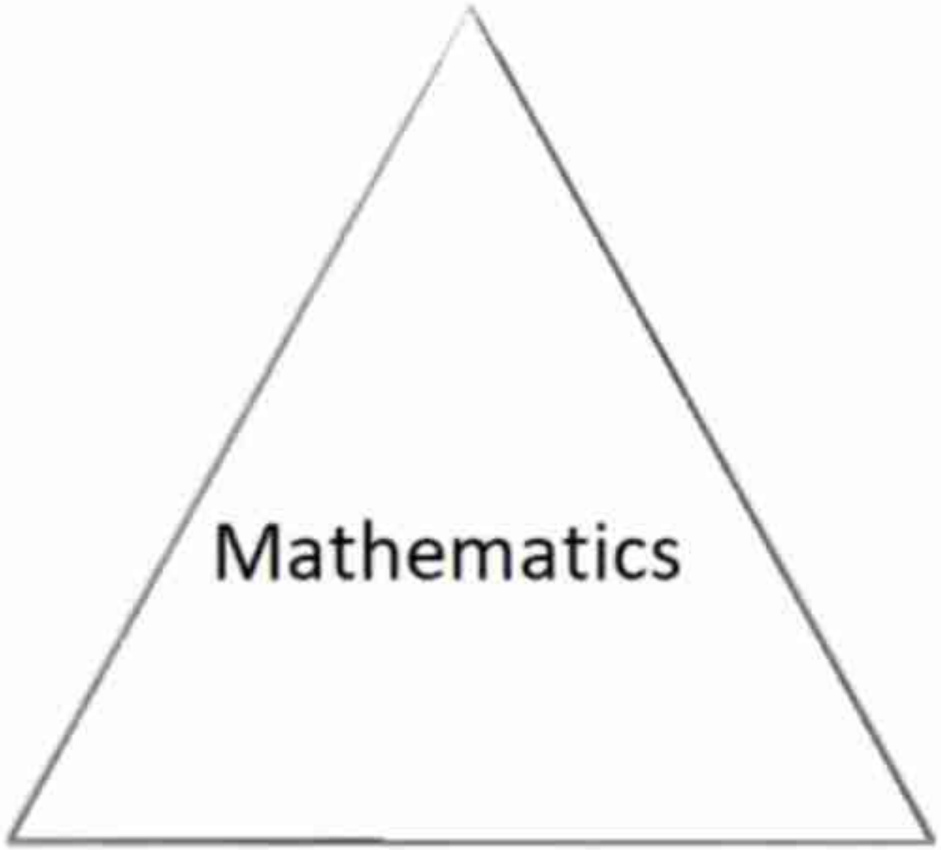
Pattern Recognition
Machine Learning

Artificial Intelligence

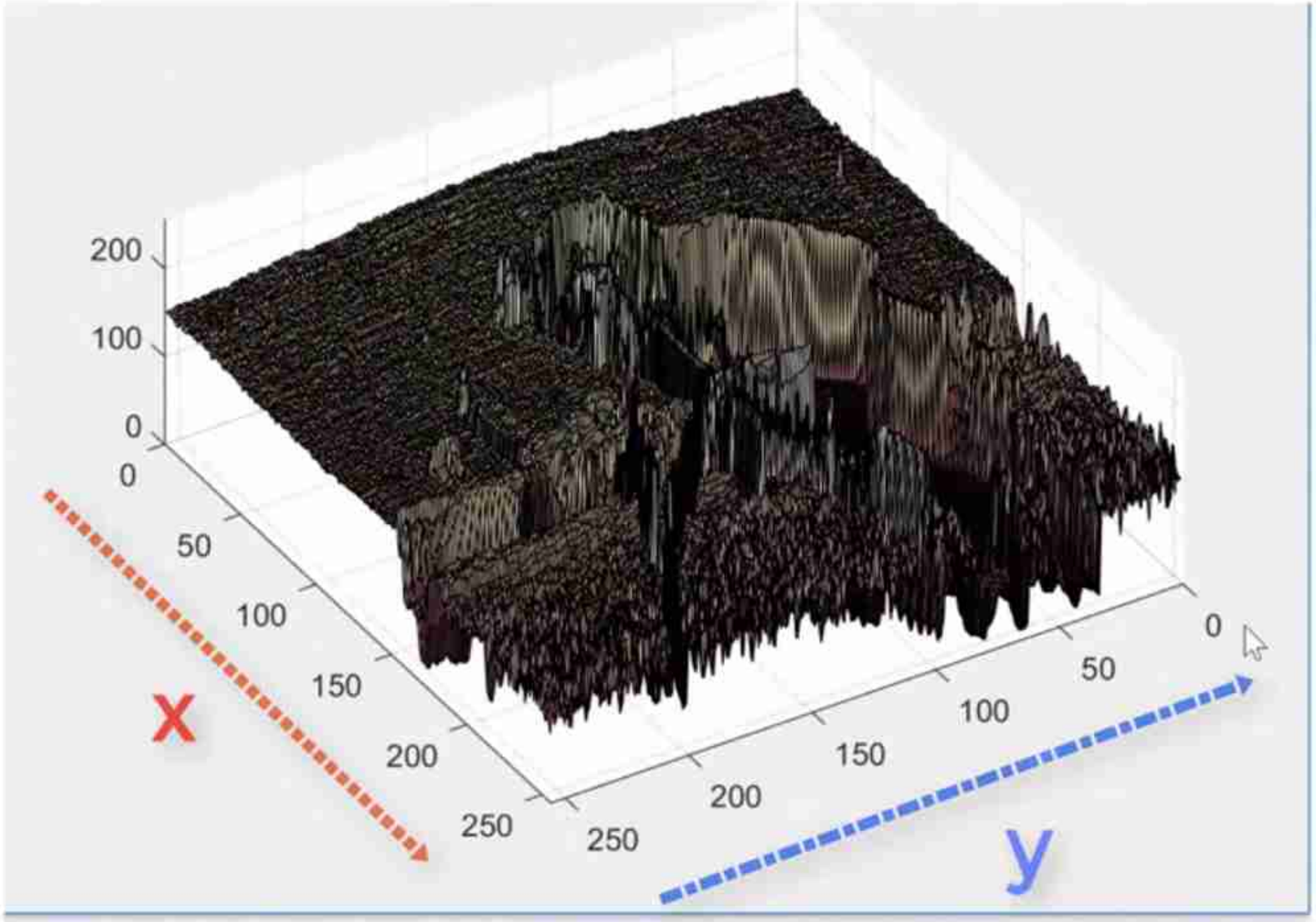

Computer Vision
Pattern Recognition
Machine Learning
Artificial Intelligence

# Computer Vision

can be studied from a purely mathematical point of view




Statistics
Optimization
Geometry


200
100
0
0
50
100
150
200
250
X
250
200
150
100
50
0
y

# Digital Image Processing

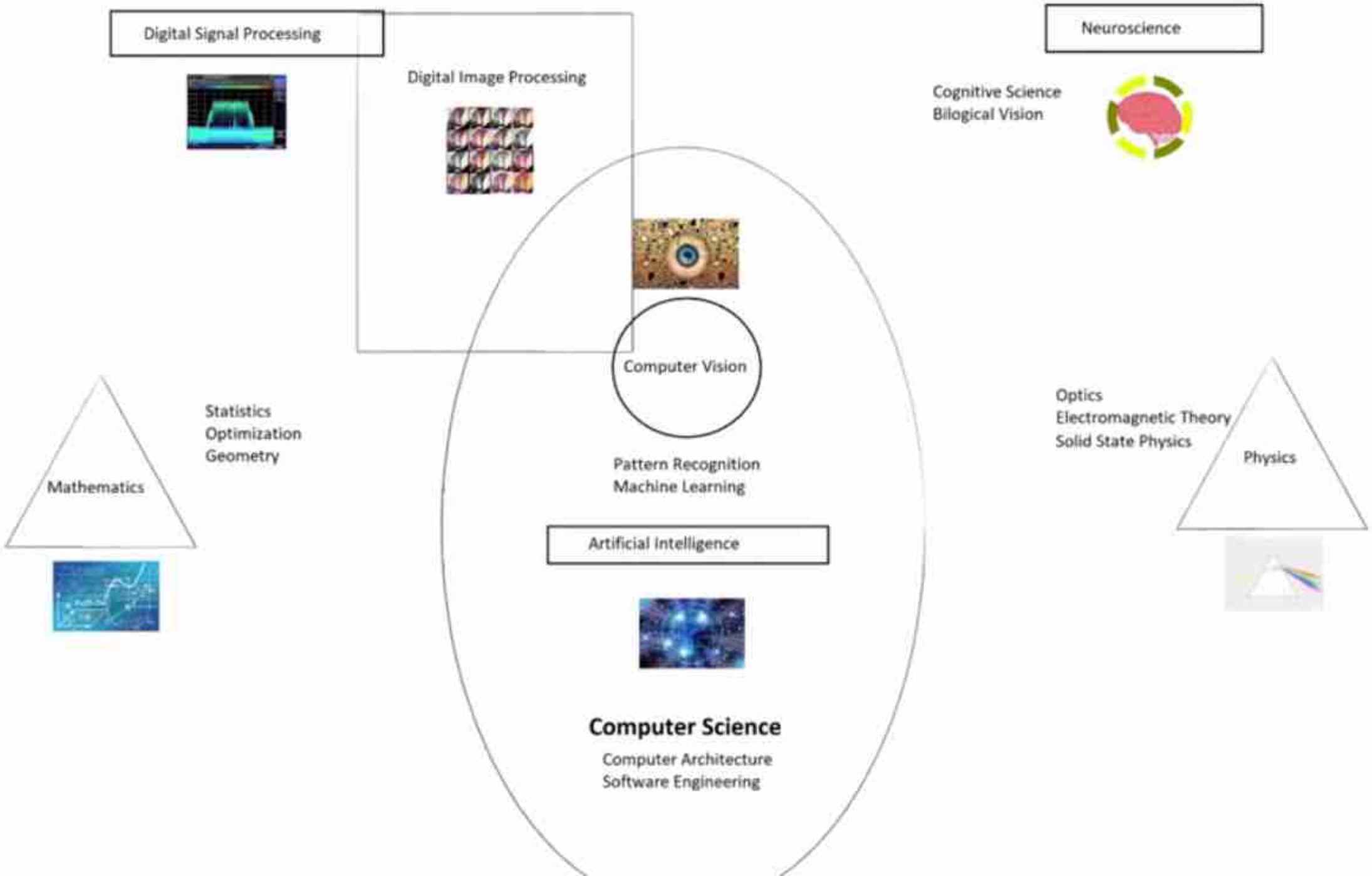
image-to-image transformation

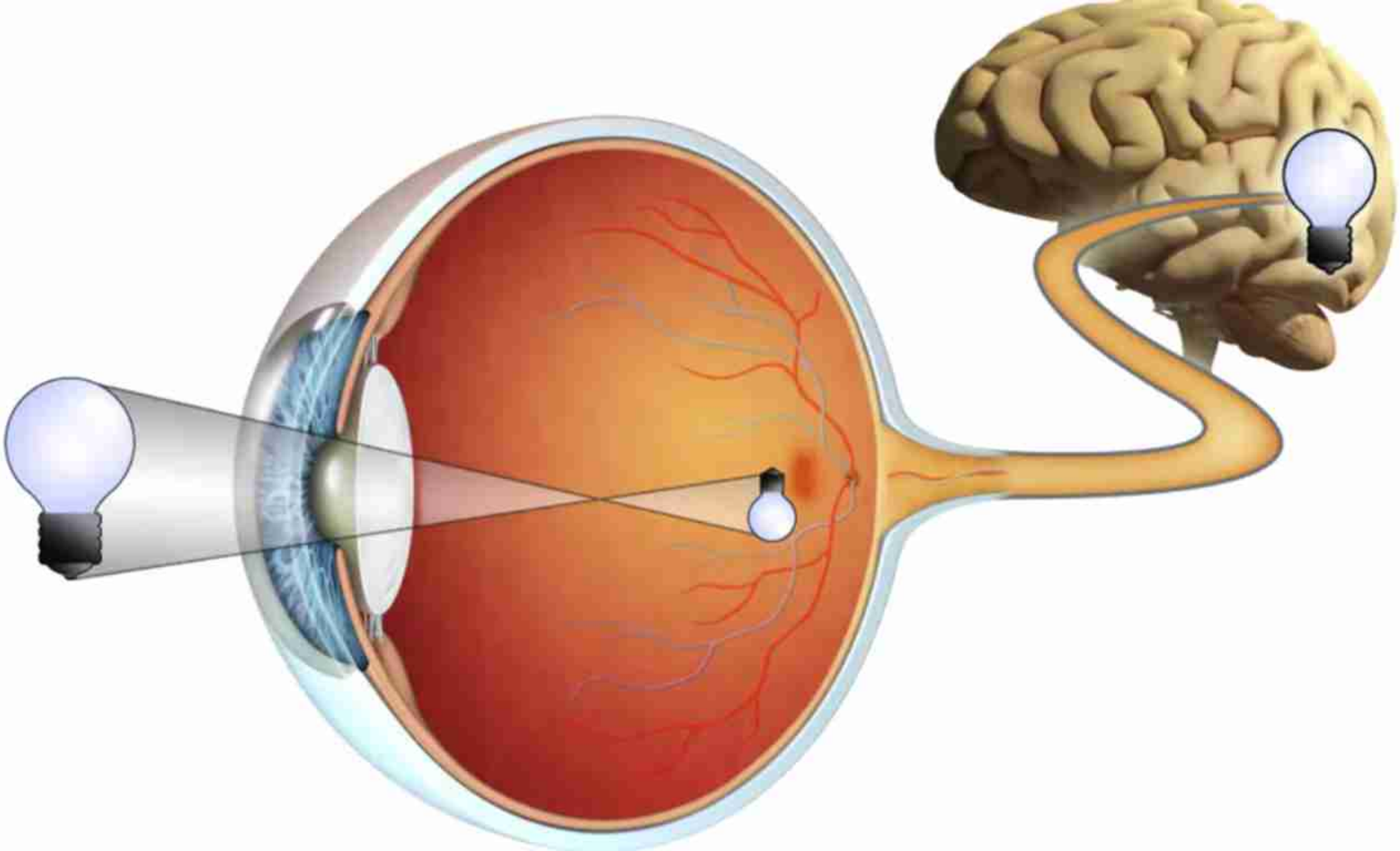
# Digital Image Processing

image compression

image restoration

image enhancement


Digital Signal Processing
Digital Image Processing
Neuroscience
Cognitive Science
Bilogical Vision
Computer Vision
Statistics
Optimization
Geometry
Mathematics
Optics
Electromagnetic Theory
Solid State Physics
Physics
Pattern Recognition
Machine Learning
Artificial Intelligence
Computer Science
Computer Architecture
Software Engineering



Share

Geometry (3D)

+

Photometry

Computer Graphics

Images (2D)

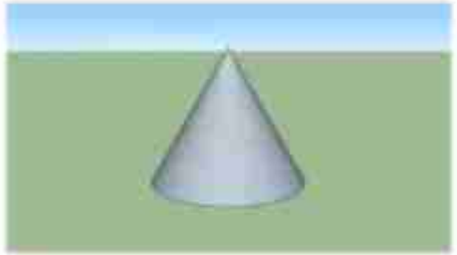
3:28 / 4:53

Geometry (3D)

Computer Vision

Images (2D)

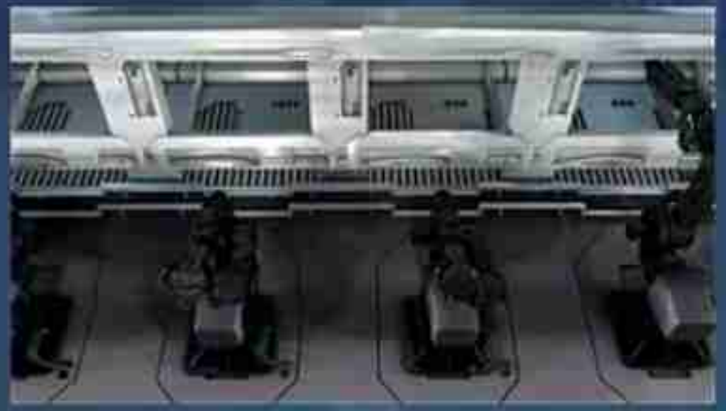
Photometry

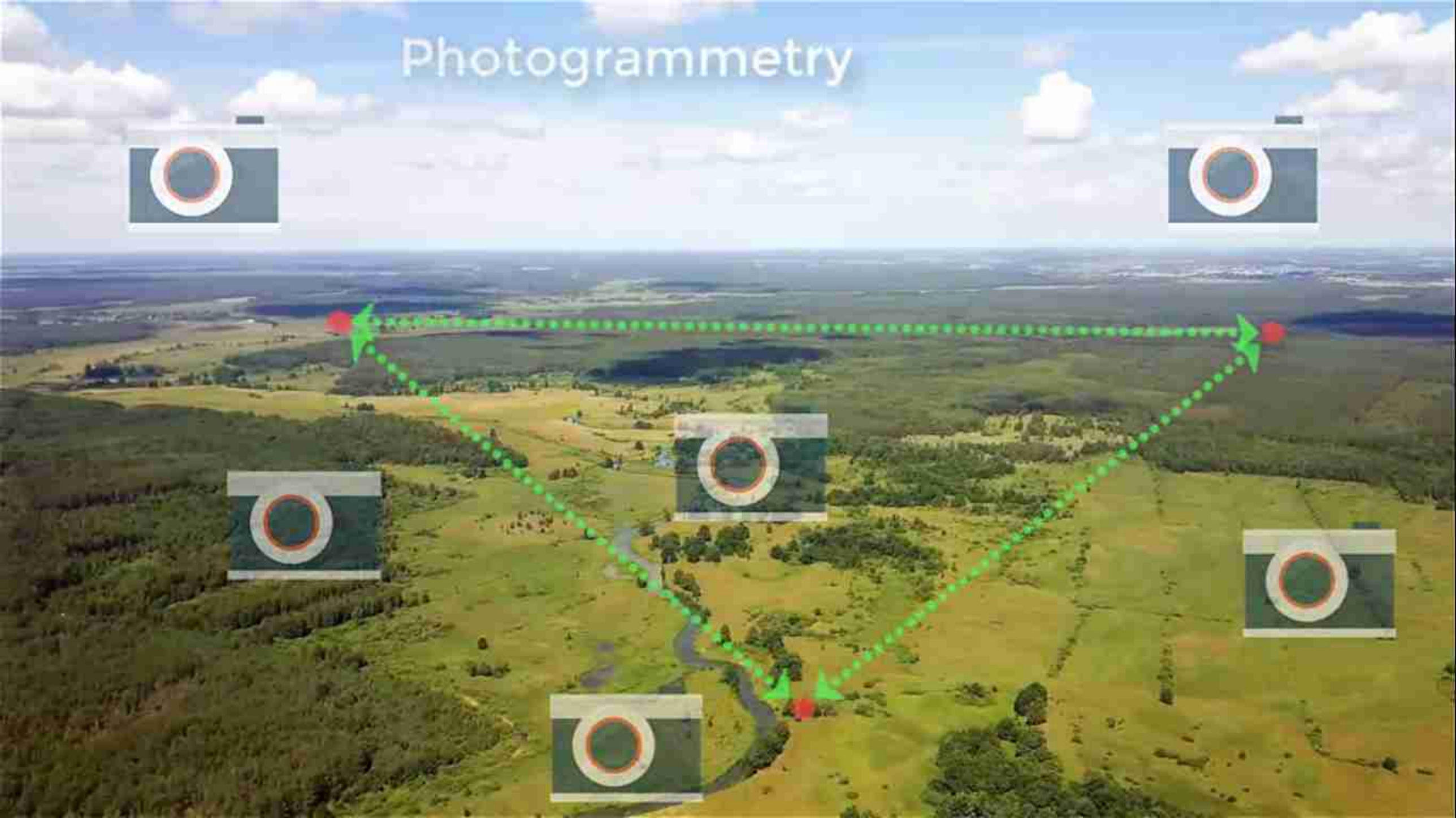


# Machine Vision

applying a range of technologies and methods

imaging-based automatic inspection




Photogrammetry

# Three Distinct Lines

replicating the eye

replicating the visual cortex

replicating the rest of the brain

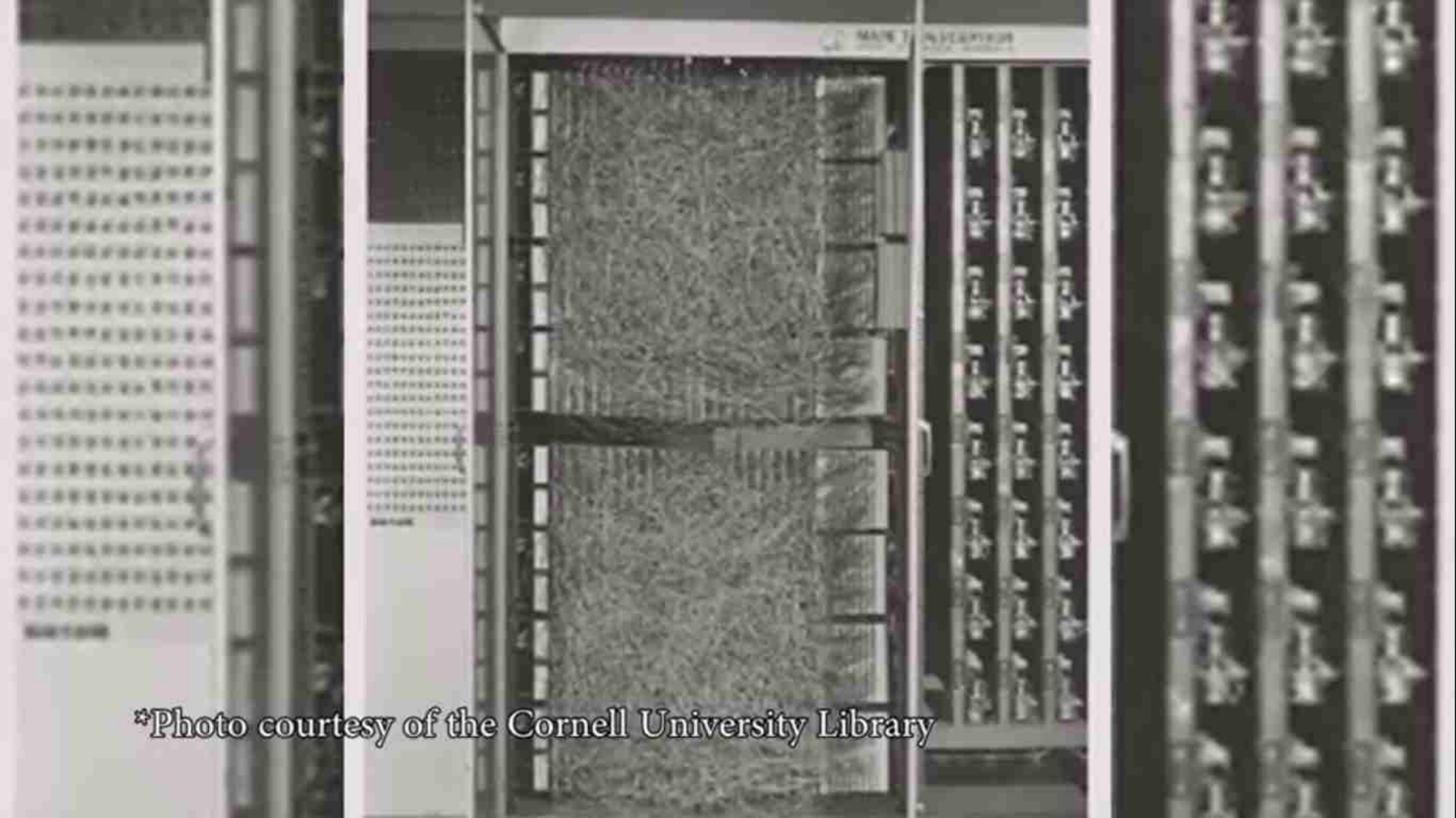
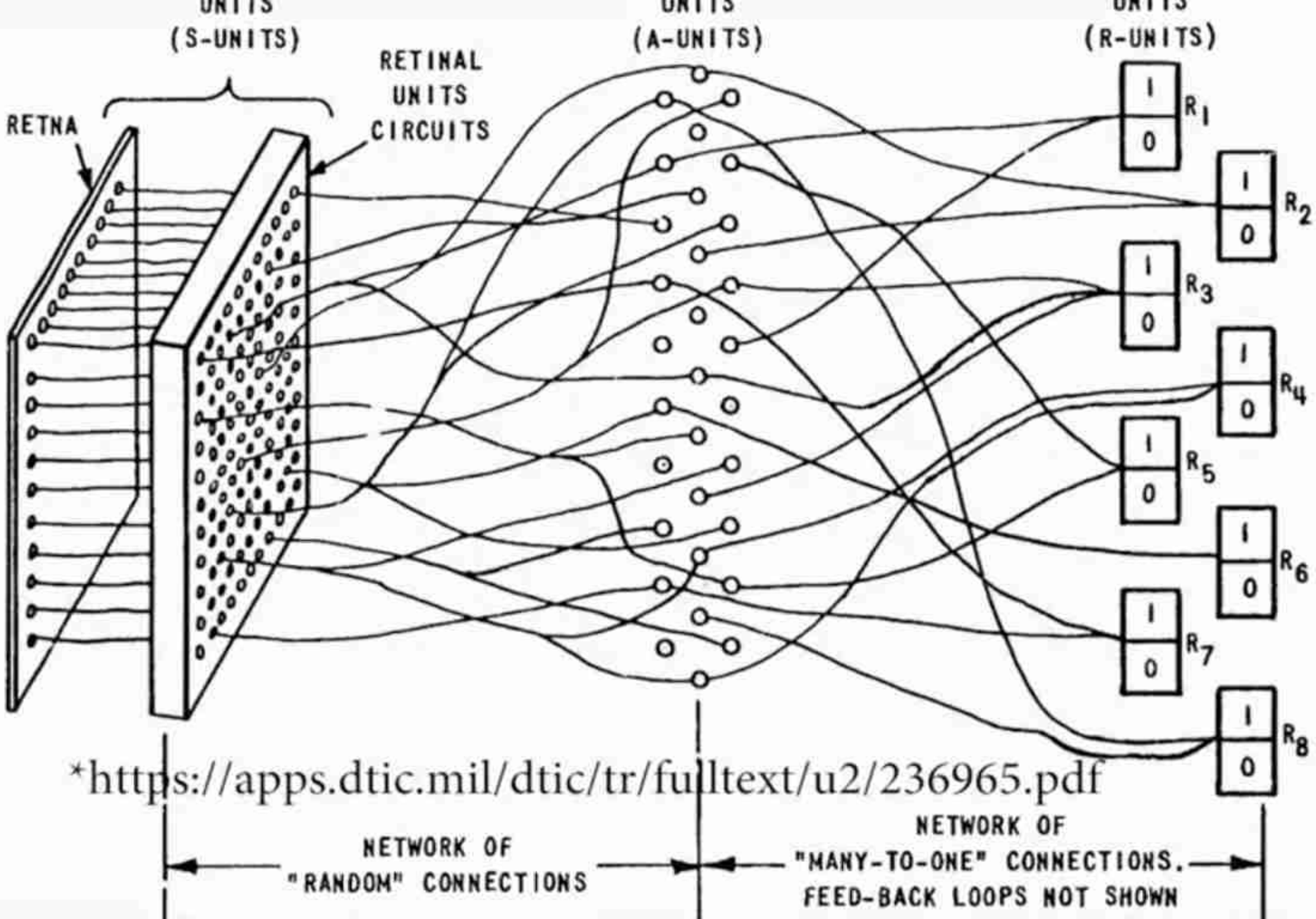


*Photo courtesy of the Cornell University Library




*https://apps.dtic.mil/dtic/tr/fulltext/u2/236965.pdf

# Recognizes Objects

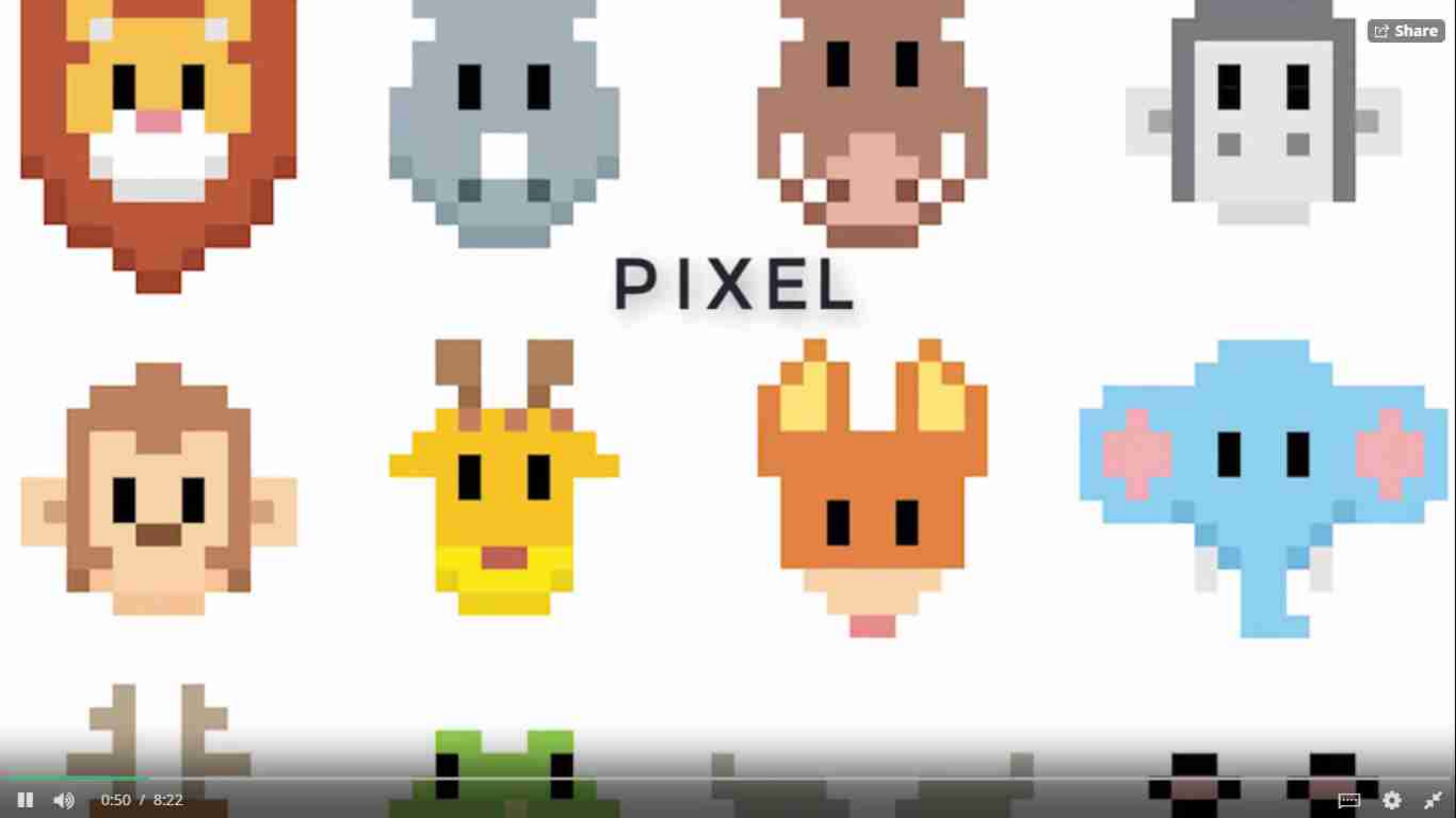
*By Scripps-Howard Service*

WASHINGTON, July 7—The first non-living mechanism able to "perceive, recognize and identify its surroundings without human training or control" has been developed for the U. S. Navy.

The mechanical brain's human characteristics could include recognizing faces, reading, translating, writing and talking.

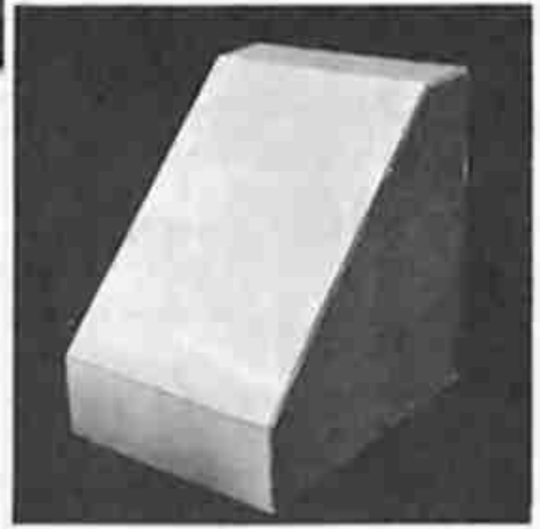
Rosenblatt, research psychologist, Cornell Aeronautical Laboratory, in Buffalo, N. Y.

The concept of Perceptron was demonstrated on a conventional IBM 704 computer,

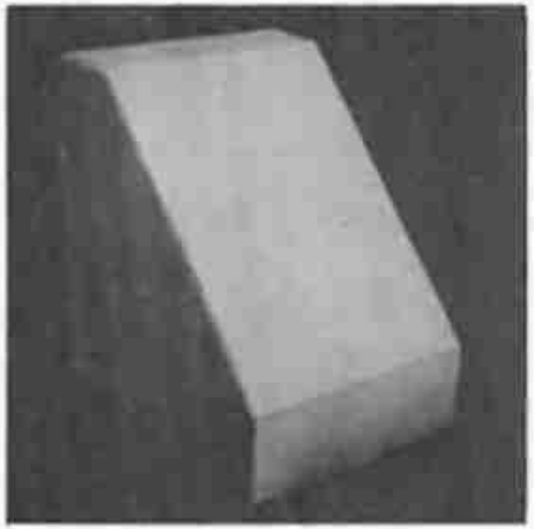

Share
PIXEL
0:50 / 8:22

*http://wafflesatnoon.com/first-digital-image/

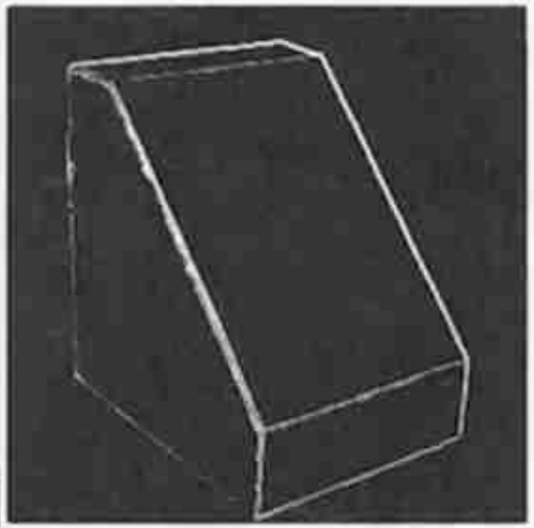
ure.



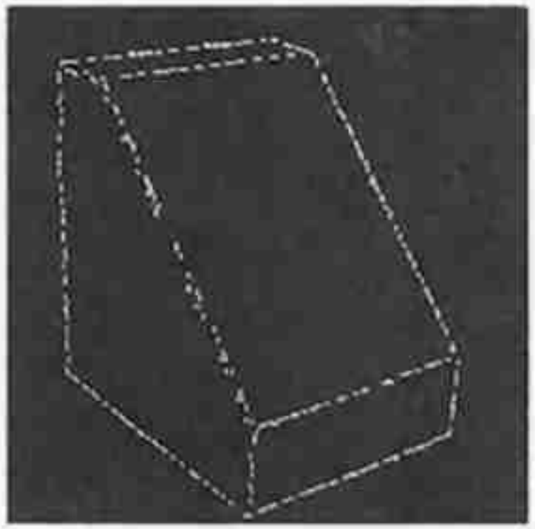
(a) Original picture.



(b) Computer display of picture (reflected by mistake).



(c) Differentiated picture.



(d) Feature points selected.

(b) Compu
(reflected

# MIT Profs Give Computer Eyes, Hands

By VICTOR K. McELHENY
Globe Science Editor

In Cambridge, there is a device that can recognize blocks of various sizes and shapes and put them together into towers or stacks — without step-by-step instruction from a human operator.

The device links a mechanical arm, a television camera and a computer. It was built as part of a long-term research program at Massachusetts Institute of Technology aimed at giving machines more "artificial intelligence" than they have now.

The program is directed by Prof. Marvin Minsky of M.I.T.'s electrical engineering department and visiting Prof. Seymour Papert of the mathematics department.

Minsky thinks a major problem with computers is to increase their ability "to deal w th mixtures of different kinds of information including general principles of problem solving."

The idea of linking the computer, the mechanical arm and the television camera, Minsky indicated, is to see if machines doing routine work can become more flexible.

With some simple kind of vision, the sort a television camera could provide, a machine in a factory might be able to hunt around for an object, notice that something was wrong with it, or pick it up when it had been dropped.

The "eye" of the M.I.T. device is a television camera which has been equipped for what engineers call "image dissection." This is different from the usual blank stare of the television camera.

Under computer control, the "image dissector" camera can be told to focus on any given part of a scene.

The arm used in the M.I.T. device is a standard industrial arm for handling dangerous materials, such as radioactive chemicals. For more complicated jobs later, the Minsky group has developed an arm with a shoulder and three elbows that can reach around corners.

Minsky and Papert direct a so-called Artificial Intelligence Group as part of M.I.T.'s large computer effort, Project MAC.

AIR CONDITIONING ANALYSIS
(No Obligation)
T. E. SNOWDEN, INC.
9 Green Street
Lynn, Mass.
599-7770
See TRANE Ad — Page 55

LIVE • WORK • PLAY

The Boston Globe, Apr 24, 1968

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

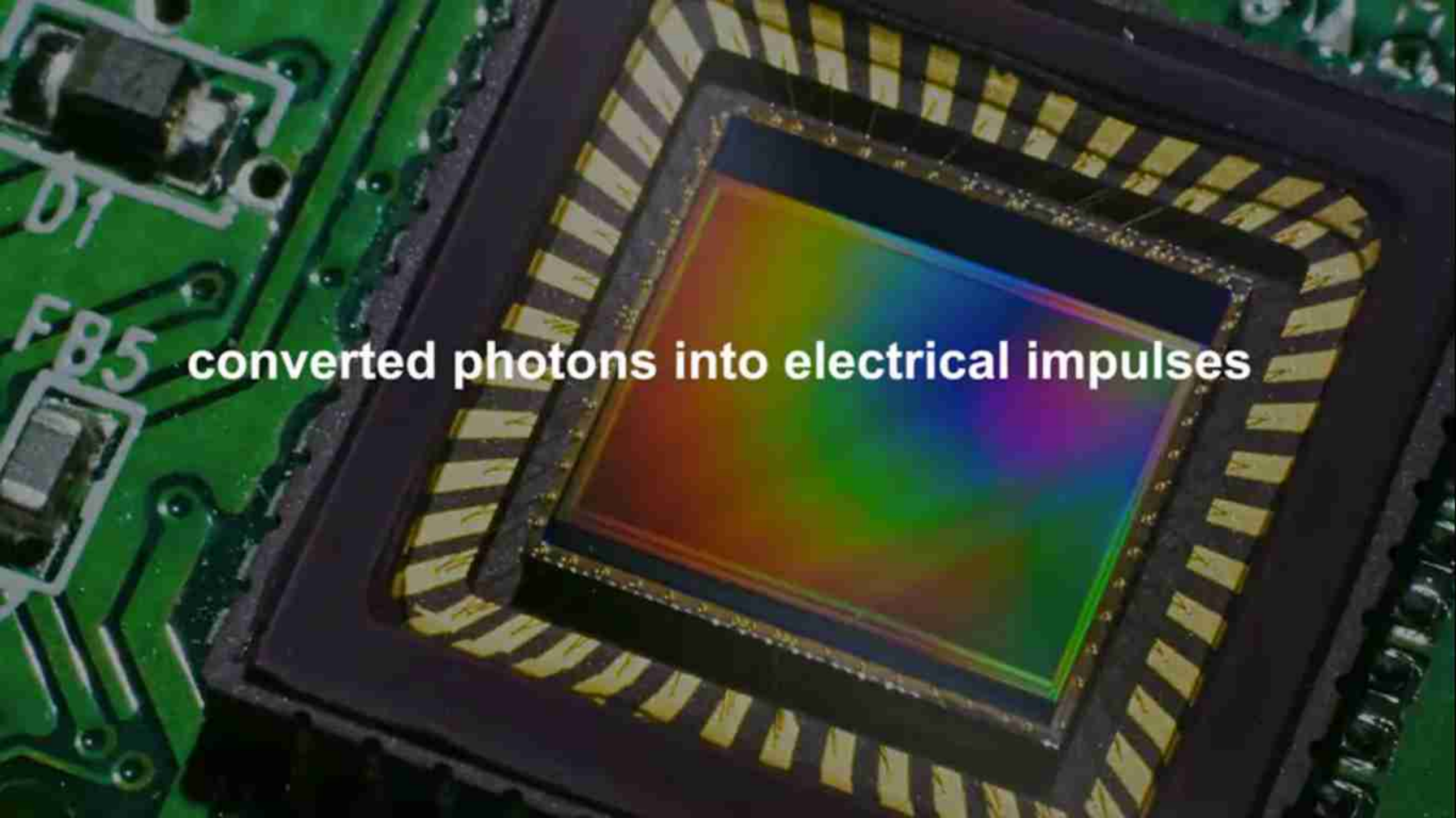
Artificial Intelligence Group
Vision Memo. No. 100.

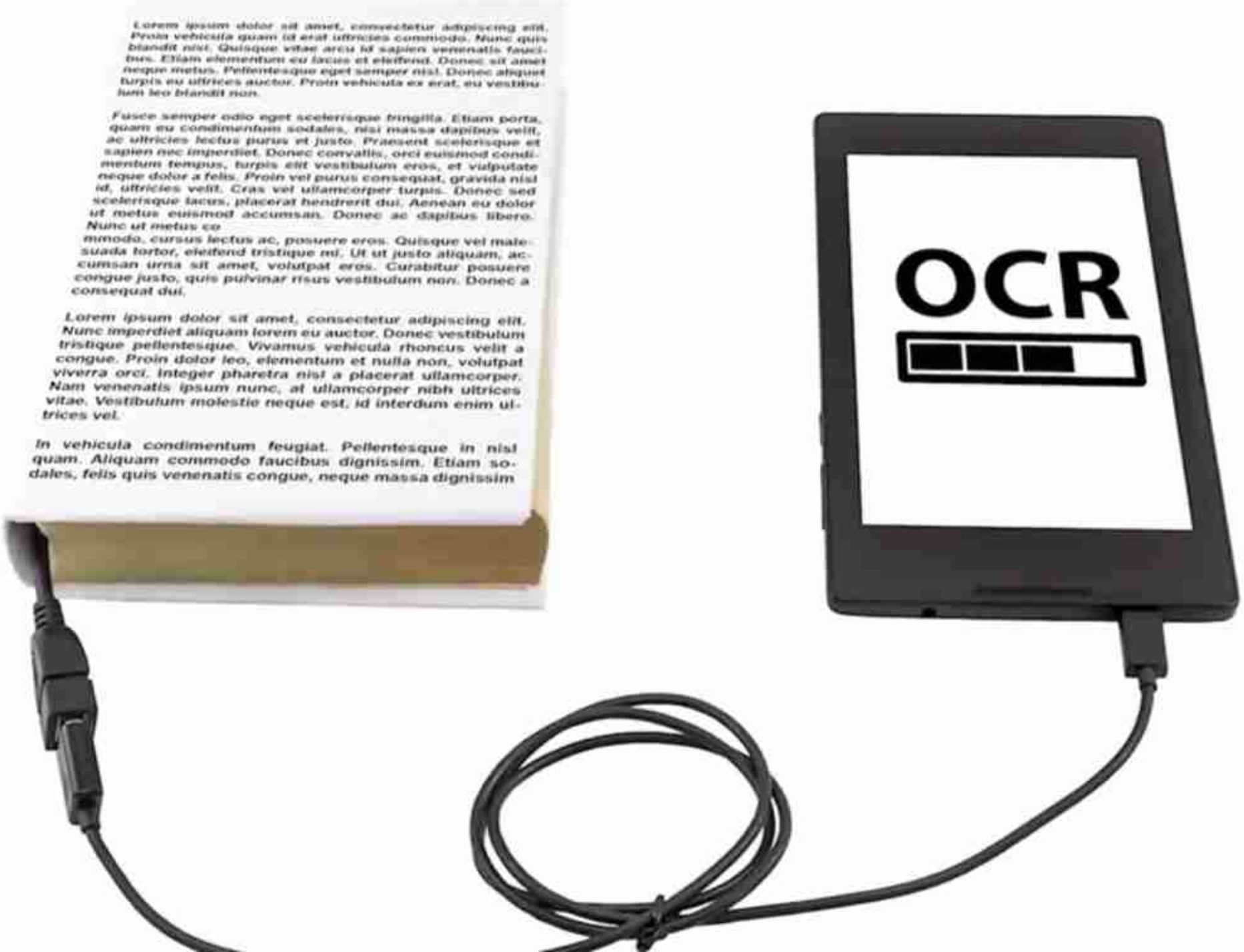
July 7, 1966

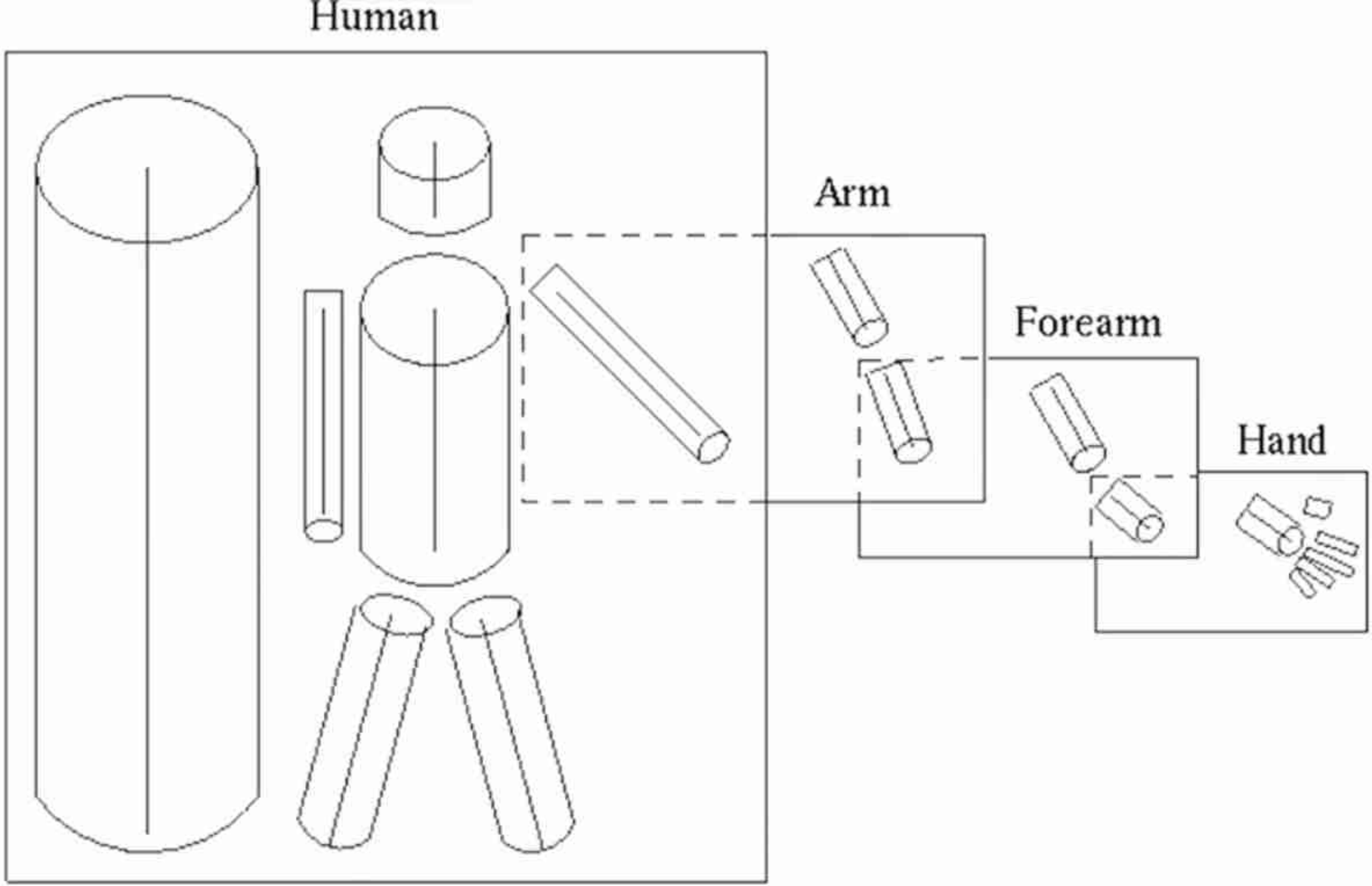
# THE SUMMER VISION PROJECT

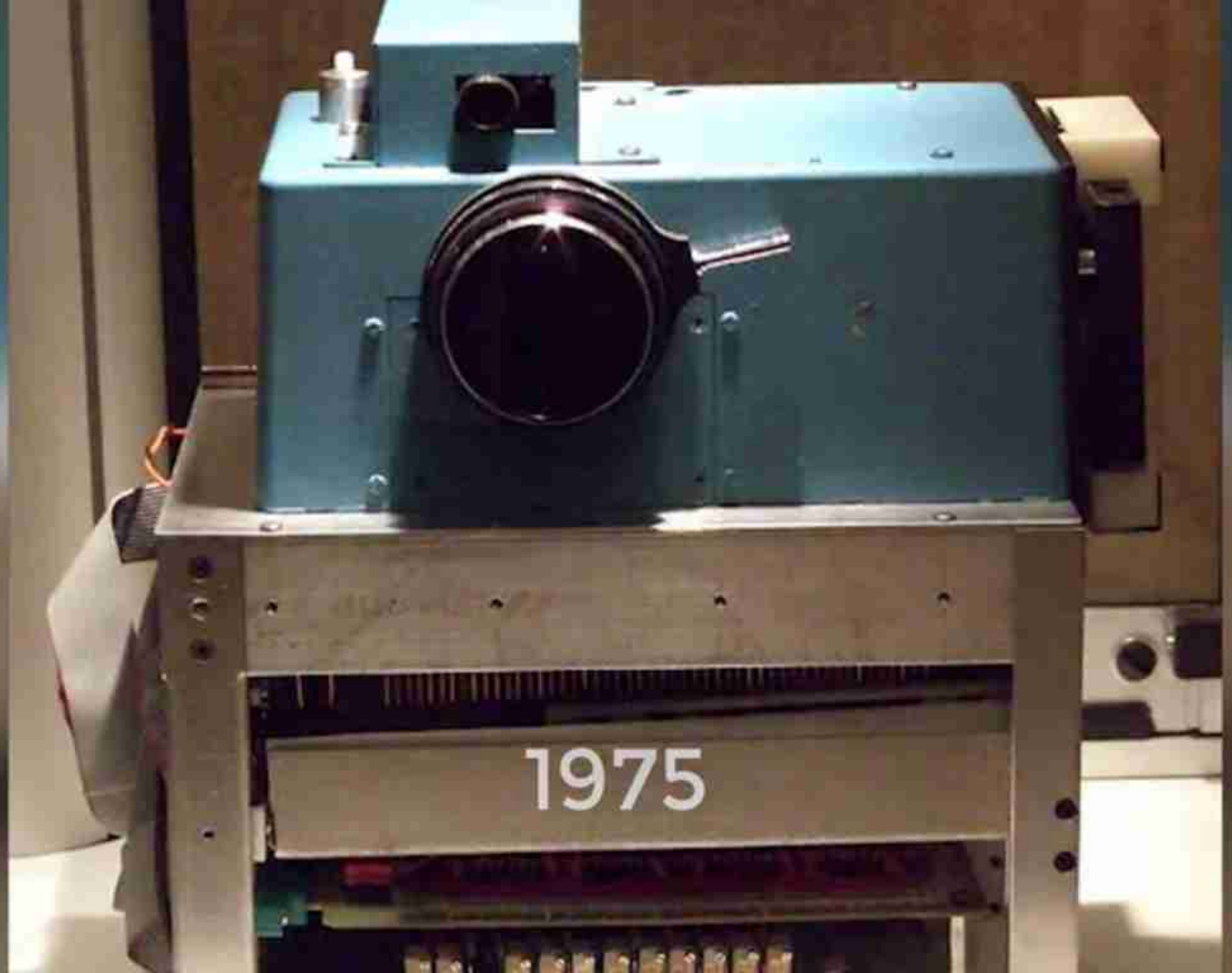
Seymour Papert

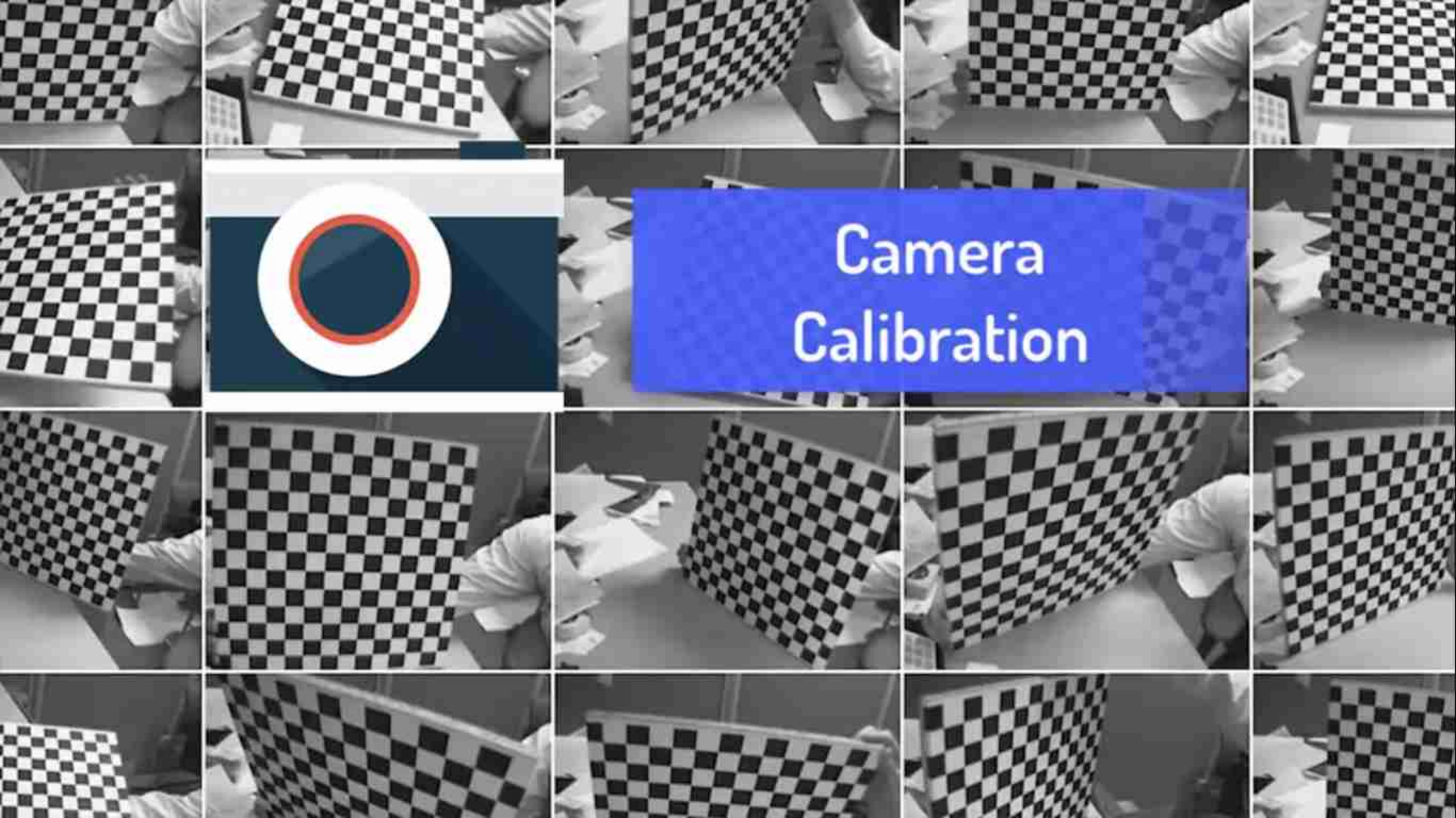
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into


converted photons into electrical impulses

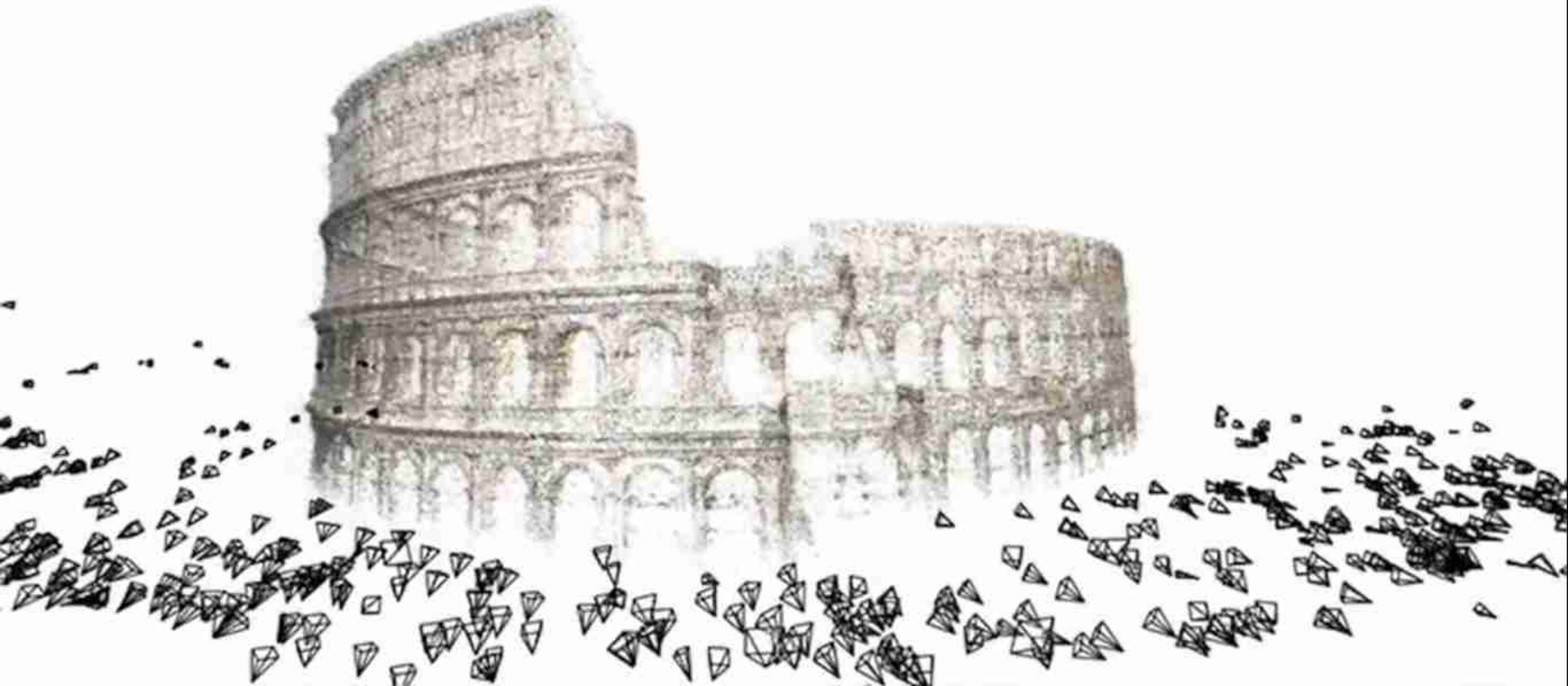

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Proin vehicula quam id erat ultrices commodo. Nunc quis blandit nisl. Quisque vitae arcu id sapien venenatis faucibus. Etiam elementum eu lacus et eleifend. Donec sit amet neque metus. Pellentesque eget semper nisl. Donec aliquet turpis eu ultrices auctor. Proin vehicula ex erat, eu vestibulum leo blandit non.
Fusce semper odio eget scelerisque tringilla. Etiam porta, quam eu condimentum sodales, nisi massa dapibus velit, ac ultricies lectus purus et justo. Praesent scelerisque et sapien nec imperdiet. Donec convallis, orci euismod condimentum tempus, turpis elit vestibulum eros, et vulputate neque dolor a felis. Proin vel purus consequat, gravida nisl id, ultricies velit. Cras vel ullamcorper turpis. Donec sed scelerisque lacus, placerat hendrerit dui. Aenean eu dolor ut metus euismod accumsan. Donec ac dapibus libero. Nunc ut metus co
mmodo, cursus lectus ac, posuere eros. Quisque vel malesuada tortor, eleifend tristique mi. Ut ut justo aliquam, accumsan urna sit amet, volutpat eros. Curabitur posuere congue justo, quis pulvinar risus vestibulum non. Donec a consequat dui.
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc imperdiet aliquam lorem eu auctor. Donec vestibulum tristique pellentesque. Vivamus vehicula rhoncus velit a congue. Proin dolor leo, elementum et nulla non, volutpat viverra orci. Integer pharetra nisl a placerat ullamcorper. Nam venenatis ipsum nunc, at ullamcorper nibh ultrices vitae. Vestibulum molestie neque est, id interdum enim ultrices vel.
In vehicula condimentum feugiat. Pellentesque in nisl quam. Aliquam commodo faucibus dignissim. Etiam sodales, felis quis venenatis congue, neque massa dignissim
OCR


Human
Arm
Forearm
Hand

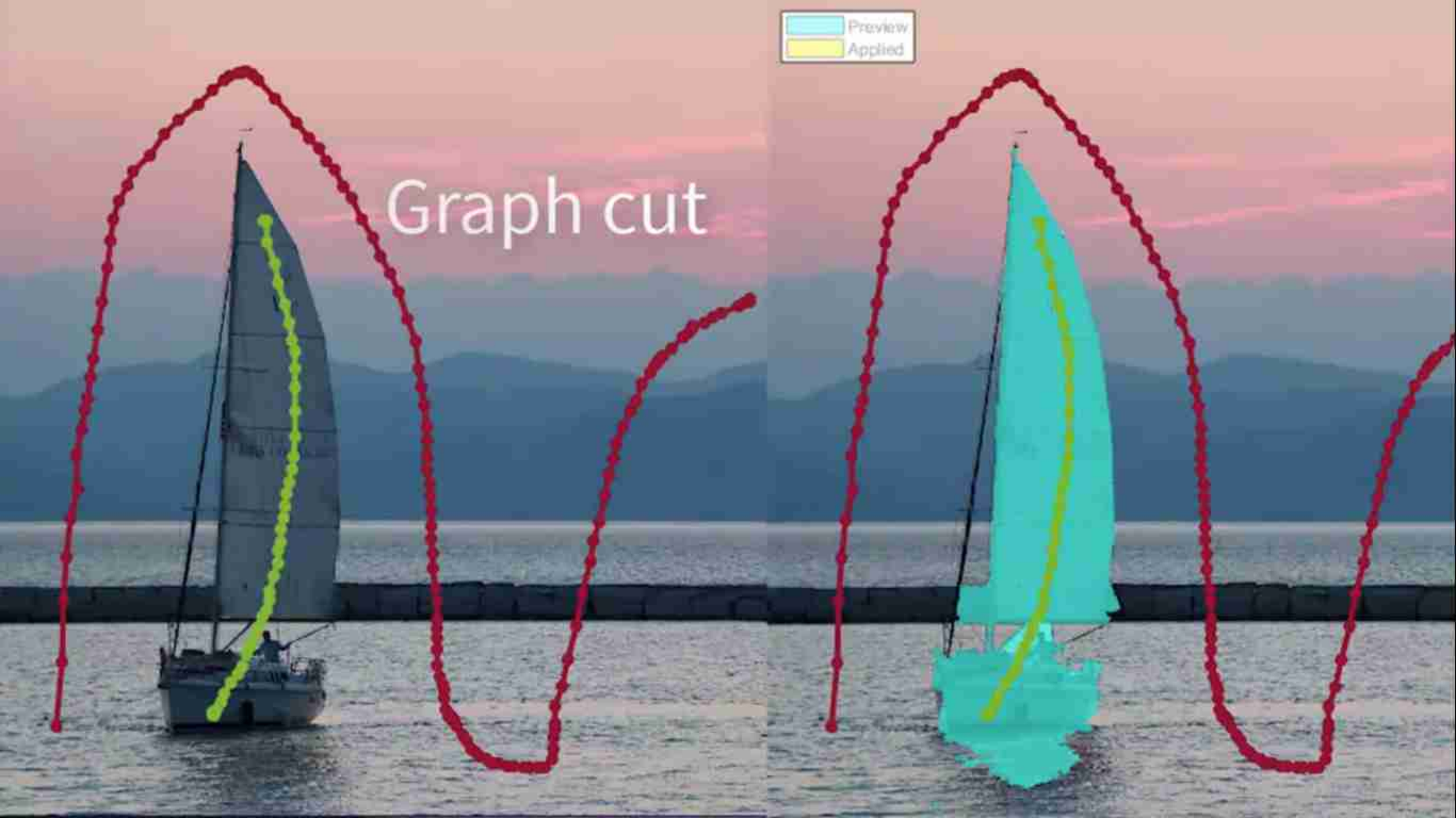

1975


Camera
Calibration

Courtesy: Sameer Agarwal, Building Rome in a Day!



Kodak STEREO Camera


Graph cut
Preview
Applied

# Computer Graphics and Computer Vision

image-based rendering

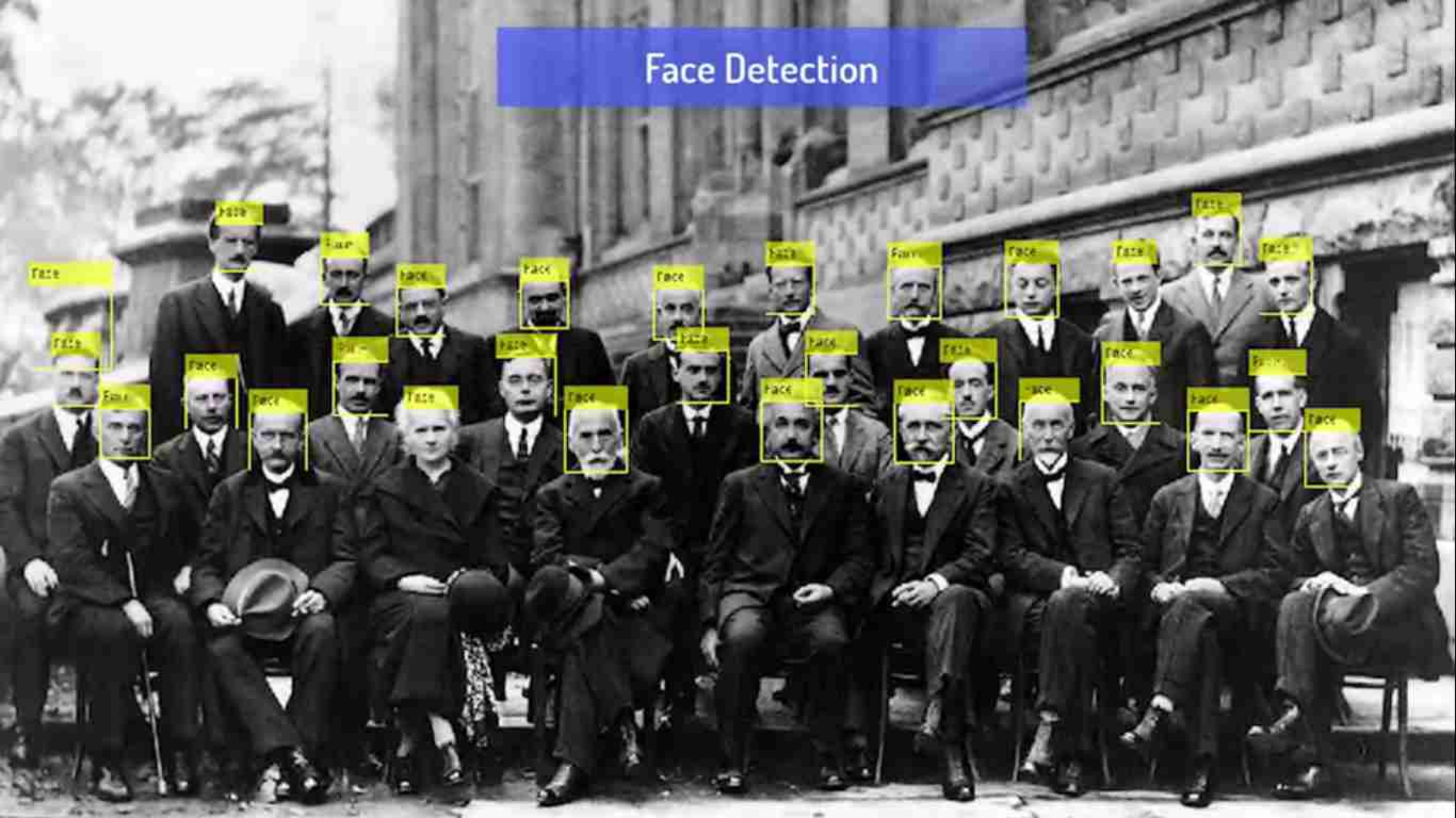
image morphing

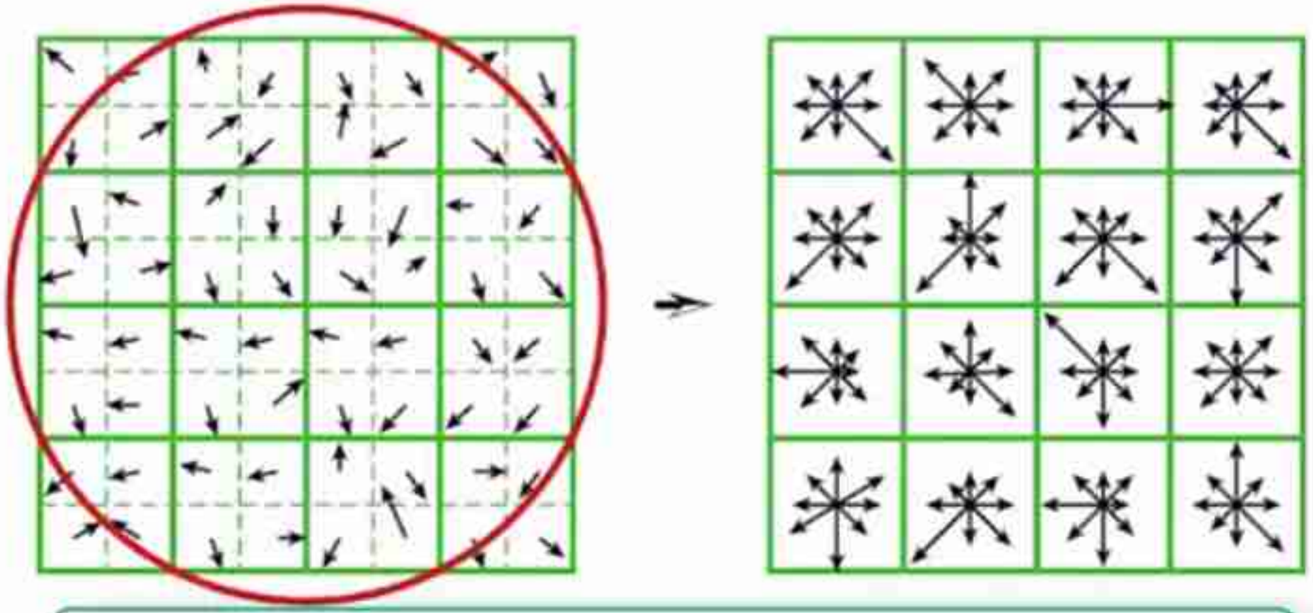
view interpolation

panoramic image stitching

early light-field rendering

# Face Detection

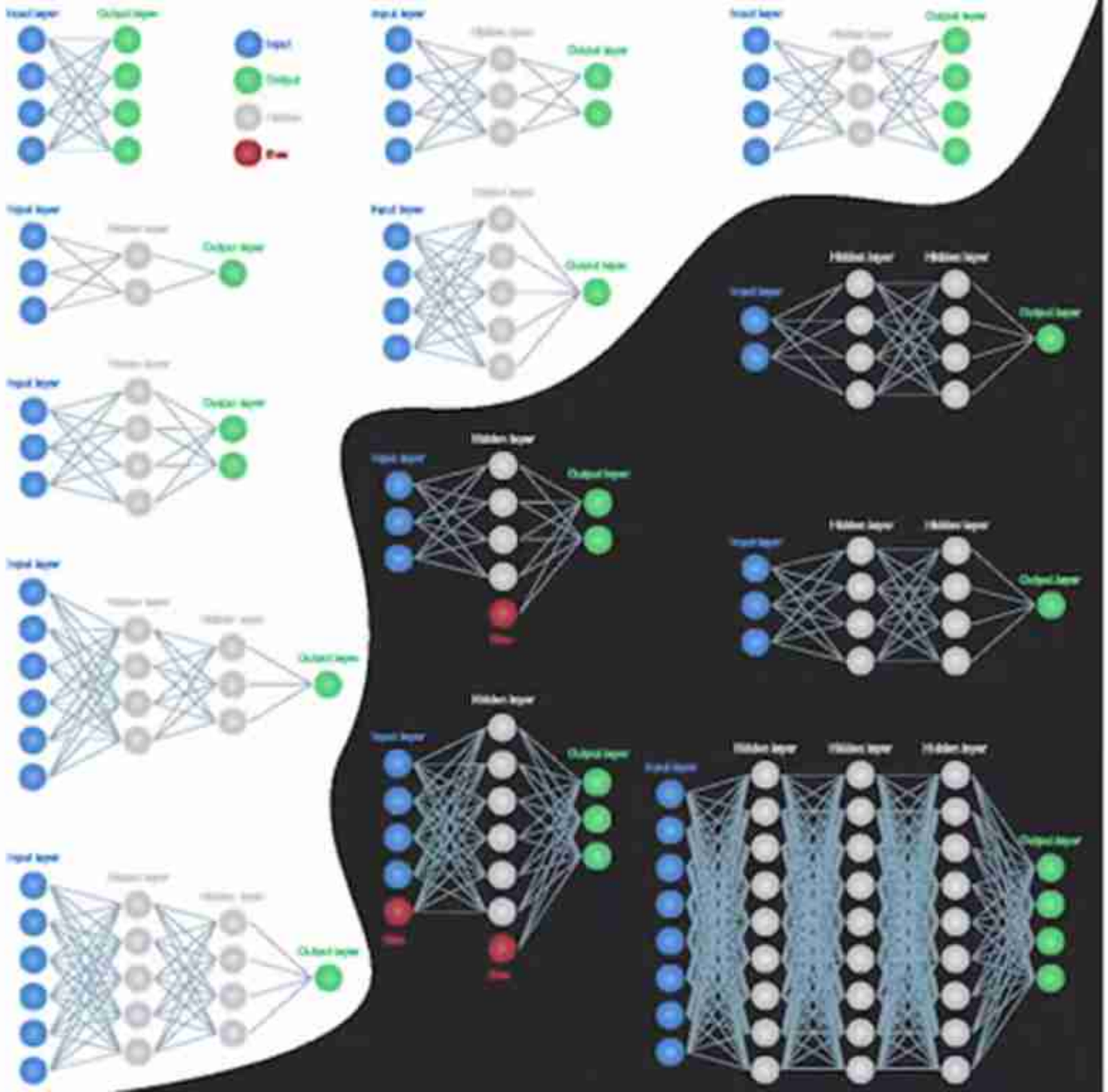


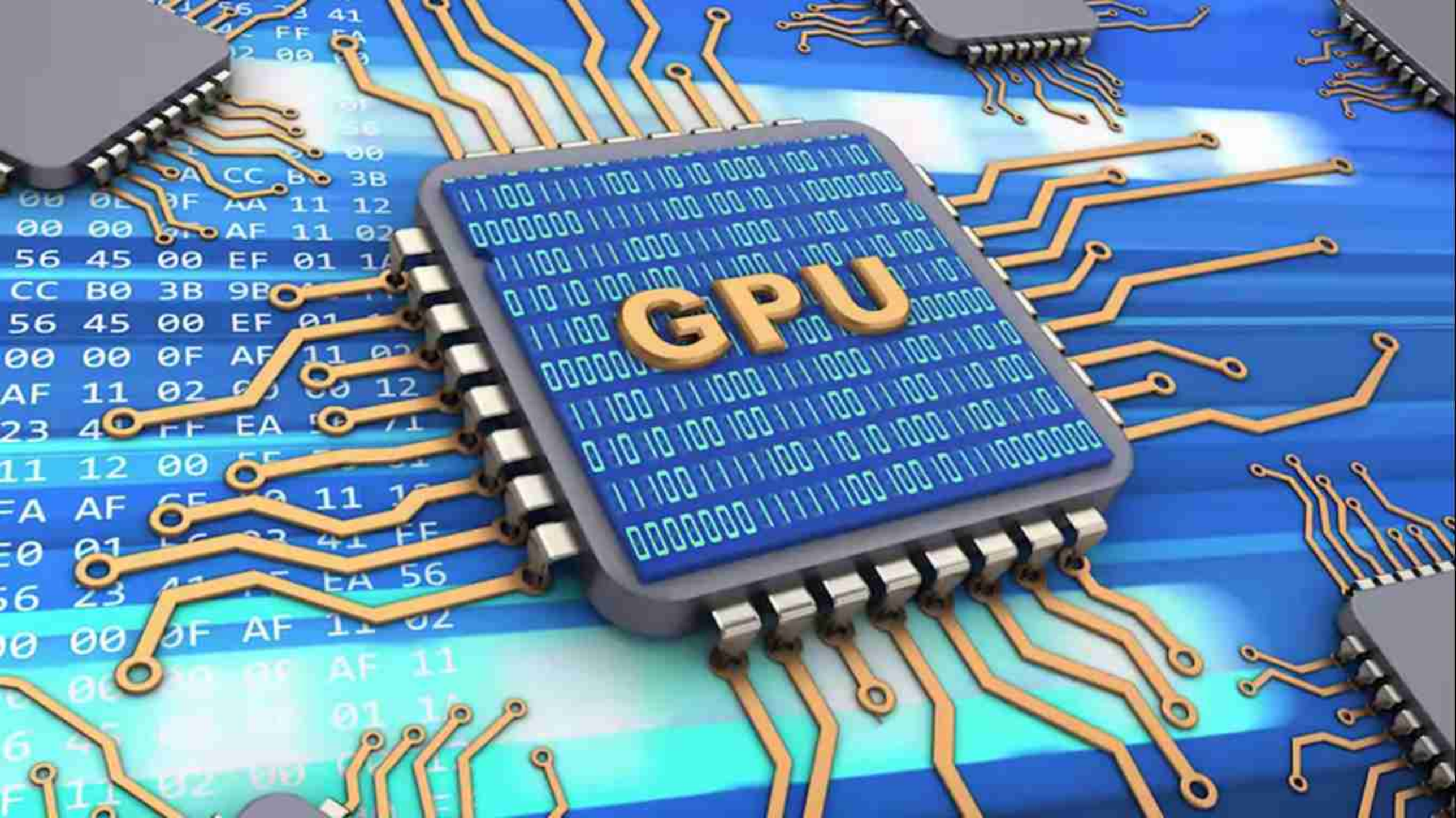


Scale Invariant Feature Transform (SIFT)

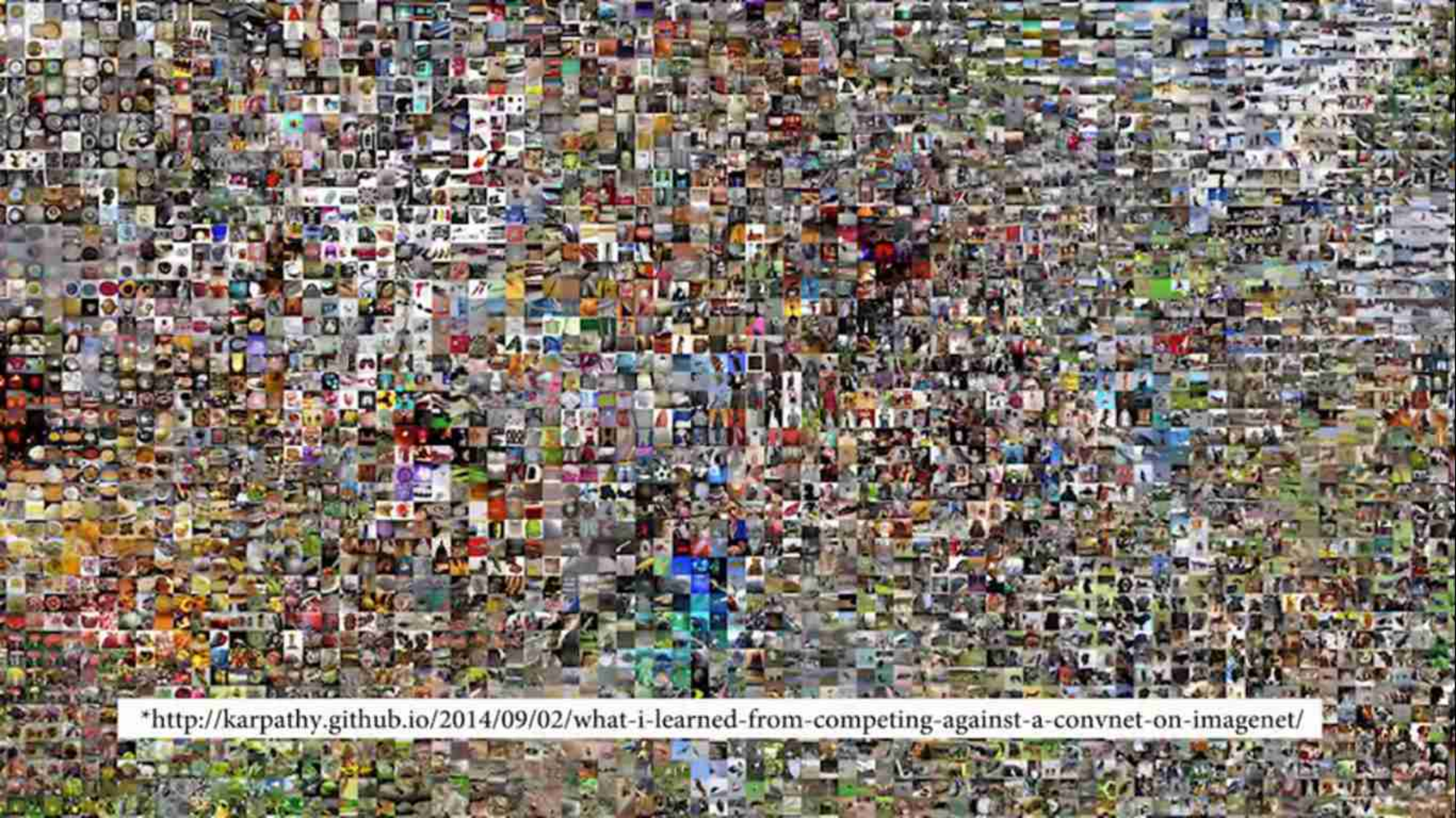
# Correspondence Problem

Scale Invariant Feature Transform (SIFT)

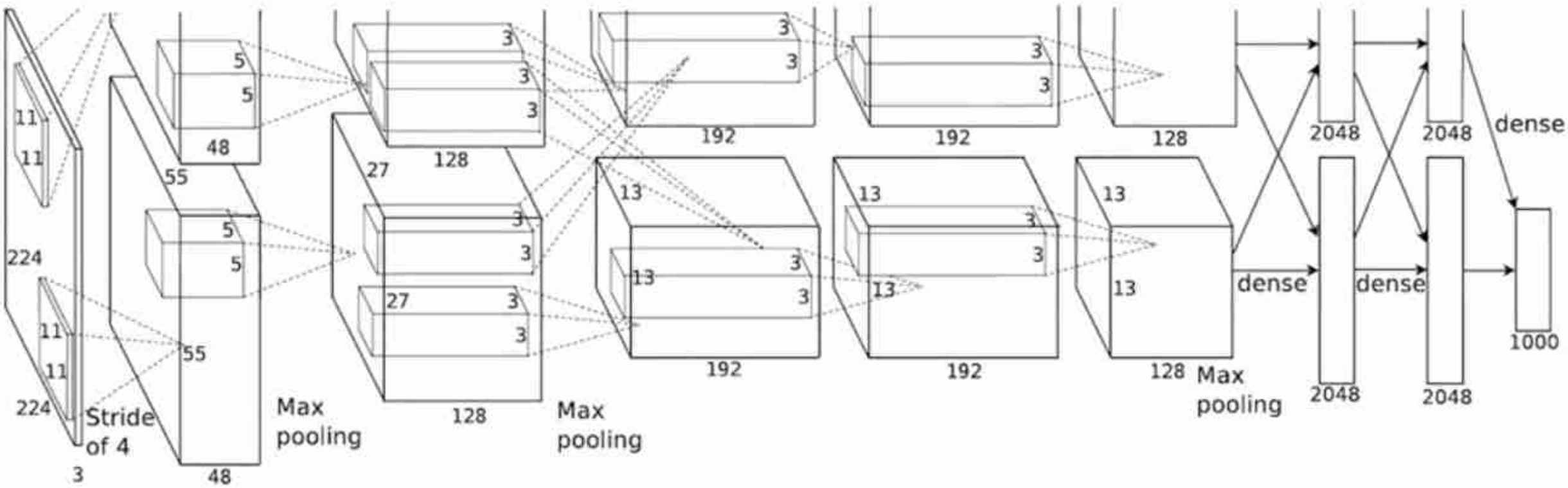

Input layer
Output layer
Input
Output
Hidden layer


GPU

*http://www.multimedia-computing.de/flickrlogos/

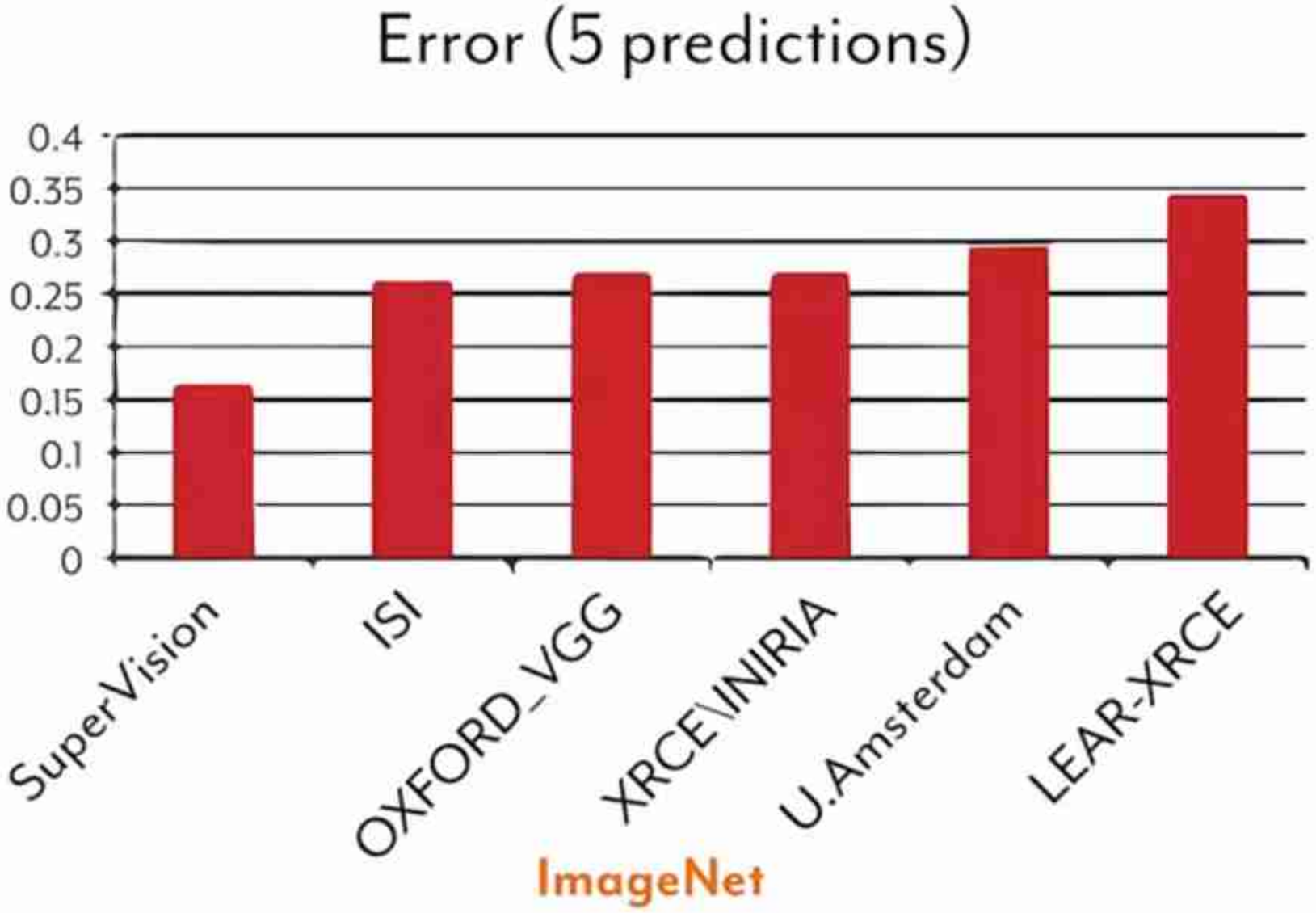


*http://karpathy.github.io/2014/09/02/what-i-learned-from-competing-against-a-convnet-on-imagenet/

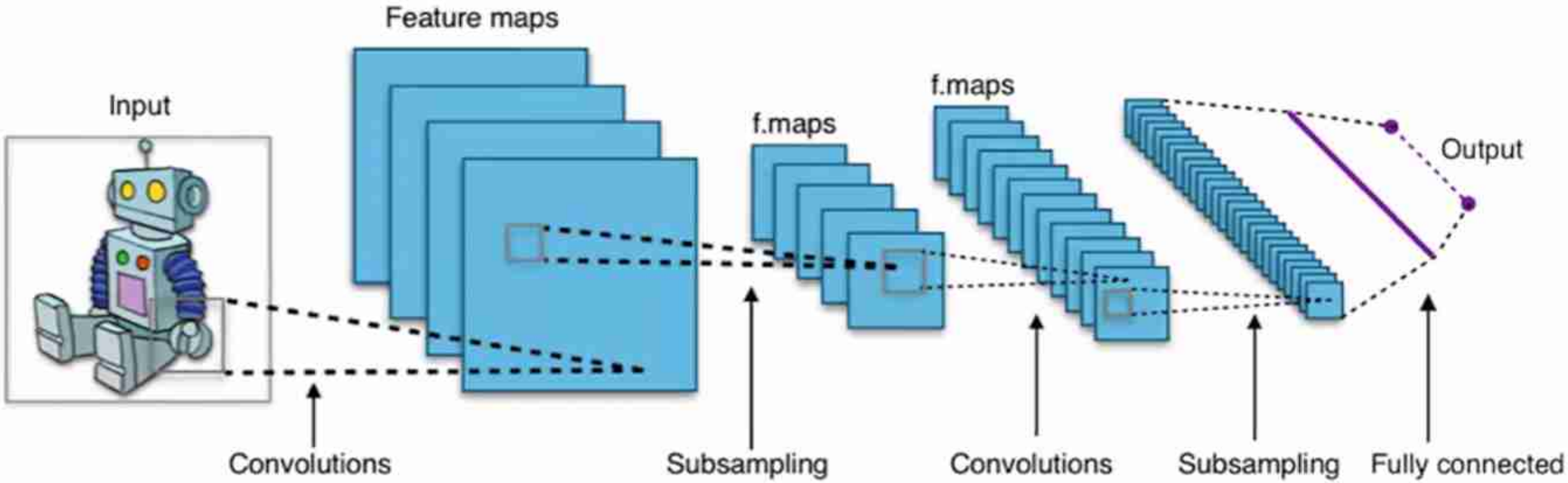



*https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

# Ranking of the best results from each team



# Convolutional Neural Networks (CNNs)



# Convolutional Neural Networks

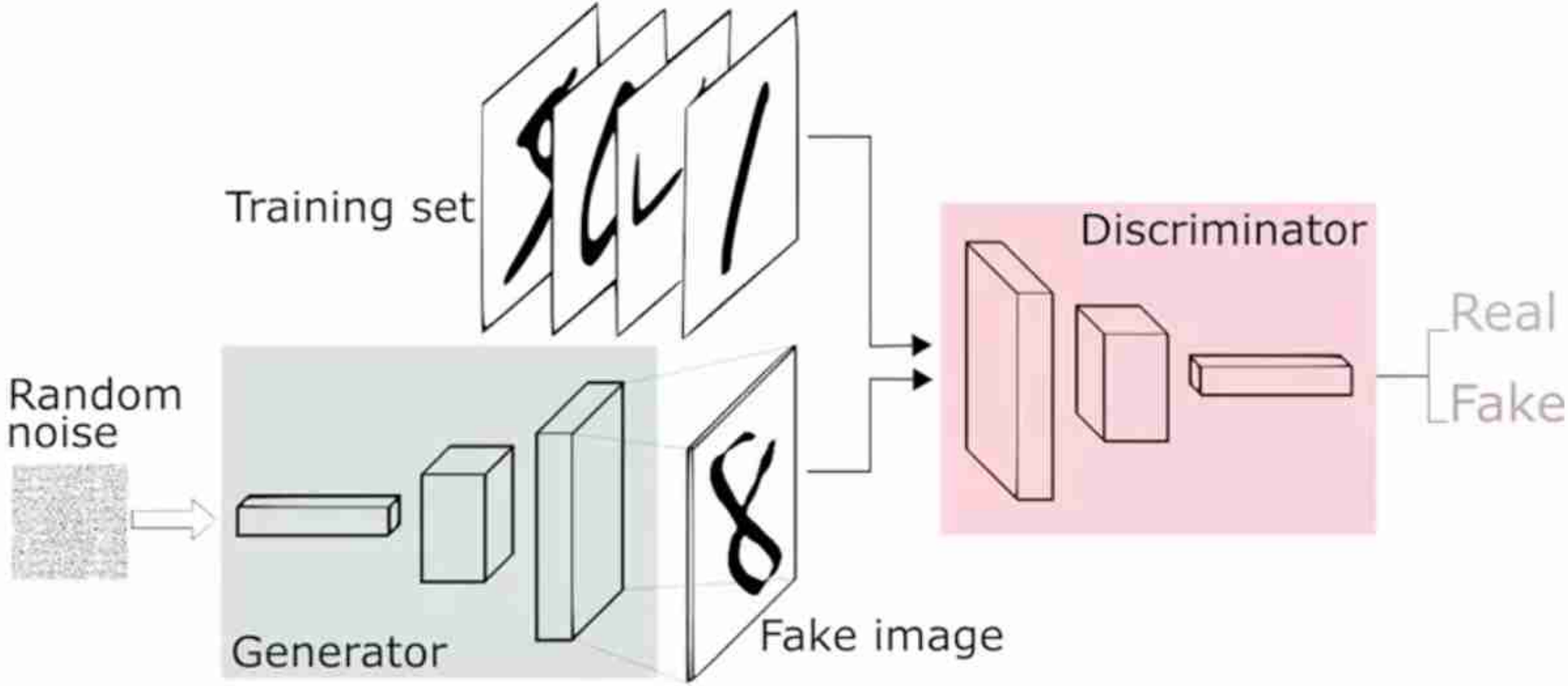
identify faces

objects

traffic signs

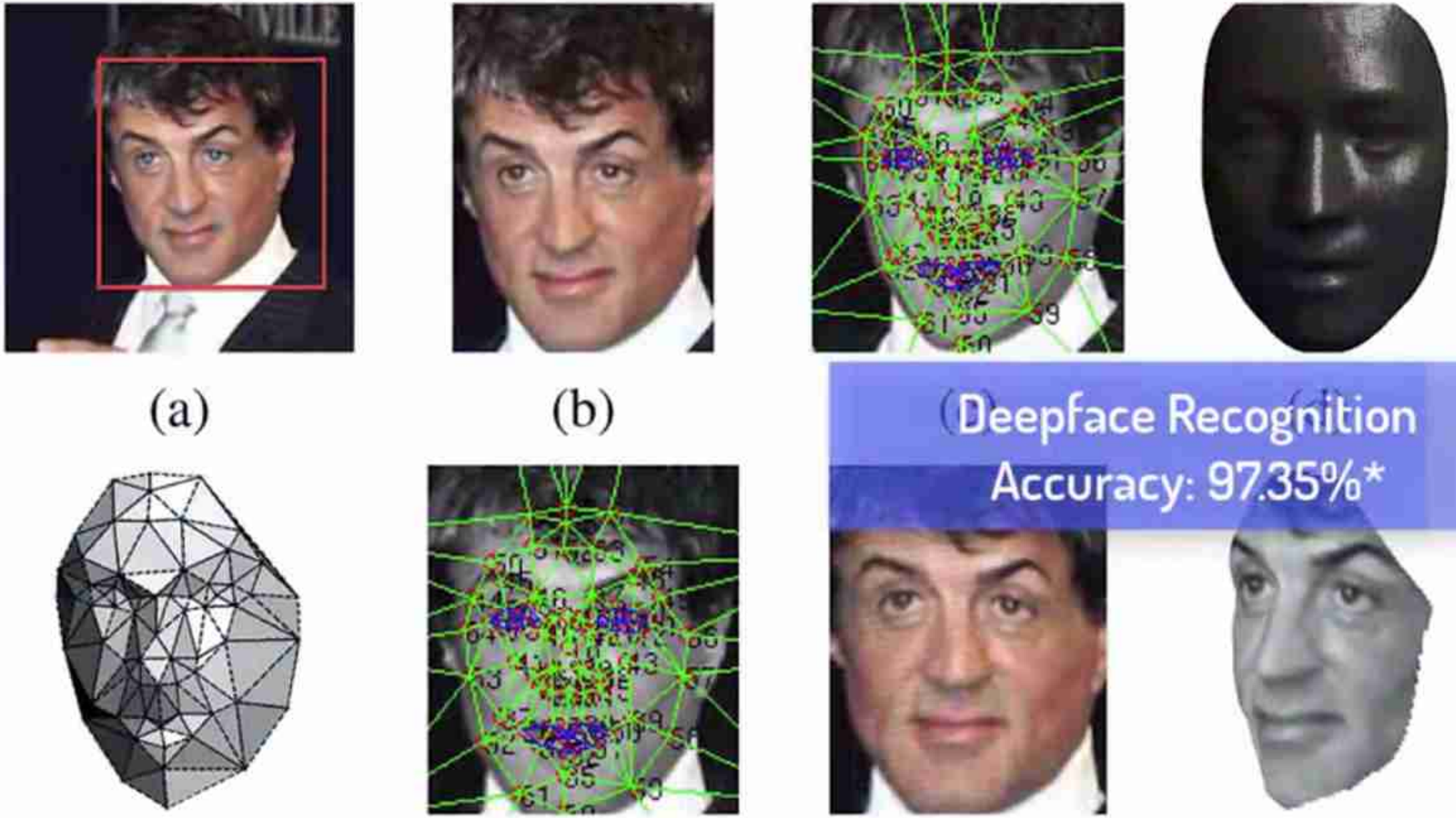
powering vision in robots

self-driving cars




*https://skymind.ai/wiki/generative-adversarial-network-gan

www.facebook.com
Deepface Recognition
Accuracy: 97.35%*
facebook




*http://time.com/29910/facebook-can-identify-faces-as-well-as-you/

erest Is Rare 'Unicorn' oaring an I.P.O. hout Hemorrhaging h

Warnings of a Dark Side to A.I. in Health Care

TECH FIX
I Deleted Facebook Last Year. Here's What Changed (and What...

BITS
The Week i Future Robo Super-Smar

TECHNOLOGY

# *How Many Computers to Identify a Cat? 16,000*

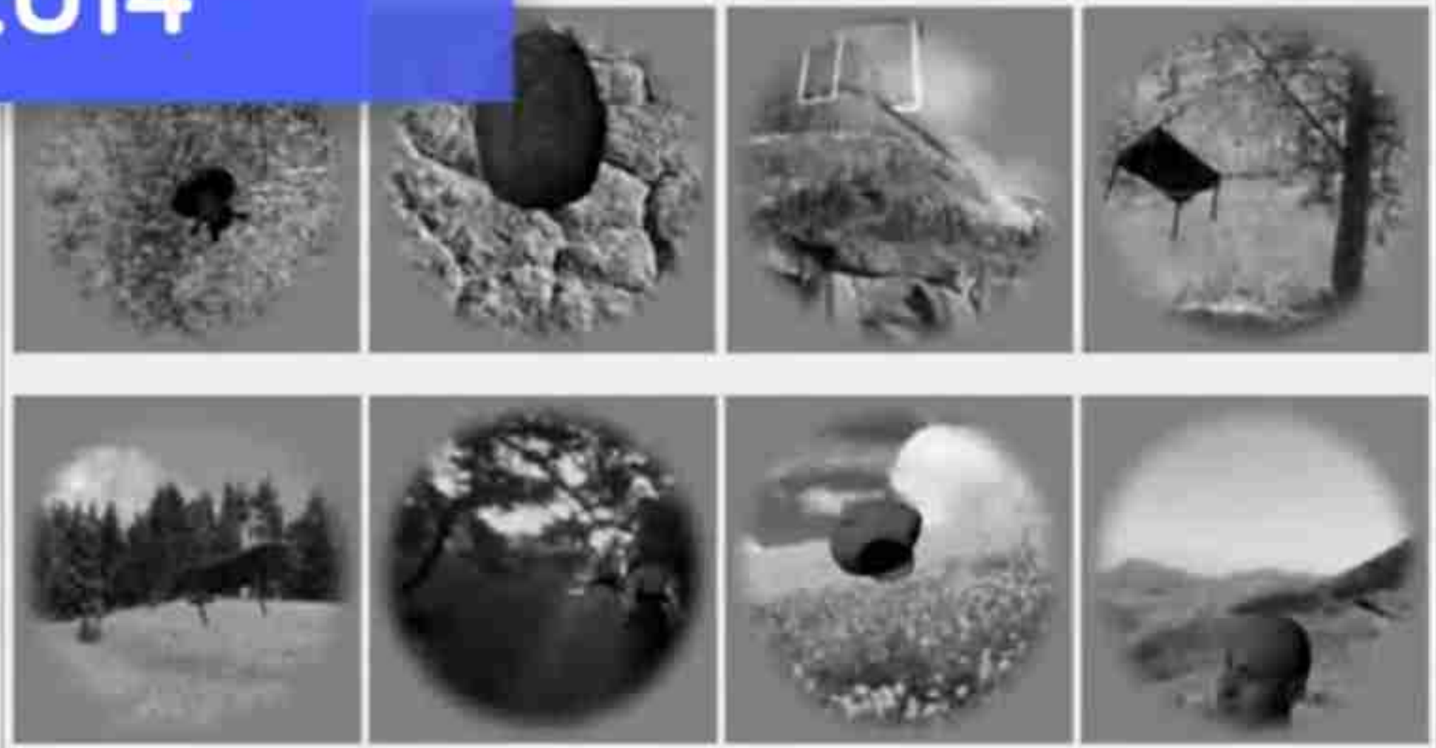
By JOHN MARKOFF JUNE 25, 2012

# MIT News

THE WORLD

Browse or Search

2014



FULL SCREEN

A team of MIT neuroscientists has found that some computer programs can identify the objects in these images just as well as the primate brain.

Image courtesy of the researchers

In one aspect of vision, computers catch up to primate brain

Intelligent Machines

# AI Has Beaten Humans at Lip-reading

2016

A pair of new studies show that a machine can understand what you're saying without hearing a sound.

by Jamie Condliffe November 21, 2016

SPLOID

# Watching a Computer Identify Objects in a Movie Trailer Is Entertaining as Hell

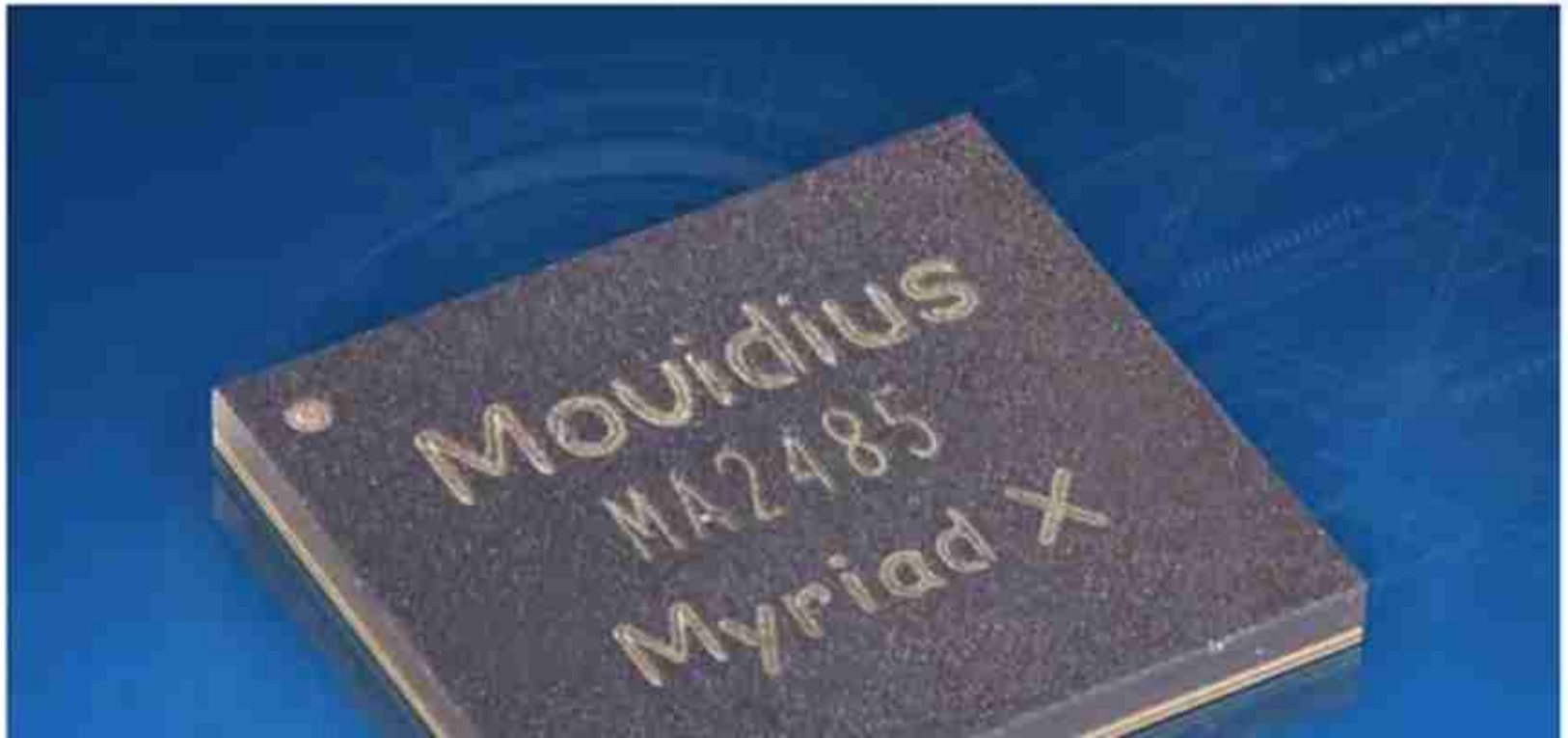
Sidney Fussell
1/17/17 12:10pm • Filed to: OBJECT RECOGNITION

2017

# Intel shows off the Movidius Myriad X, a computer vision chip with deep learning baked-in

Lucas Matney @lucasmtny / 2 years ago

2017



2019

SMART HOME | LEER EN ESPAÑOL

# Facial recognition: Apple, Ama
# Google and the race for your f

Facial recognition technology is both innovative and worrisome.
how it works and what you need to know.

BY ANDREW GEBHART | MARCH 18, 2019 3:00 PM PDT

2019

MS TECH

Intelligent Machines

# Trained neural nets perform much like humans on classic psychological tests

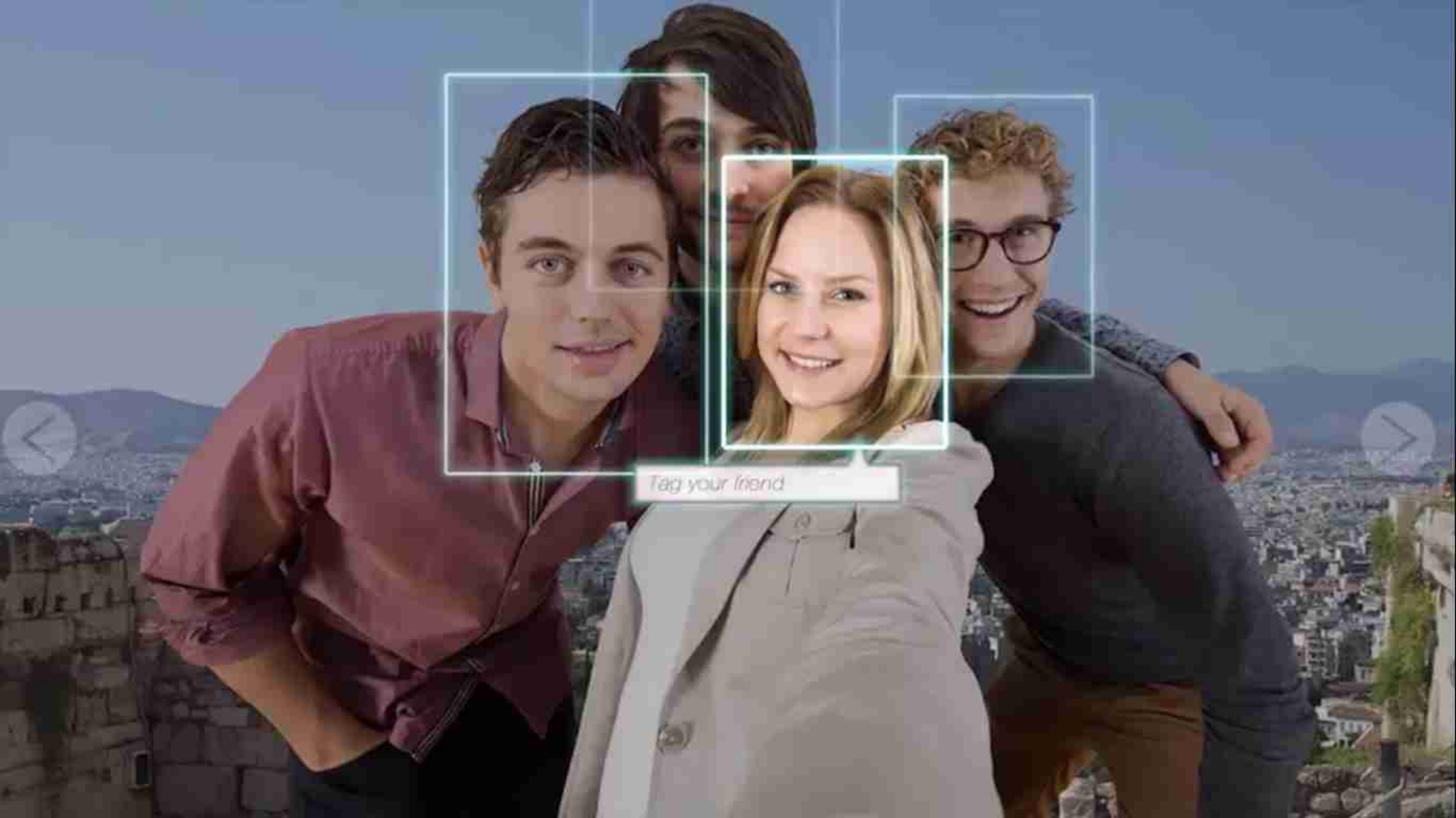
Applications
Video Monitoring Center
Full Video Control

# Applications of Computer Vision

Industrial Vision Systems

inspect bottles on a production line

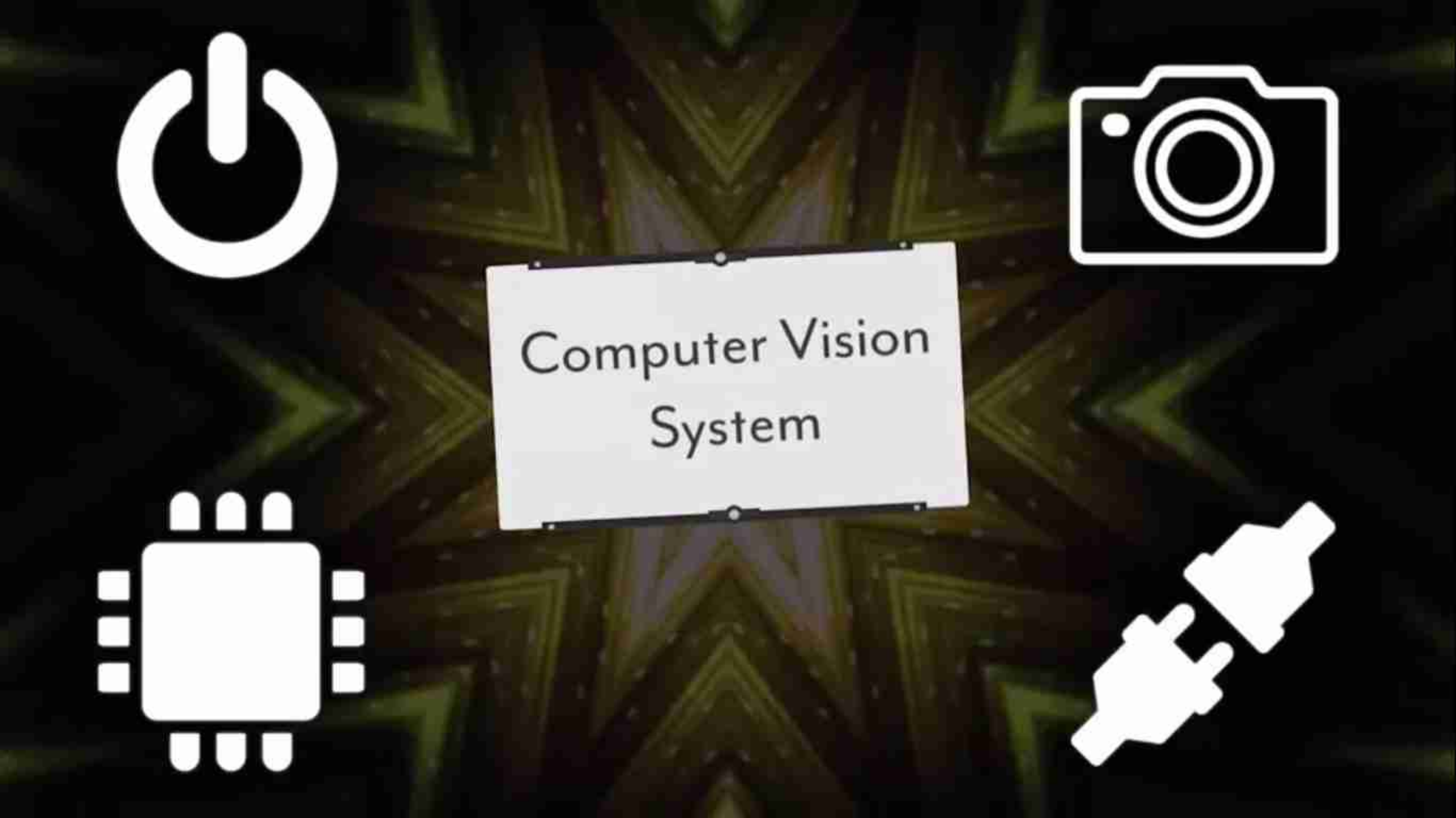
robots that can comprehend the world around them

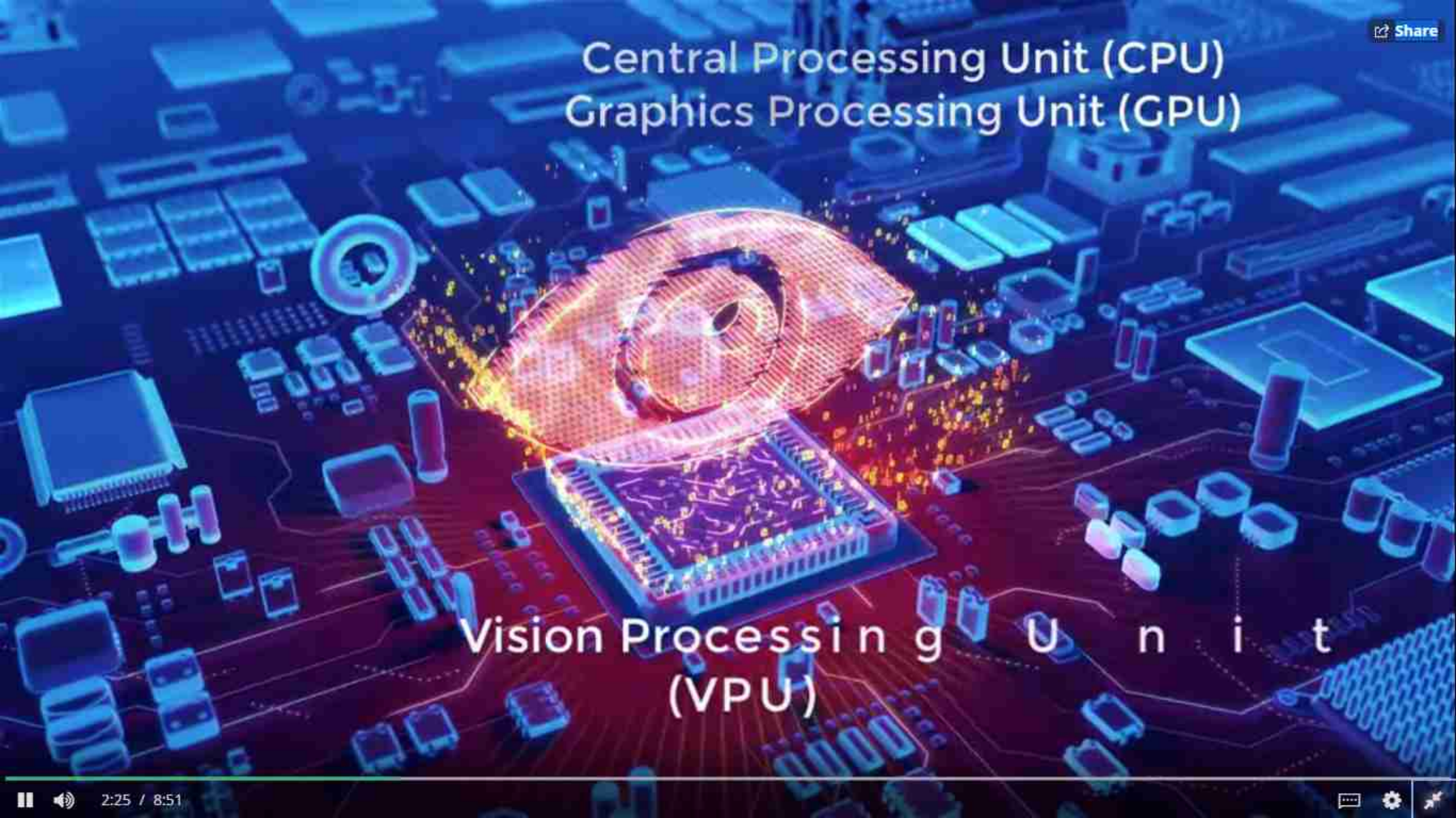

Tag your friend

# Non-Traditional Areas

astronomy

nanotechnology


Computer Vision
System


Share
Central Processing Unit (CPU)
Graphics Processing Unit (GPU)
Vision Processing Unit
(VPU)
2:25 / 8:51

Share
2:31 / 8:51

Visual Surveillance

CAM1
05;25;32;28
CAM2
05;25;32;28
CAM3
05;25;32;28
CAM4
05;25;32;28

BIOMETRIC FINGERPRINT SCAN

1 INDEX

2 MIDDLE

3 RING

4 LITTLE

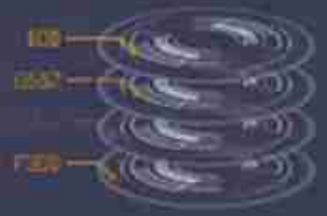
5 THUMB

STATUS: ANALYZING

%6

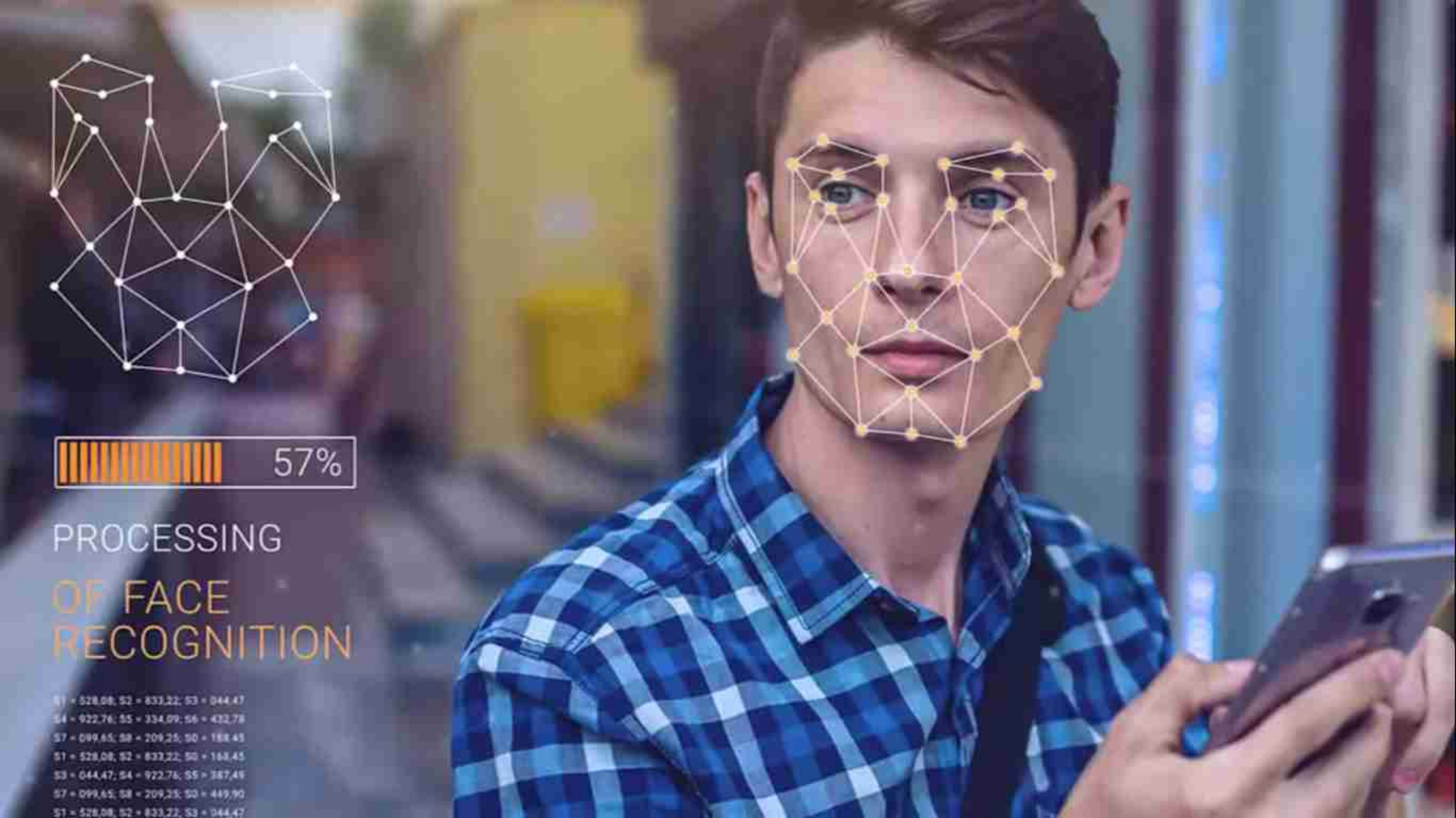
PATTERN ANALYSIS

DATABASE ACCESS



Iris Recognition

How the Afghan Girl was Identified by Her Iris Patterns
Courtesy: John Daugman


57%
PROCESSING
OF FACE
RECOGNITION
S1 = 528,08; S2 = 833,22; S3 = 044,47
S4 = 922,76; S5 = 334,09; S6 = 432,78
S7 = 099,65; S8 = 209,25; S0 = 188,45
S1 = 528,08; S2 = 833,22; S0 = 188,45
S3 = 044,47; S4 = 922,76; S5 = 387,49
S7 = 099,65; S8 = 209,25; S0 = 449,90
S1 = 528,08; S2 = 833,22; S3 = 044,47

# Computer Vision in Multimedia and Entertainment

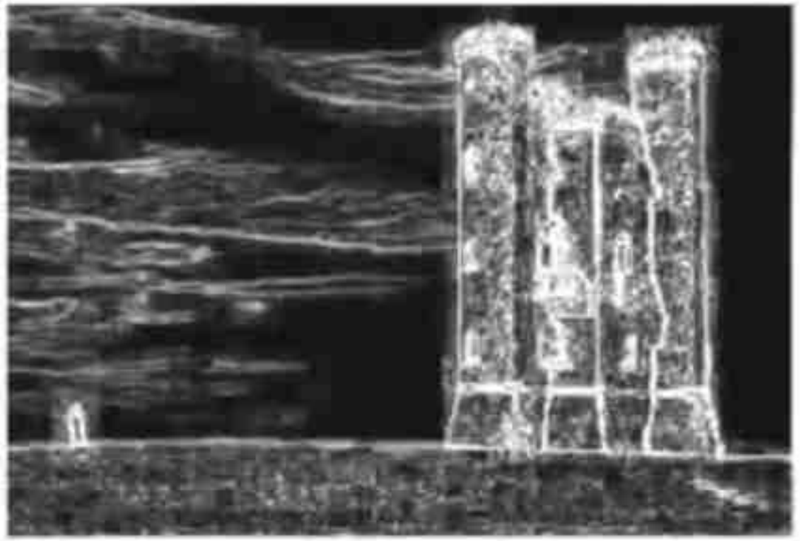
# Face Filters

# Mixed Reality Games

Cropping

Courtesy: Author Newton2, Wikipedia, Seam Carving

Scaling



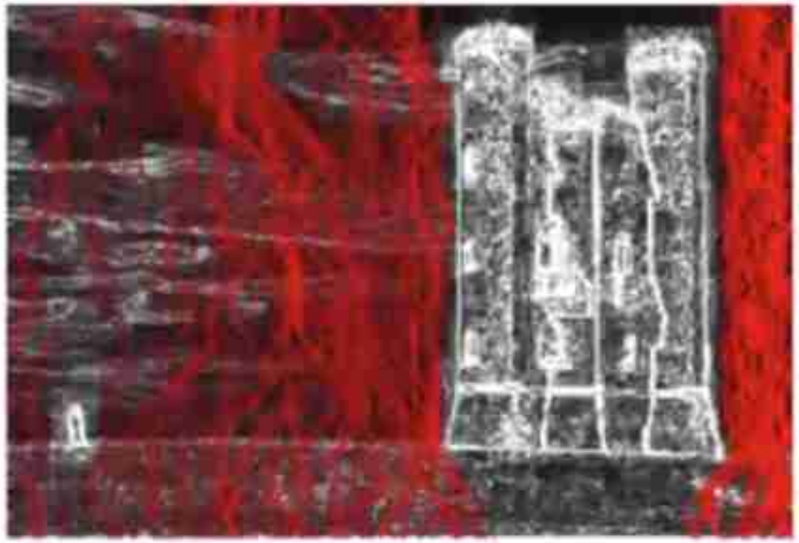
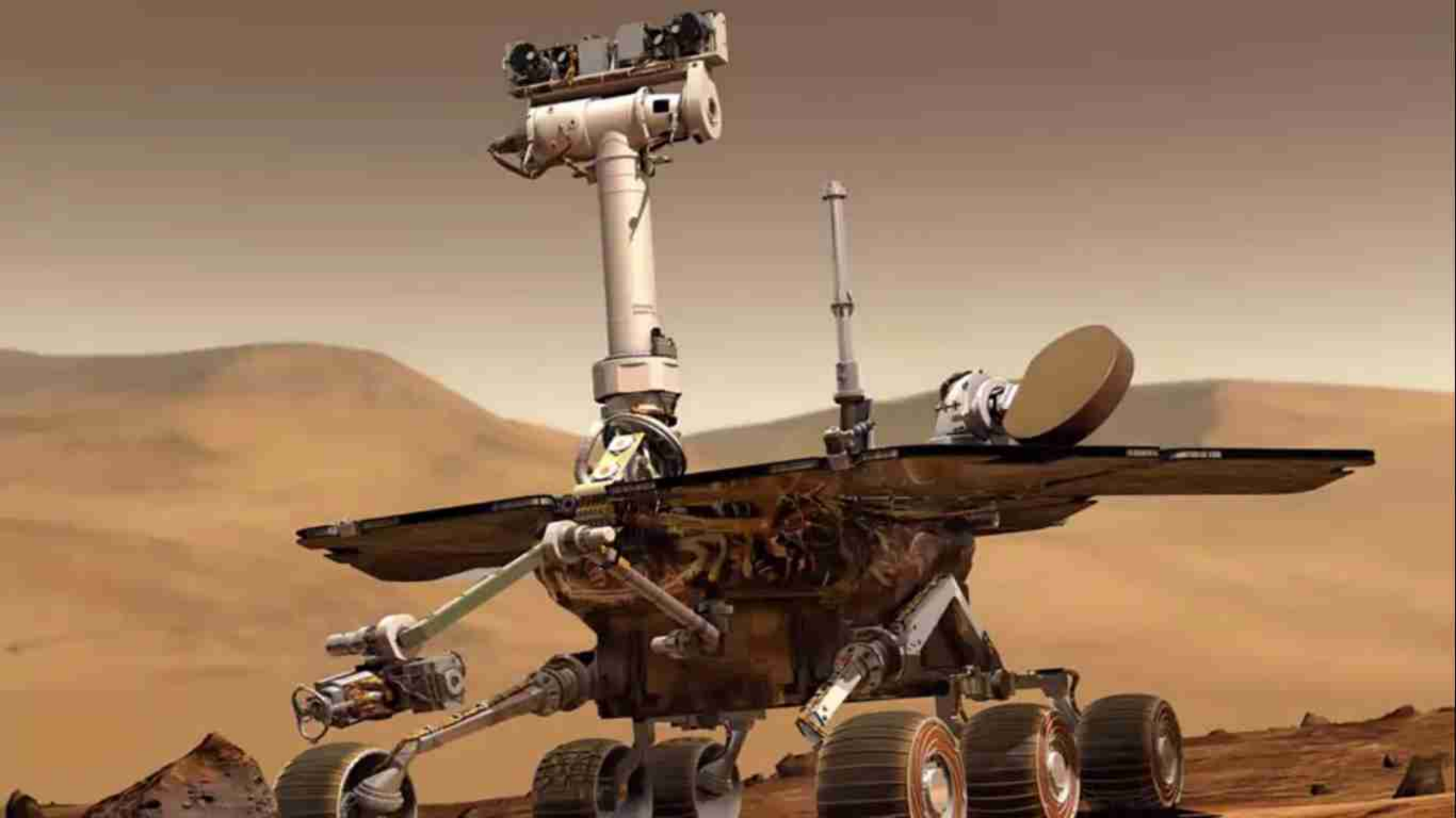


Image Retargeting

Making of "Planet
Educatio

# Computer Vision for Navigation



# Robot Navigation

Courtesy: Maira Sab

Google maps
Street View

Barbara and Jack
Davis Hall
W White Rd
Google
Image capture: Aug 2016 © 2018 Google United States Terms Report a problem

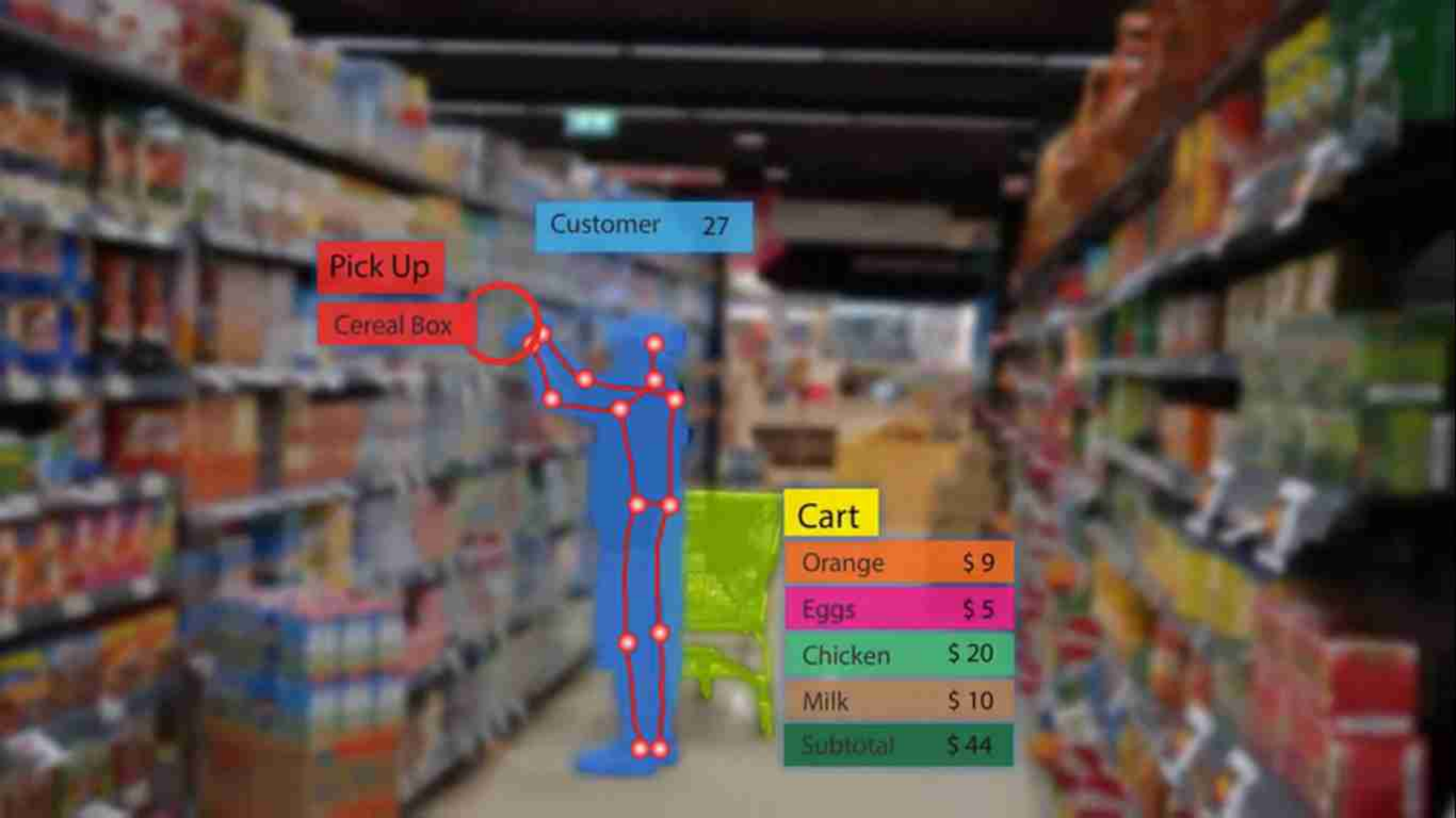
Lane Detection, Courtesy: Dr. Panya Chanawangsa

Moving vehicle
40 km/h
Human

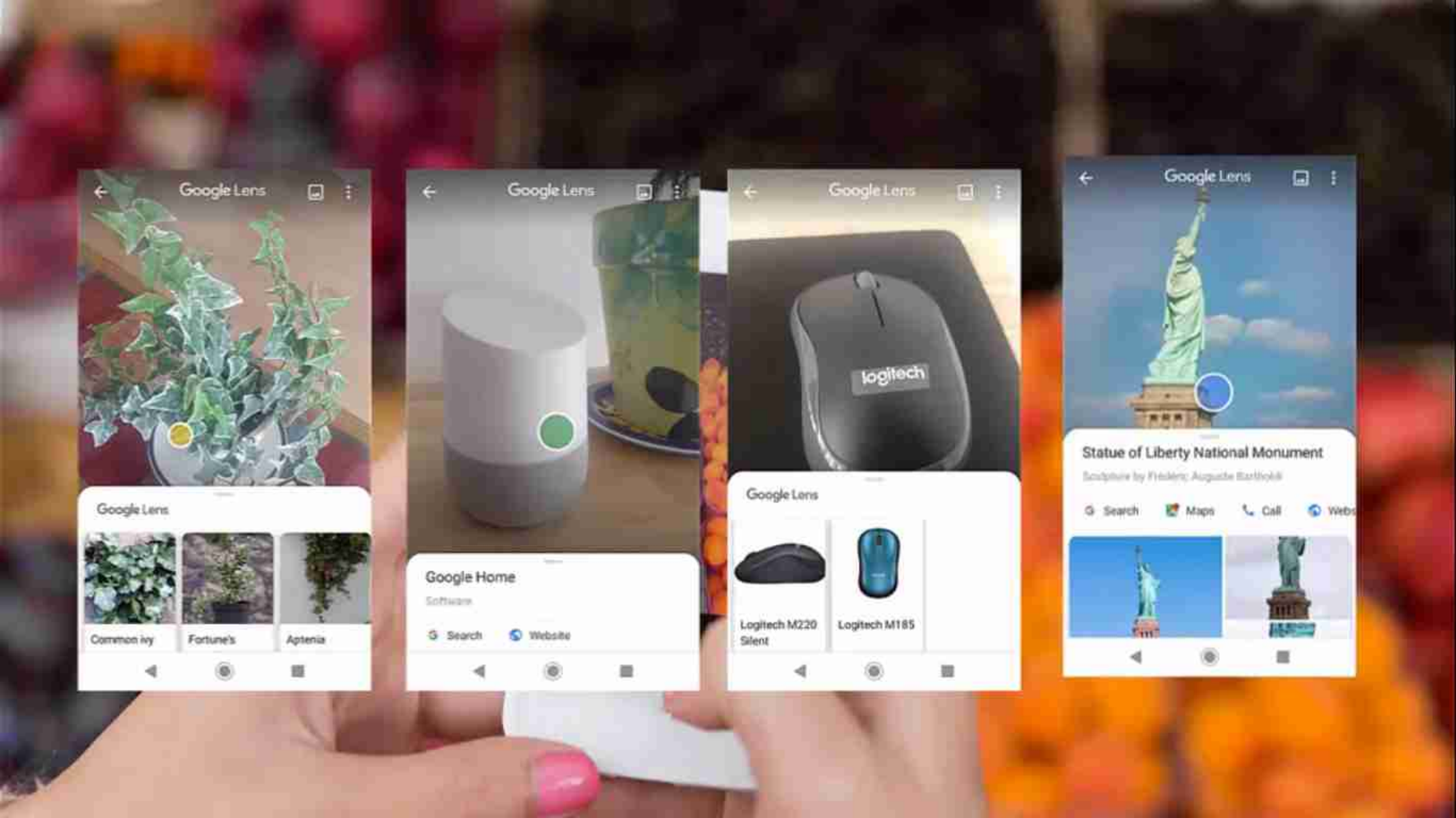
Share

# Computer Vision in Retail

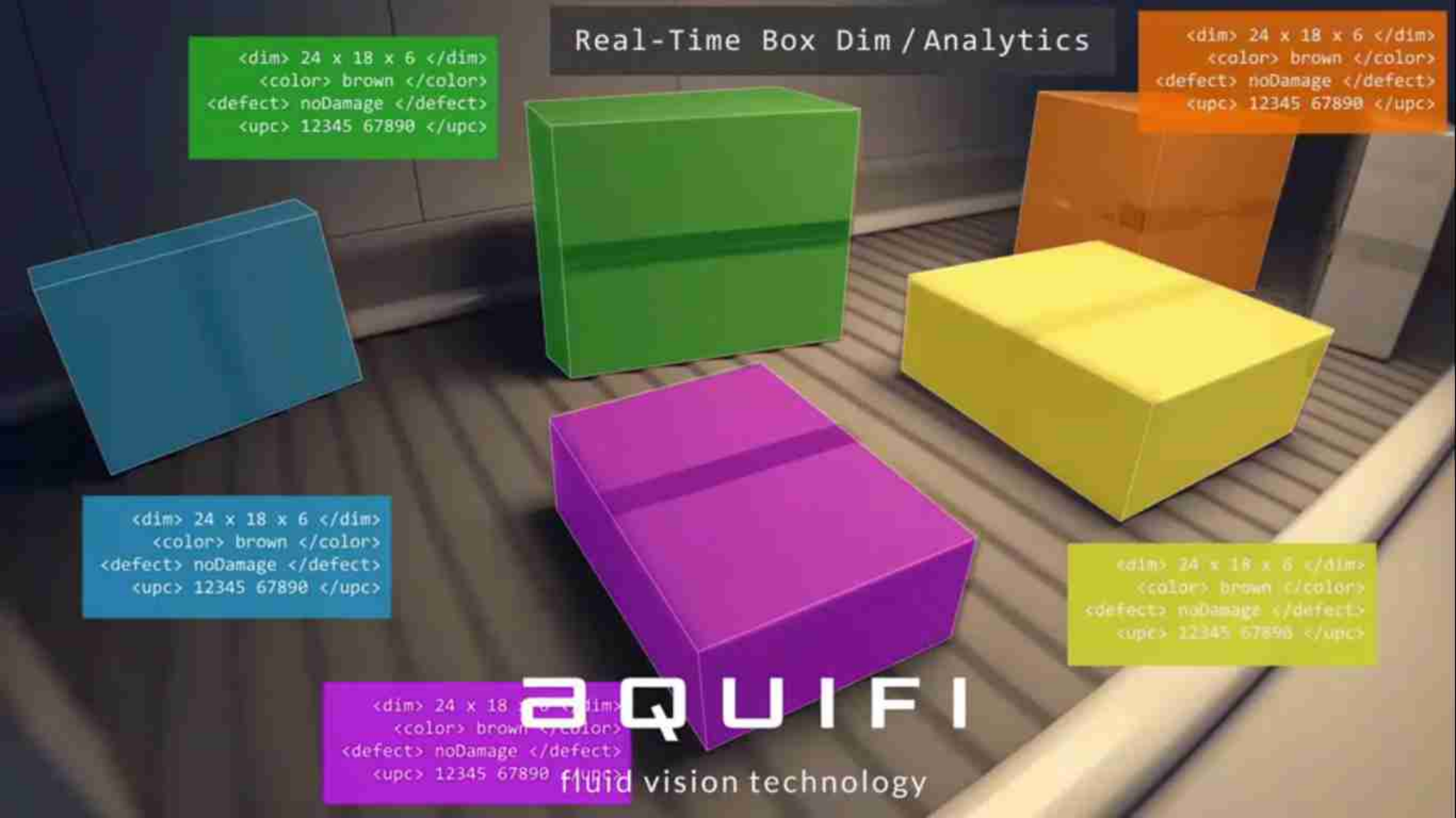
5:14 / 8:51

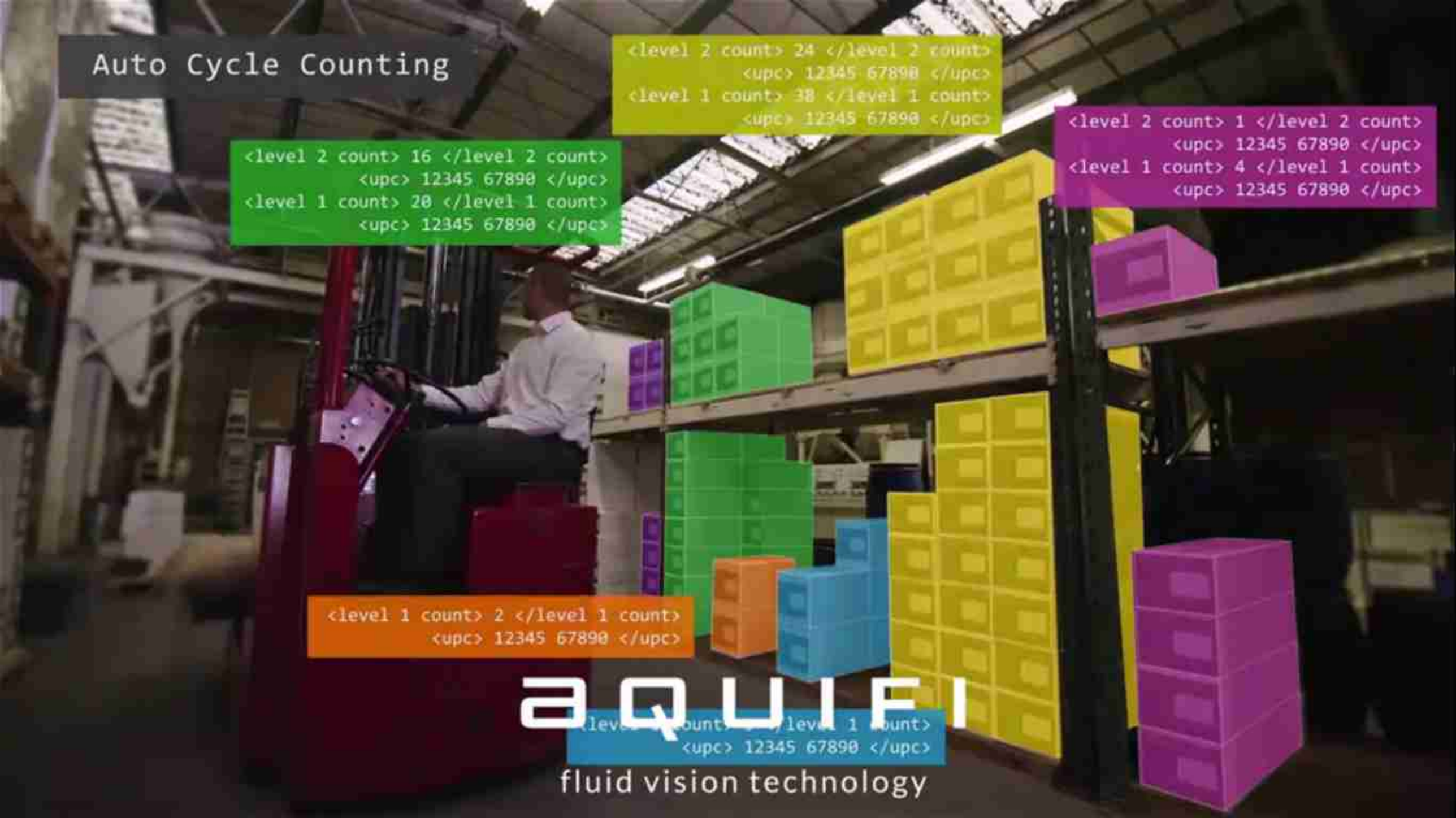

Customer 27
Pick Up
Cereal Box
Cart
Orange $ 9
Eggs $ 5
Chicken $ 20
Milk $ 10
Subtotal $ 44

Google Lens
Google Lens
Common ivy
Fortune's
Aptenia


Google Lens
Google Lens
Common ivy
Fortune's
Aptenia
Google Lens
Google Home
Software
Search
Website
Google Lens
logitech
Google Lens
Logitech M220 Silent
Logitech M185
Google Lens
Statue of Liberty National Monument
Search
Maps
Call

# Computer Vision
# Applications in Industry


Real-Time Box Dim / Analytics
<dim> 24 x 18 x 6 </dim>
<color> brown </color>
<defect> noDamage </defect>
<upc> 12345 67890 </upc>
<dim> 24 x 18 x 6 </dim>
<color> brown </color>
<defect> noDamage </defect>
<upc> 12345 67890 </upc>
<dim> 24 x 18 x 6 </dim>
<color> brown </color>
<defect> noDamage </defect>
<upc> 12345 67890 </upc>
<dim> 24 x 18 x 6 </dim>
<color> brown </color>
<defect> noDamage </defect>
<upc> 12345 67890 </upc>
<dim> 24 x 18 x 6 </dim>
<color> brown </color>
<defect> noDamage </defect>
<upc> 12345 67890 </upc>
aquifi
fluid vision technology

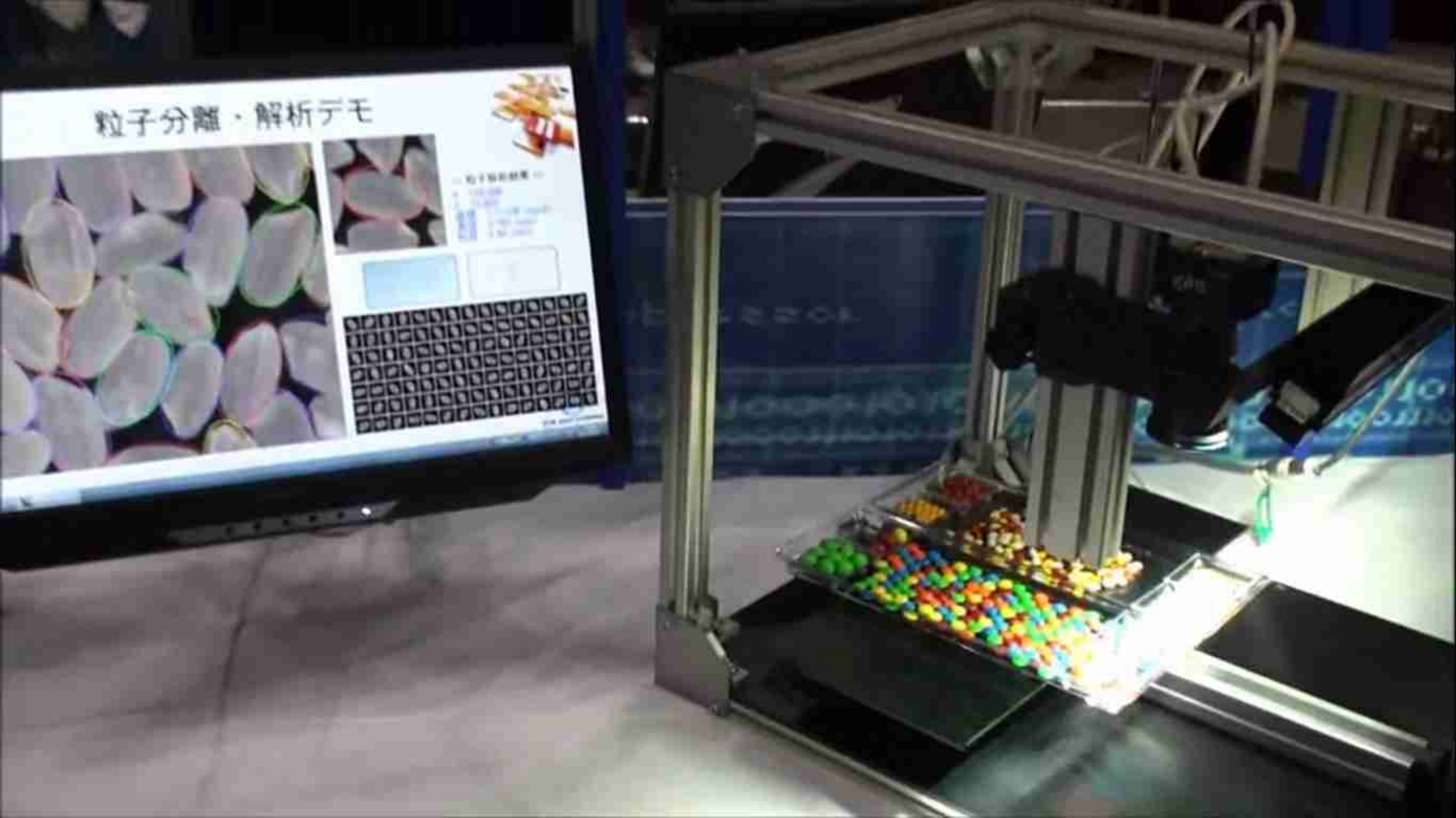

Auto Cycle Counting
<level 2 count> 24 </level 2 count>
<upc> 12345 67890 </upc>
<level 1 count> 38 </level 1 count>
<upc> 12345 67890 </upc>
<level 2 count> 16 </level 2 count>
<upc> 12345 67890 </upc>
<level 1 count> 20 </level 1 count>
<upc> 12345 67890 </upc>
<level 2 count> 1 </level 2 count>
<upc> 12345 67890 </upc>
<level 1 count> 4 </level 1 count>
<upc> 12345 67890 </upc>
<level 1 count> 2 </level 1 count>
<upc> 12345 67890 </upc>
<upc> 12345 67890 </upc>
aQUIFI
fluid vision technology

Auto Cycle Counting
<level 3 count> 26 </level 4 count>
<upc> 12345 67890 </upc>
<level 3 count> 6 </level 4 count>
<upc> 12345 67890 </upc>
<level 3 count> 11 </level 4 count>
<upc> 12345 67890 </upc>
AQUIFI
fluid vision technology

Identify, Inspect, Analyze
<model_ID> RunMax </model_ID>
<side> Right </side>
<size> 9.0 </size>
<stitching> Inspect </stitching>
<shape> 80 </shape>
<holes> 10 </holes>
<model_ID> RunMax </model_ID>
<side> Left </side>
<size> 9.5 </size>
<stitching> Good </stitching>
<shape> 99 </shape>
<holes> 10 </holes>
AQUIFI
fluid vision technology

# Segmenting and Analysing Candies and Seeds by Blob Analysis with HALCON

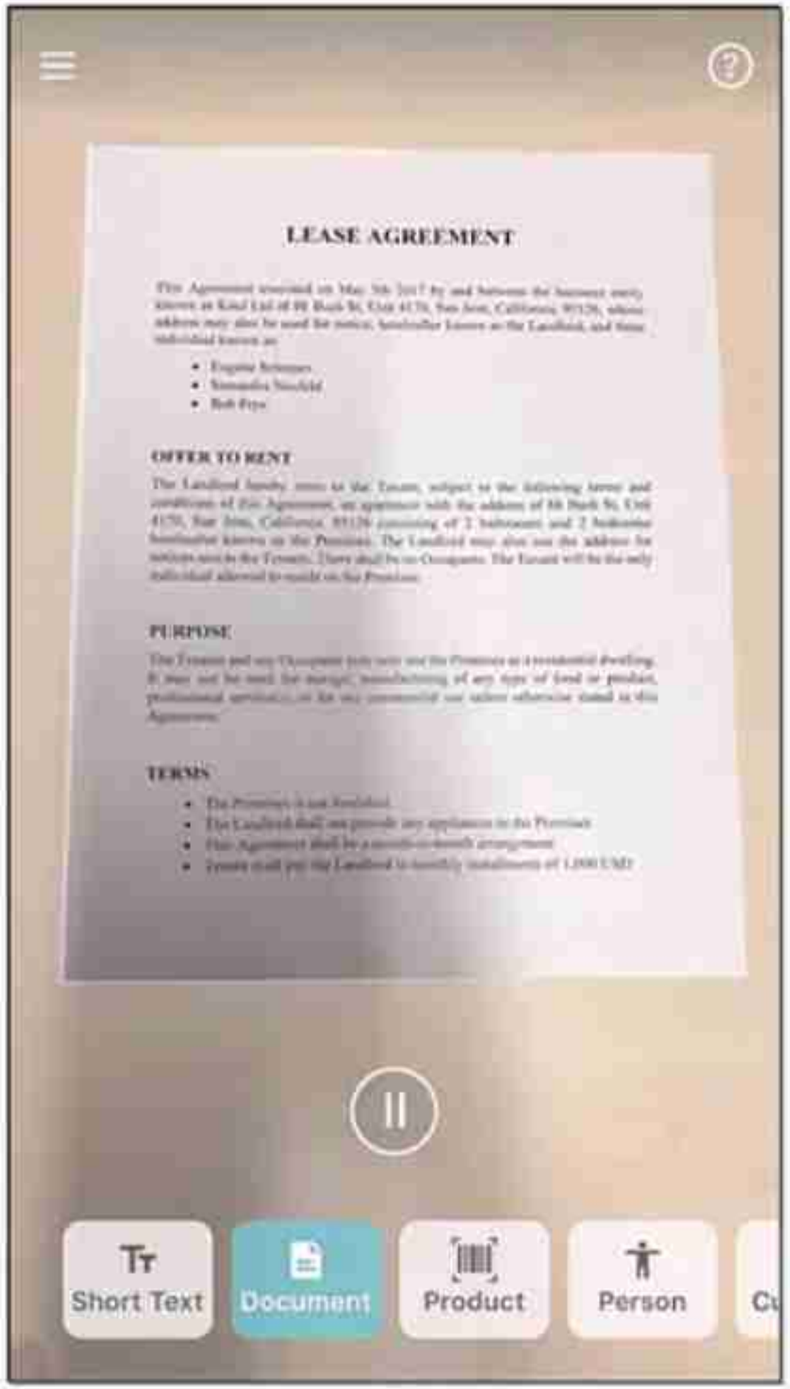
by LinX, Yokohama


粒子分離・解析デモ

# Computer Vision
# Applications for social causes

Seeing AI

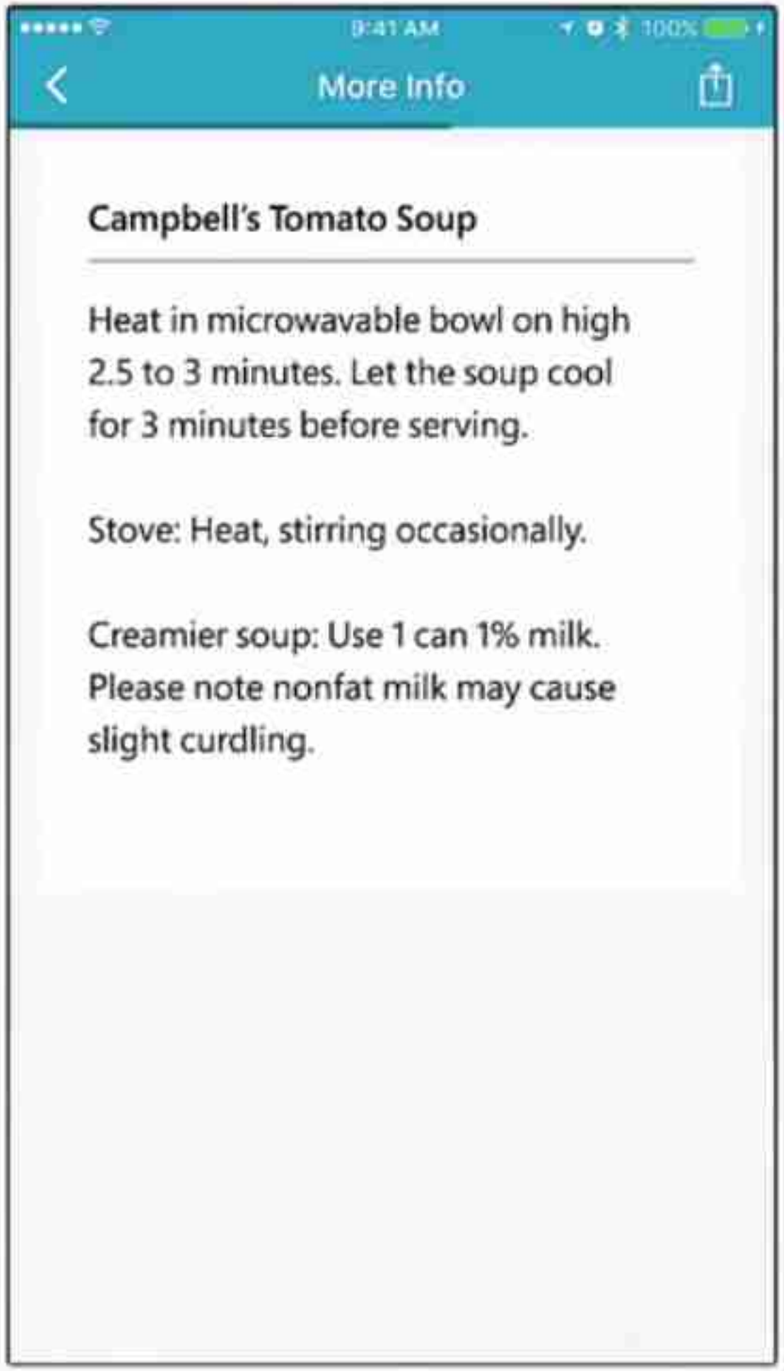
28 year old female wearing glasses
looking happy.
Short Text
Document
Product

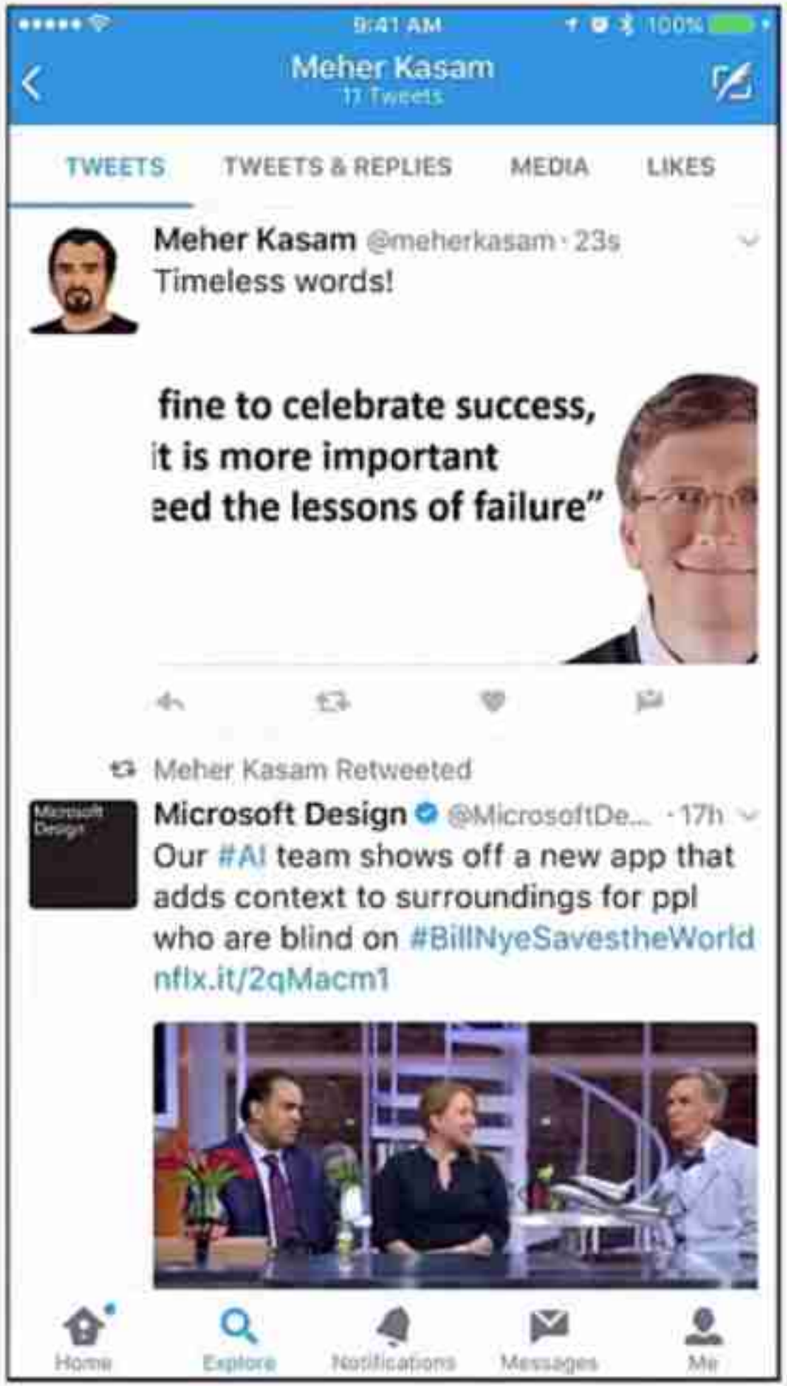

LEASE AGREEMENT
OFFER TO RENT
PURPOSE
TERMS
Short Text
Document
Product
Person

Campbell's Tomato Soup

Share

More Info


9:41 AM
100%
More Info
Campbell's Tomato Soup
Heat in microwavable bowl on high 2.5 to 3 minutes. Let the soup cool for 3 minutes before serving.
Stove: Heat, stirring occasionally.
Creamier soup: Use 1 can 1% milk. Please note nonfat milk may cause slight curdling.


9:41 AM
100%
Meher Kasam
11 Tweets
TWEETS
TWEETS & REPLIES
MEDIA
LIKES
Meher Kasam @meherkasam · 23s
Timeless words!
fine to celebrate success,
it is more important
eed the lessons of failure"
Meher Kasam Retweeted
Microsoft Design @MicrosoftDe... · 17h
Our #AI team shows off a new app that adds context to surroundings for ppl who are blind on #BillNyeSavestheWorld nflx.it/2qMacm1
Home
Explore
Notifications
Messages
Me

A young girl throwing a Frisbee in the park.

Save Photo

Share

differential mode differential amplifier2 - Multisim - [differential mode differential amplifier2]

File Edit View Place MCU Simulate Transfer Tools Reports Options Window Help

Design Toolbox

erential mode differential amplifi
differential mode differential am

Oscilloscope-XSC1

T1
T2
T2-T1

Time Channel_A Channel_B

Reverse
Save
Ext. trigger

Timebase
Scale: 1 s/Div
X pos.(Div): 0
Y/T Add B/A A/B

Channel A
Scale: 10 mV/Div
Y pos.(Div): 1
AC 0 DC

Channel B
Scale: 2 V/Div
Y pos.(Div): -1
AC 0 DC -

Trigger
Edge:
Level: 0 V
Single Normal Auto None

Hierarchy Visibility Project View

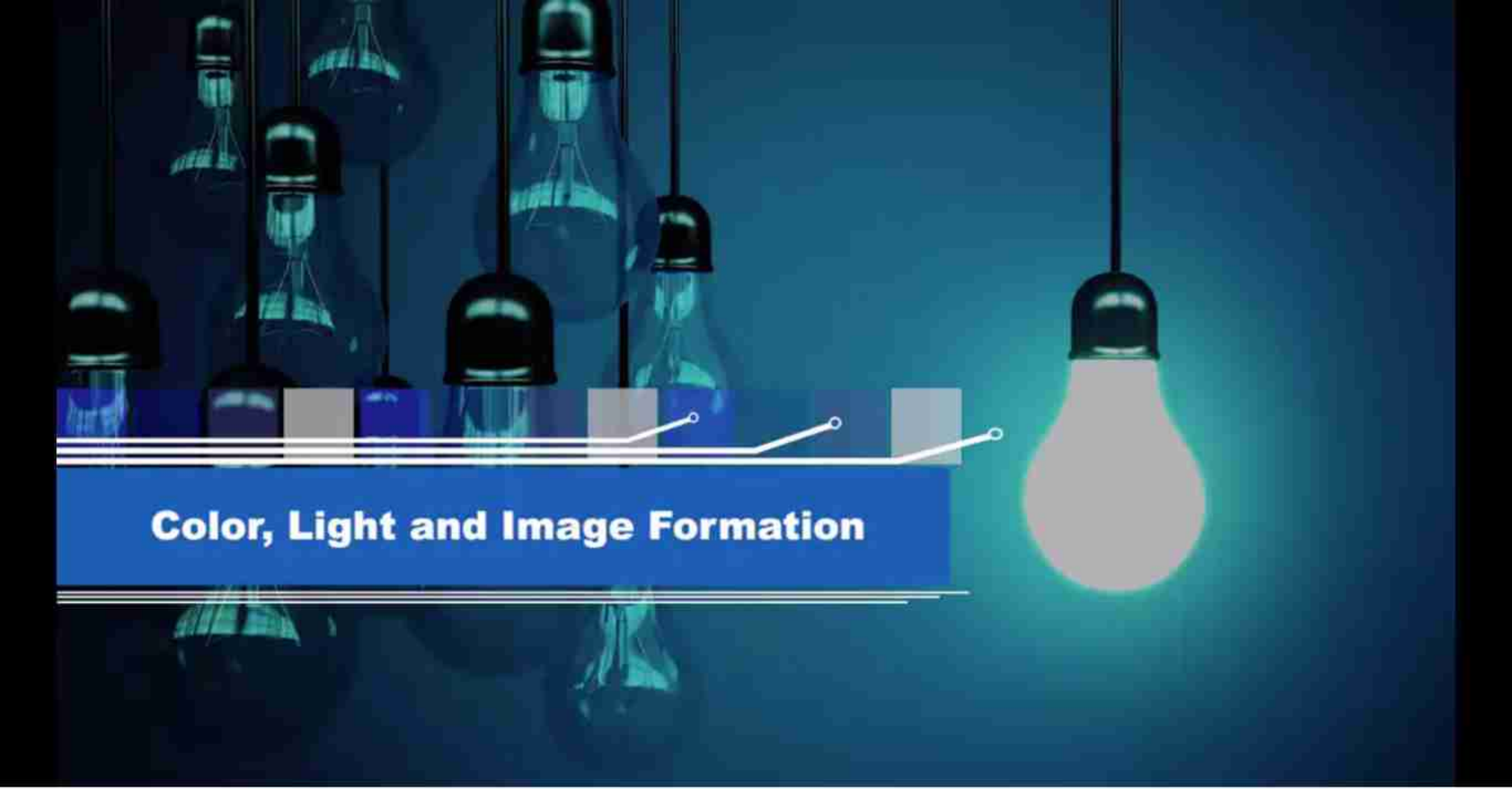
differential mode differential amplifier2

Multisim - 16 December 2020, 13:14:35

Results Nets Components Copper layers Simulation

Type here to search

ENG 13:15 16-12-2020




cameraman.png
Show all

cameraman.png
Show all

cameraman.png
Show all

# Factors That Affect "Color"

light sources

object surface properties

emittance

reflectance spectrum

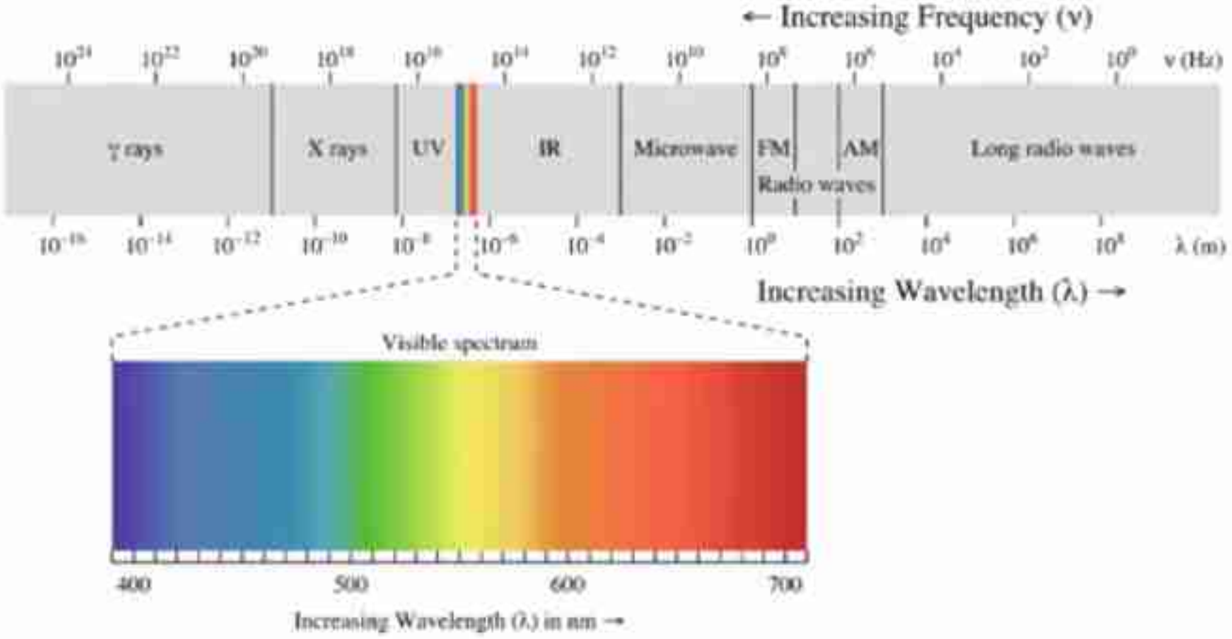
position and orientation

and their relative position
and orientation.

cameraman.png

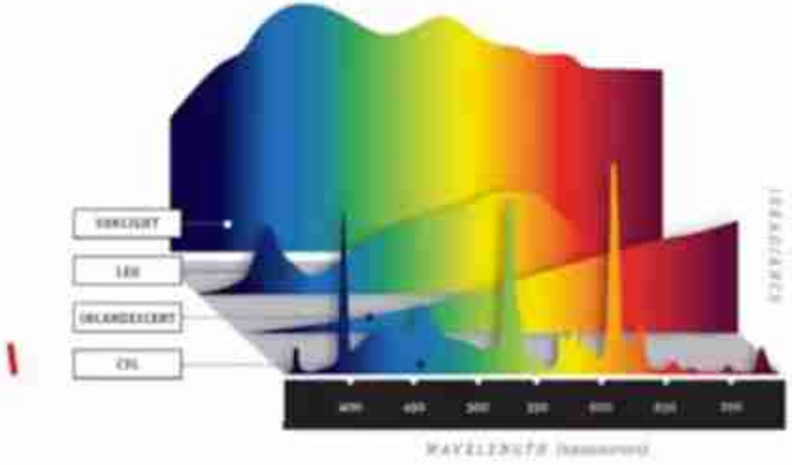
Show all

# Light

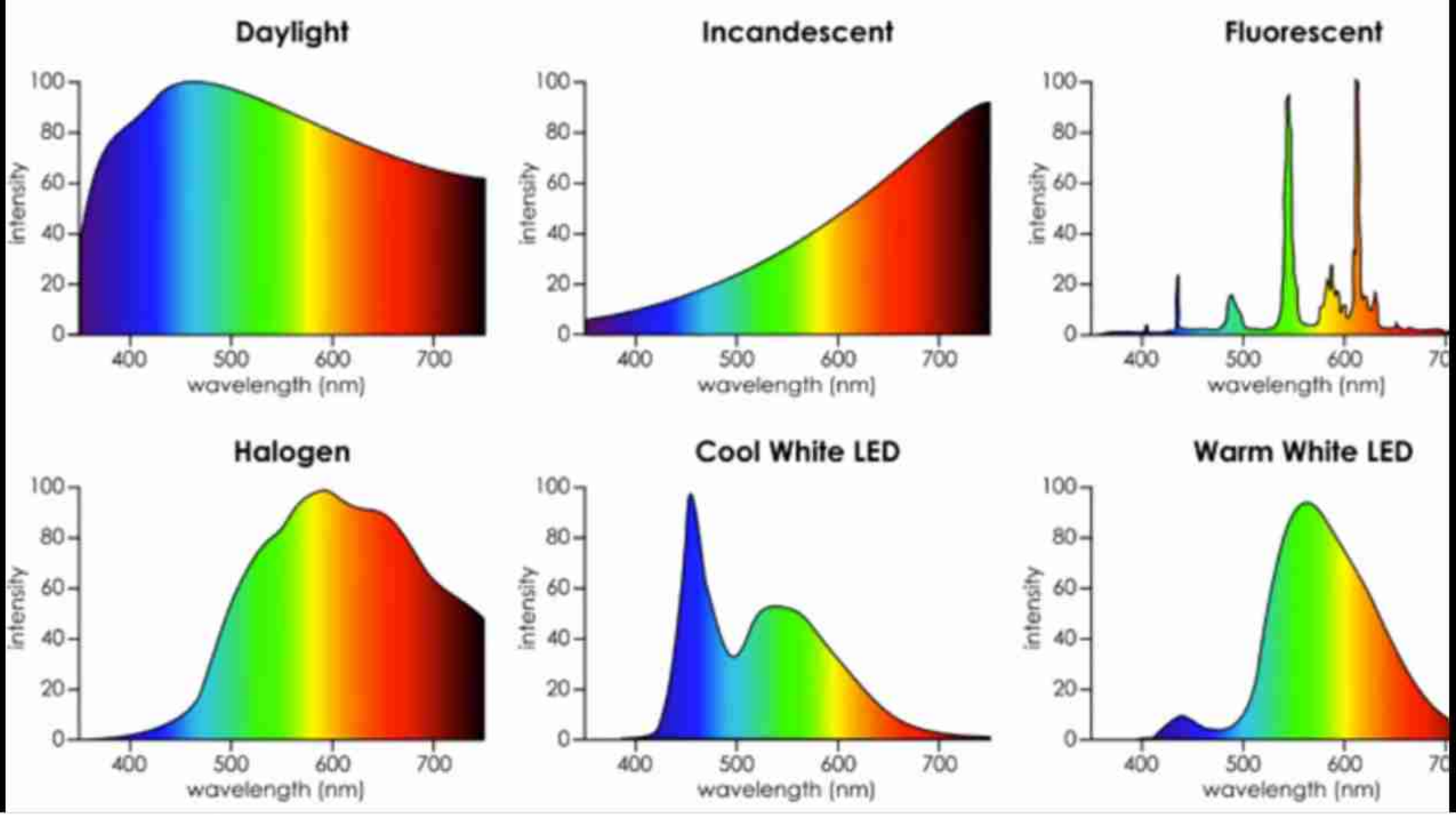



cameraman.png
Show all

## Spectral Power Distribution (SPD)

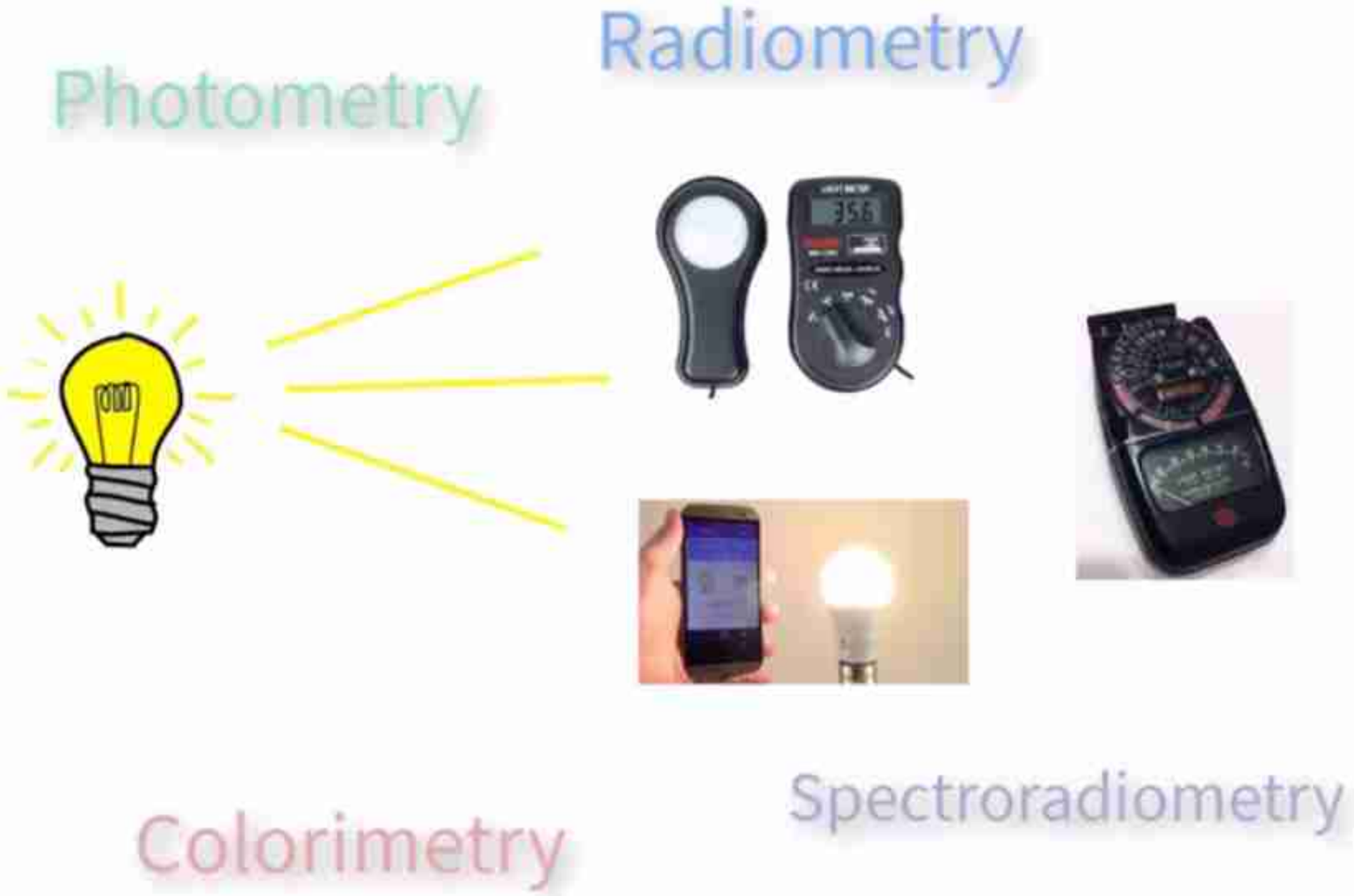



Courtesy: https://www.popularmechanics.com/



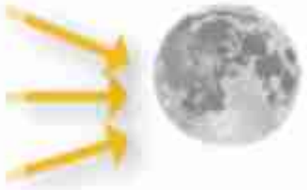

cameraman.png
Show all

cameraman.png
Show all




cameraman.png
Show all

cameraman.png
Show all

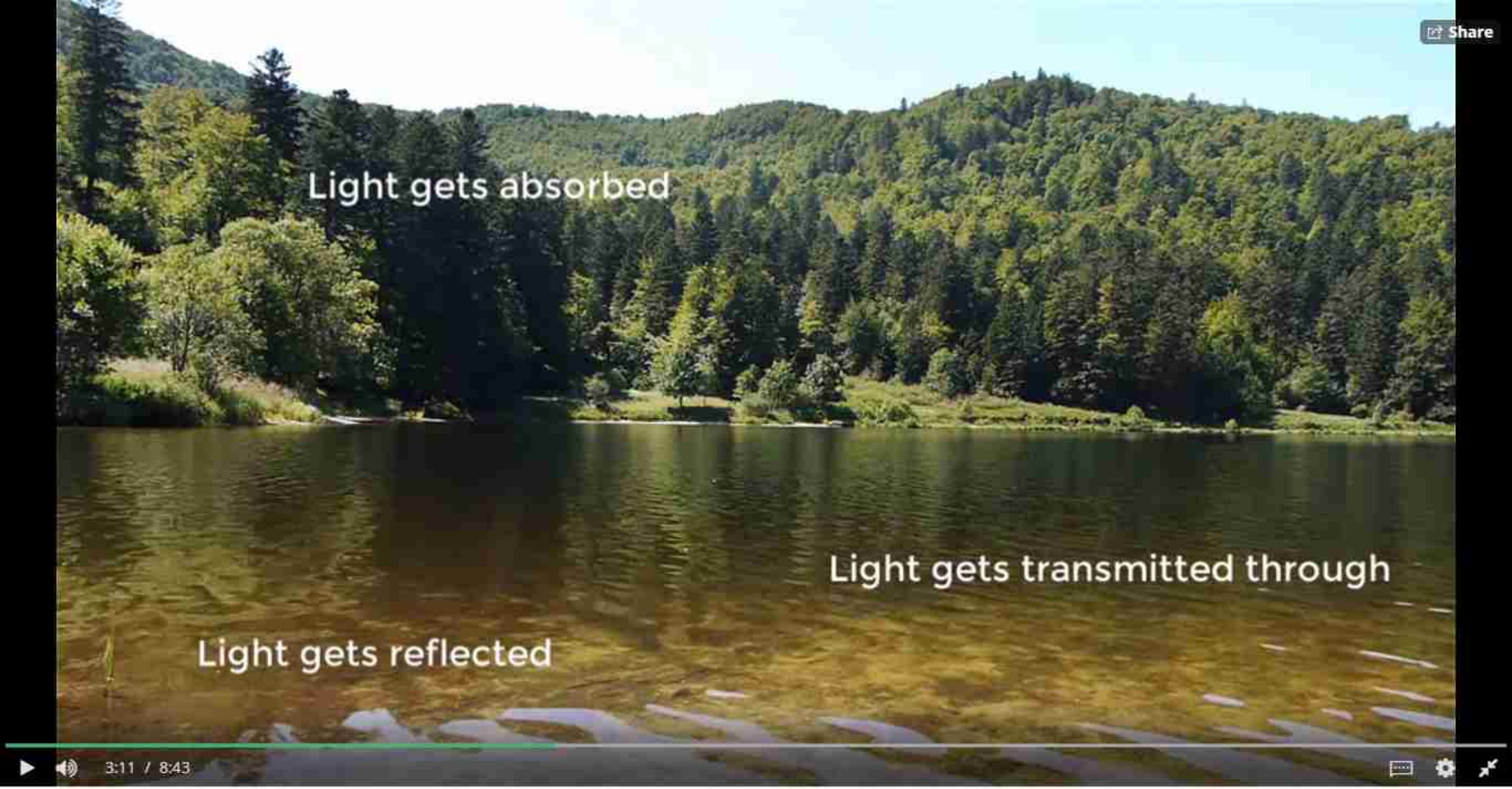


Irradiance
Illuminance

cameraman.png
Show all

cameraman.png

Show all

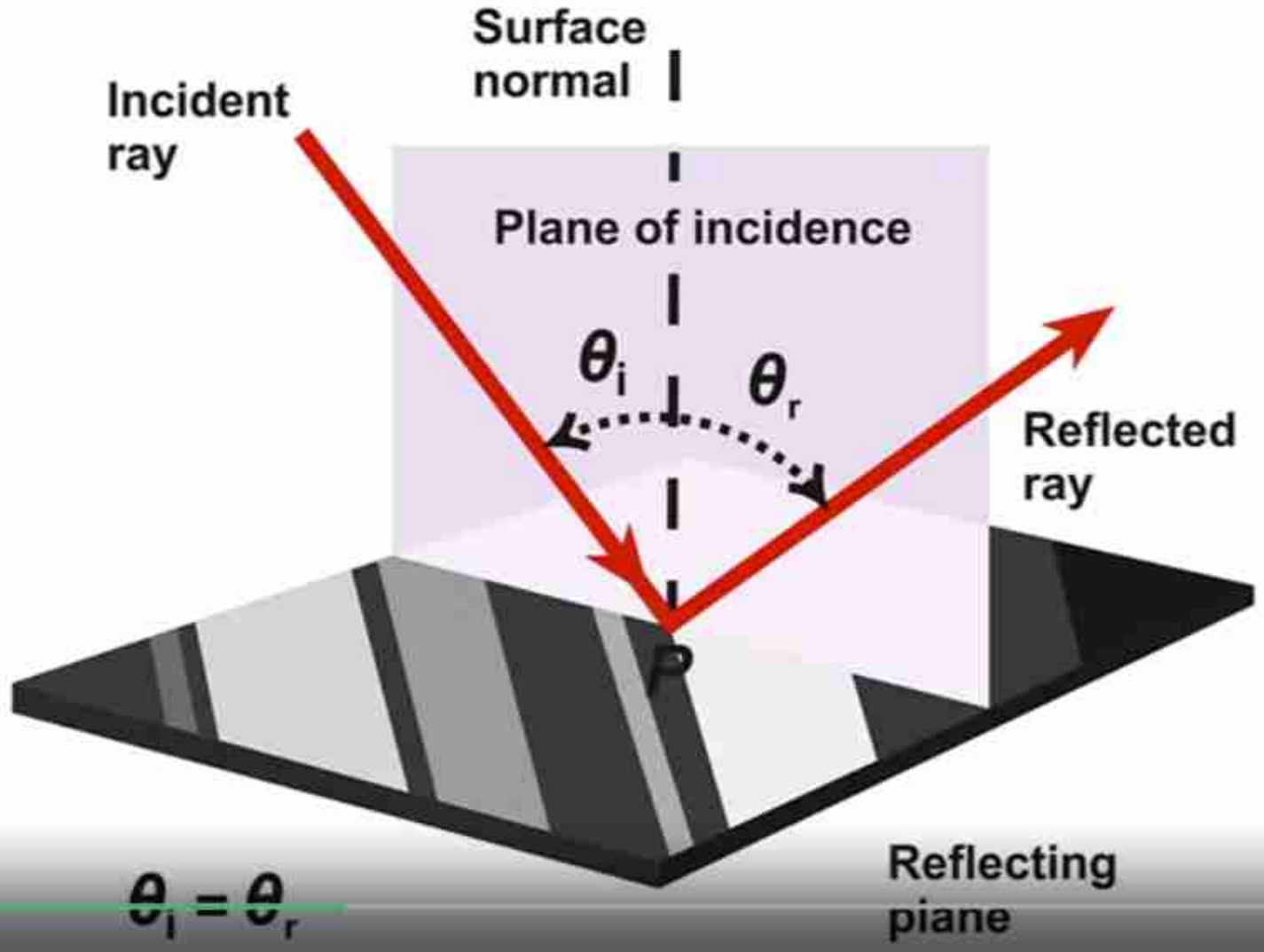



cameraman.png

Show all

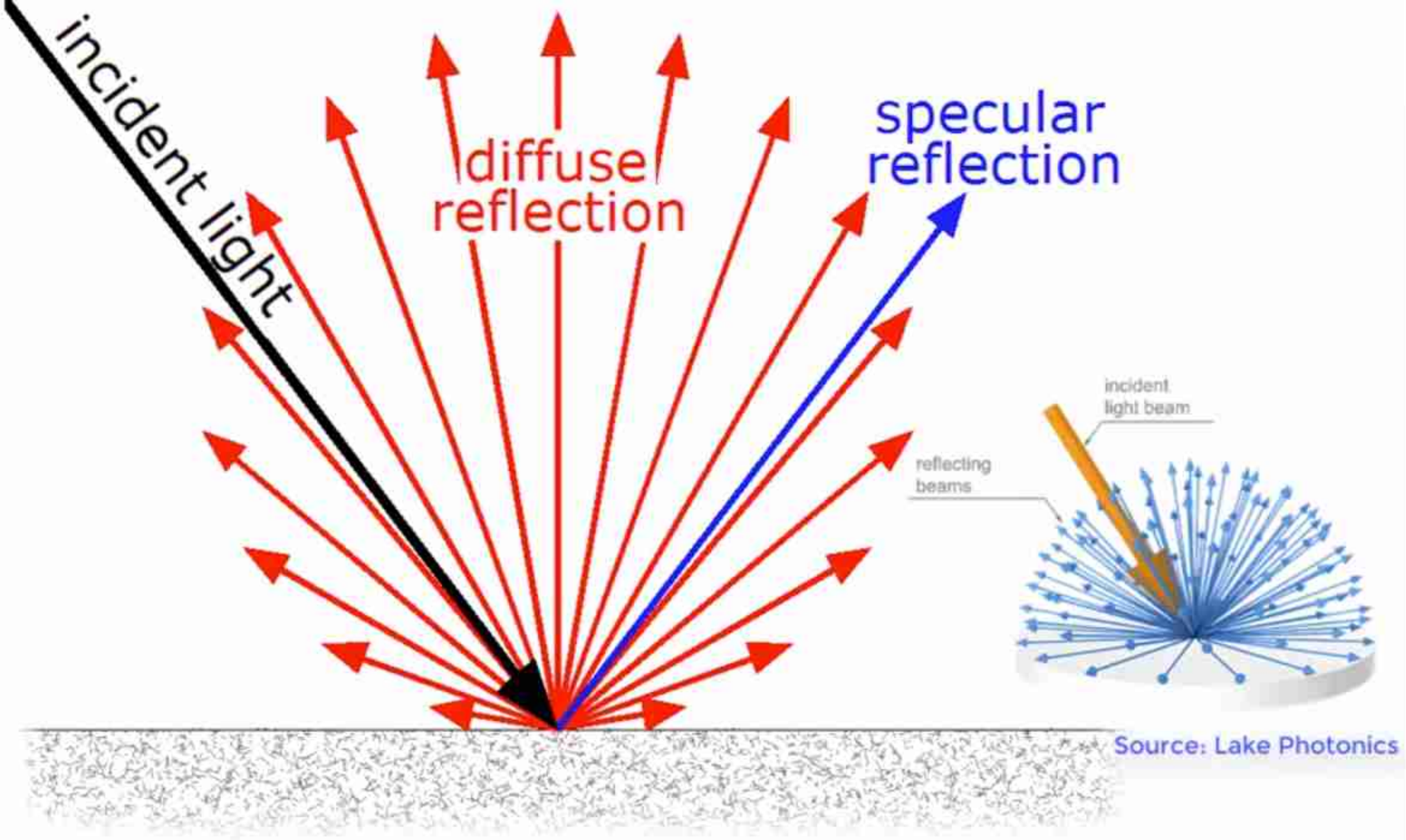
# The Reflection Law

(Reflection from a Flat Mirror)

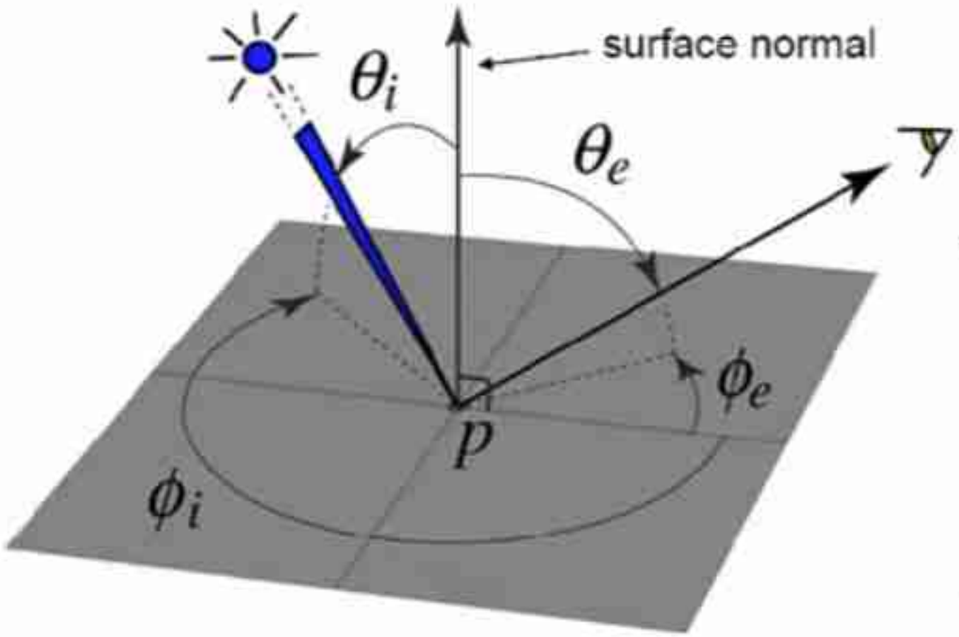



$\theta_i = \theta_r$

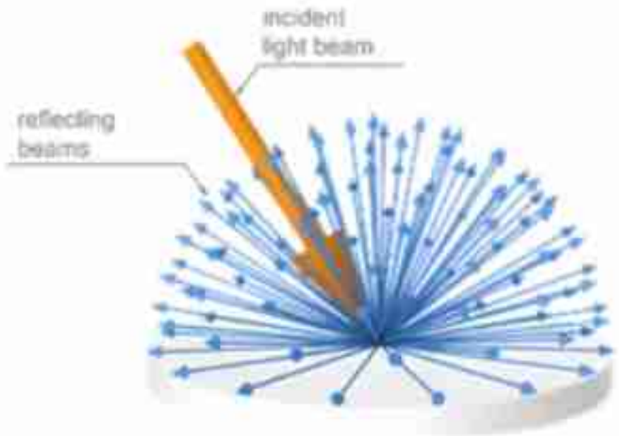
Share
3:15 / 8:43
cameraman.png
Show all


incident light
diffuse reflection
specular reflection
incident light beam
reflecting beams
Source: Lake Photonics

incoming ray $(\theta_i, \phi_i)$

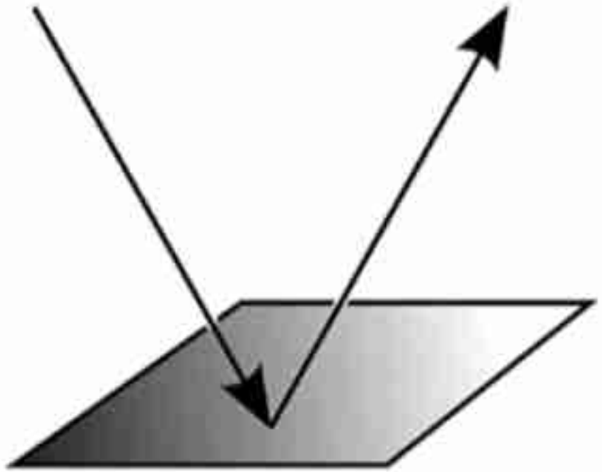



outgoing ray $(\theta_e, \phi_e)$



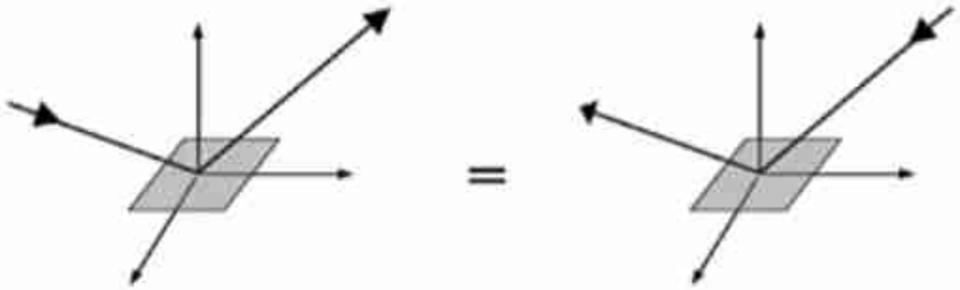

Source: Lake Photonics

Share

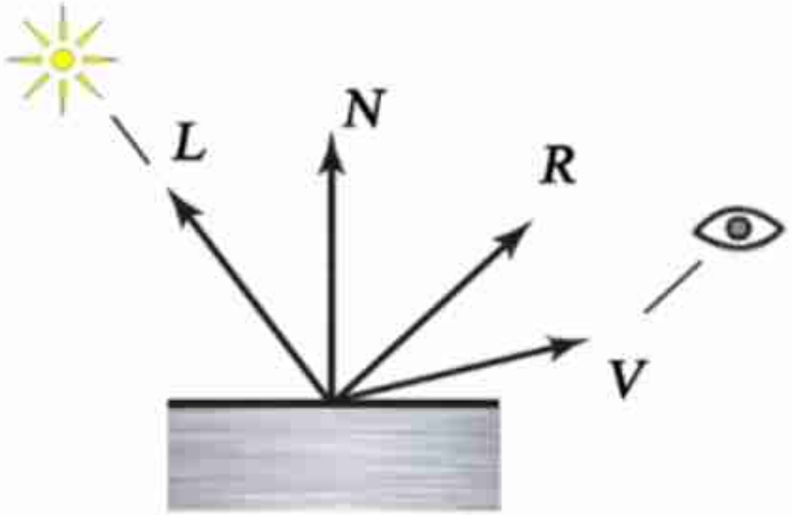


Outgoing light energy is less than or equal to incident light energy

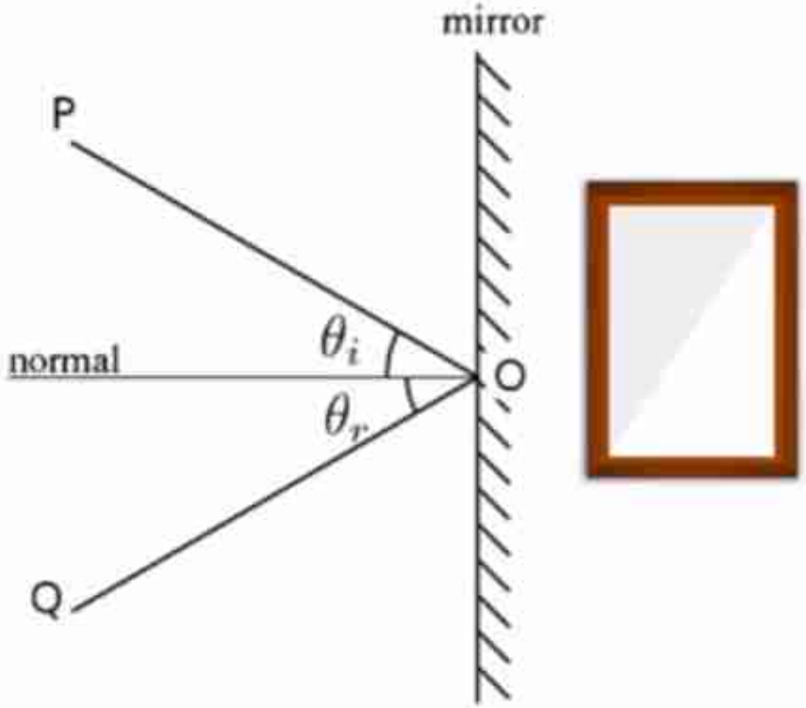
Reversing the incident light direction results in the same reflectance

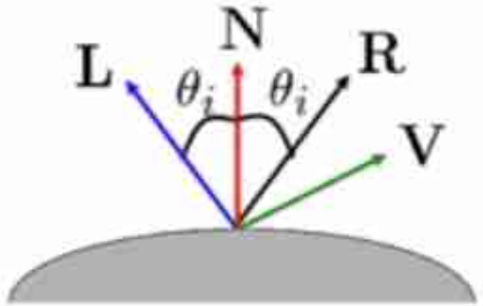


4:56 / 8:43

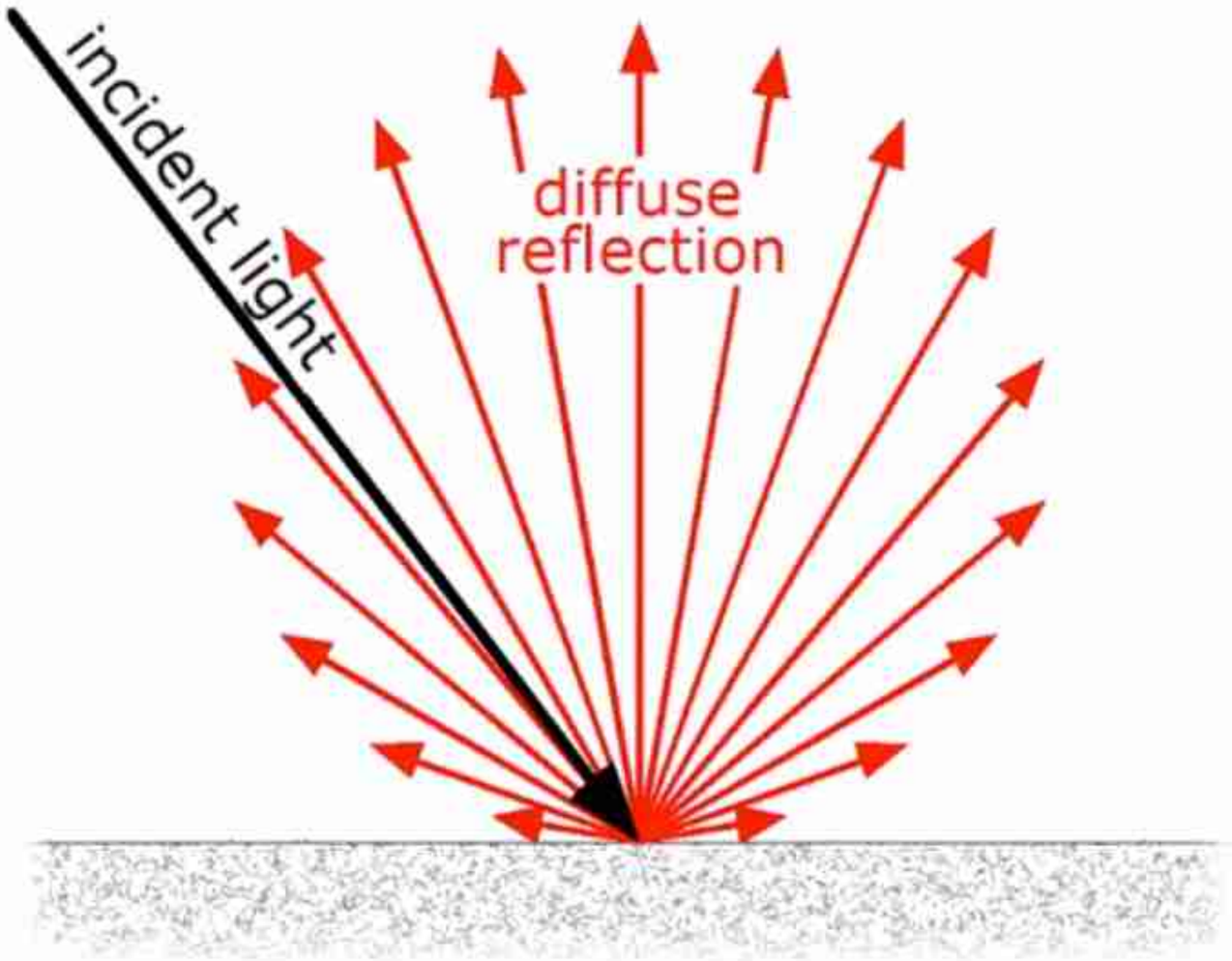



**L**, **N**, **V** unit vectors

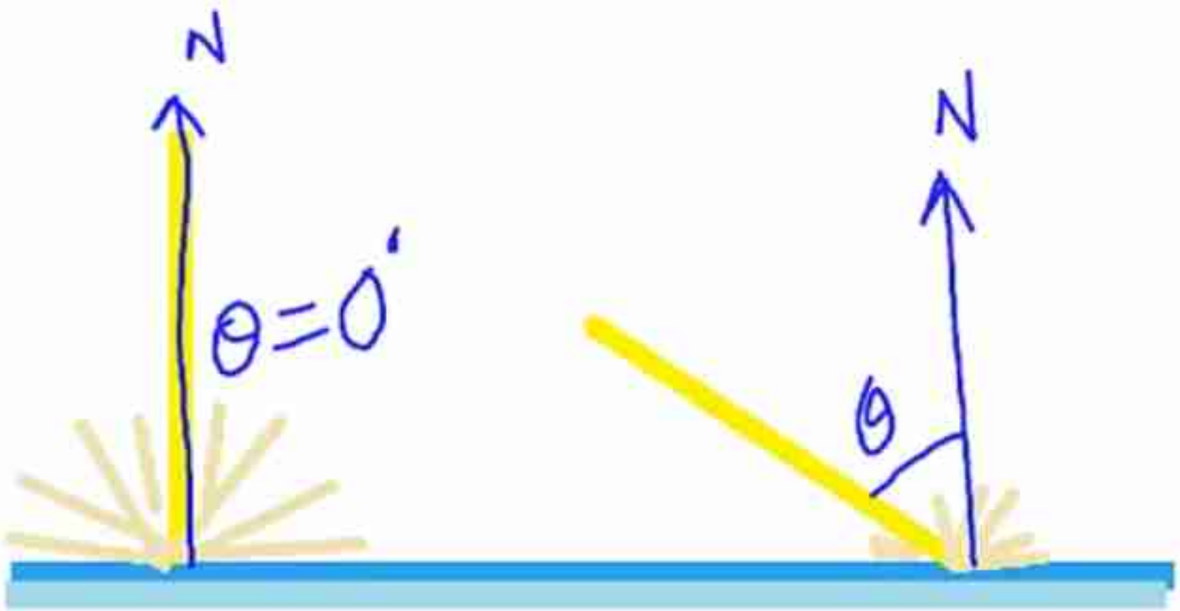



$$I_e = \begin{cases} I_i & \text{if } \mathbf{V} = \mathbf{R} \\ 0 & \text{otherwise} \end{cases}$$



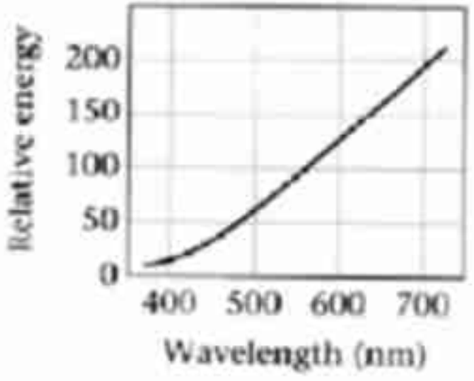

Lambertian Surface




define the diffuse
reflection model

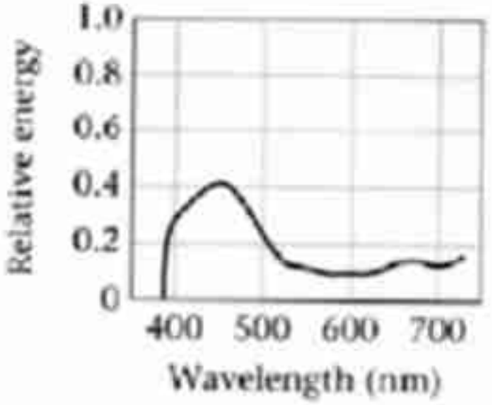
Share
6:11 / 8:43

## Spectral Irradiance



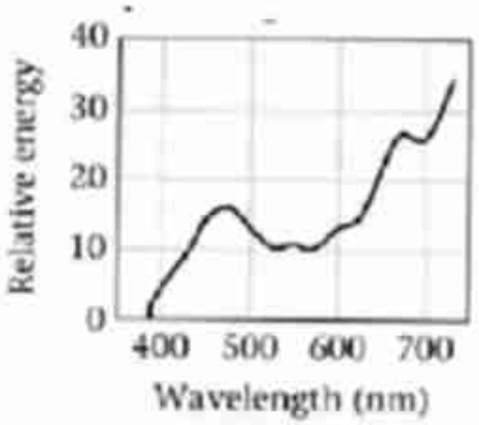

.*

## Spectral Albedo




=

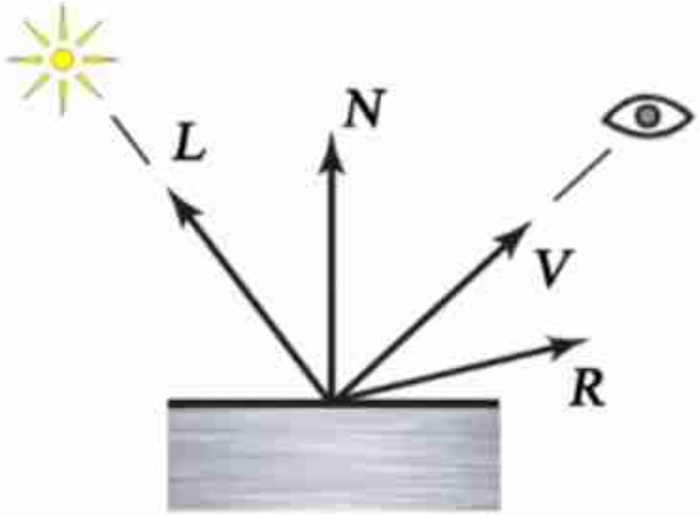
## Spectral Radiance




Foundations of Vision, by Brian Wandell, Sinauer Assoc., 1995

The albedo value may vary with

We do have ambient light

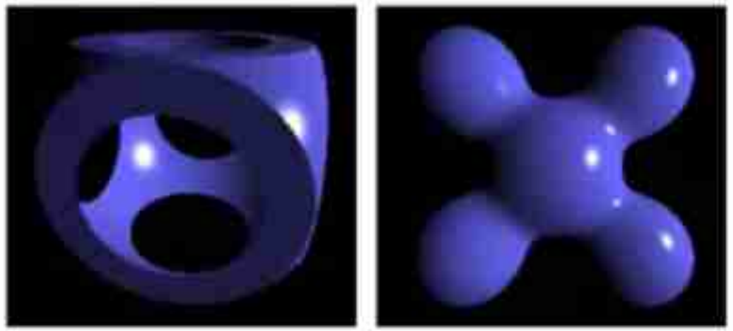


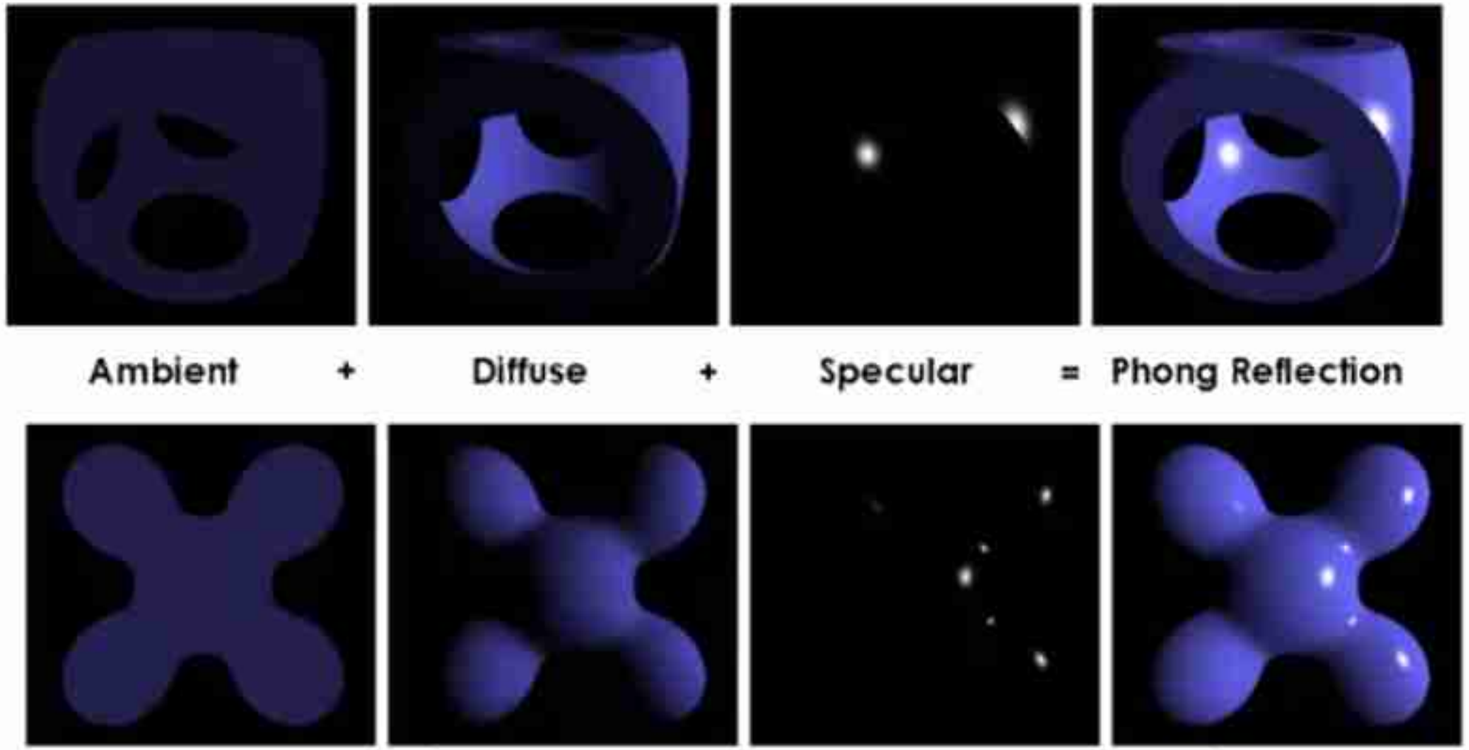

**L**, **N**, **V** unit vectors

$$I_e = k_a I_a + I_i \left[ k_d(\mathbf{N} \cdot \mathbf{L}) + k_s(\mathbf{V} \cdot \mathbf{R})^{n_s} \right]$$

$I_e$ = outgoing radiance
$I_i$ = incoming radiance
$I_a$ = ambient light
$k_a$ = ambient light reflectance factor

So this is the holistic illumination model,



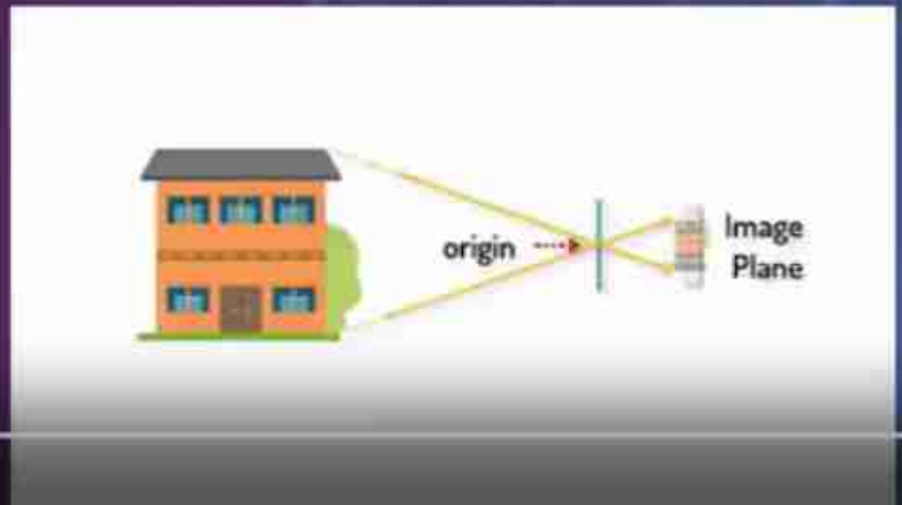



these three components
can be separated,

# Pinhole Camera Model

# Pinhole Camera Model

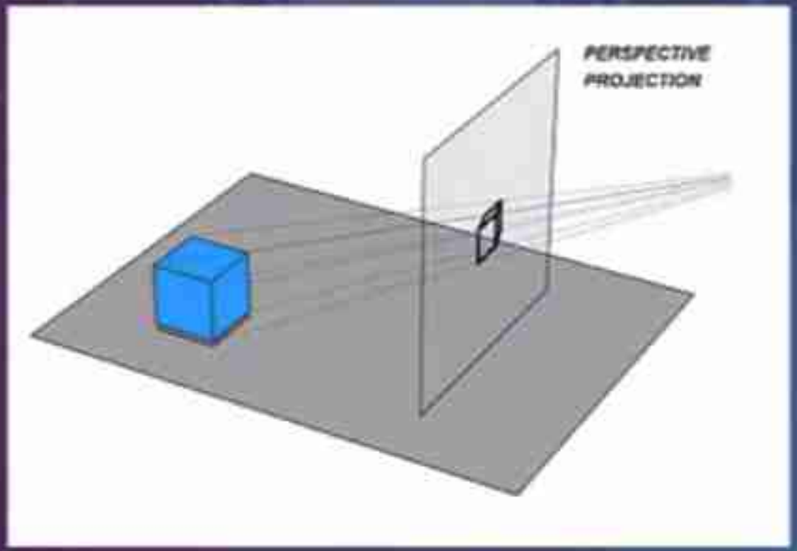
light passes through a vanishingly small pinhole placed at the origin and illuminates an image plane beneath it

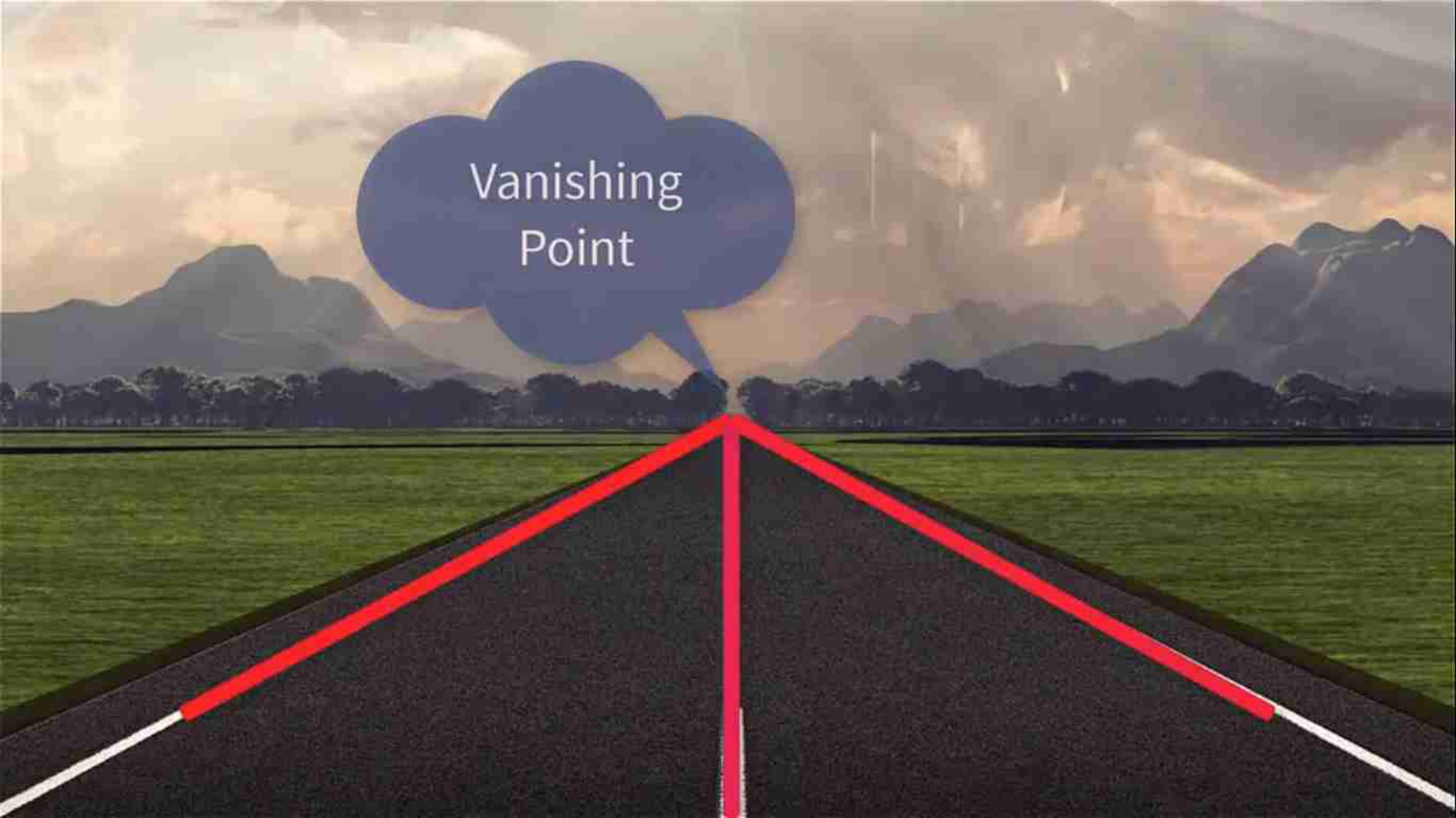



Share
0:27 / 9:03

# Perspective Projection

geometric mapping from 3D to 2D

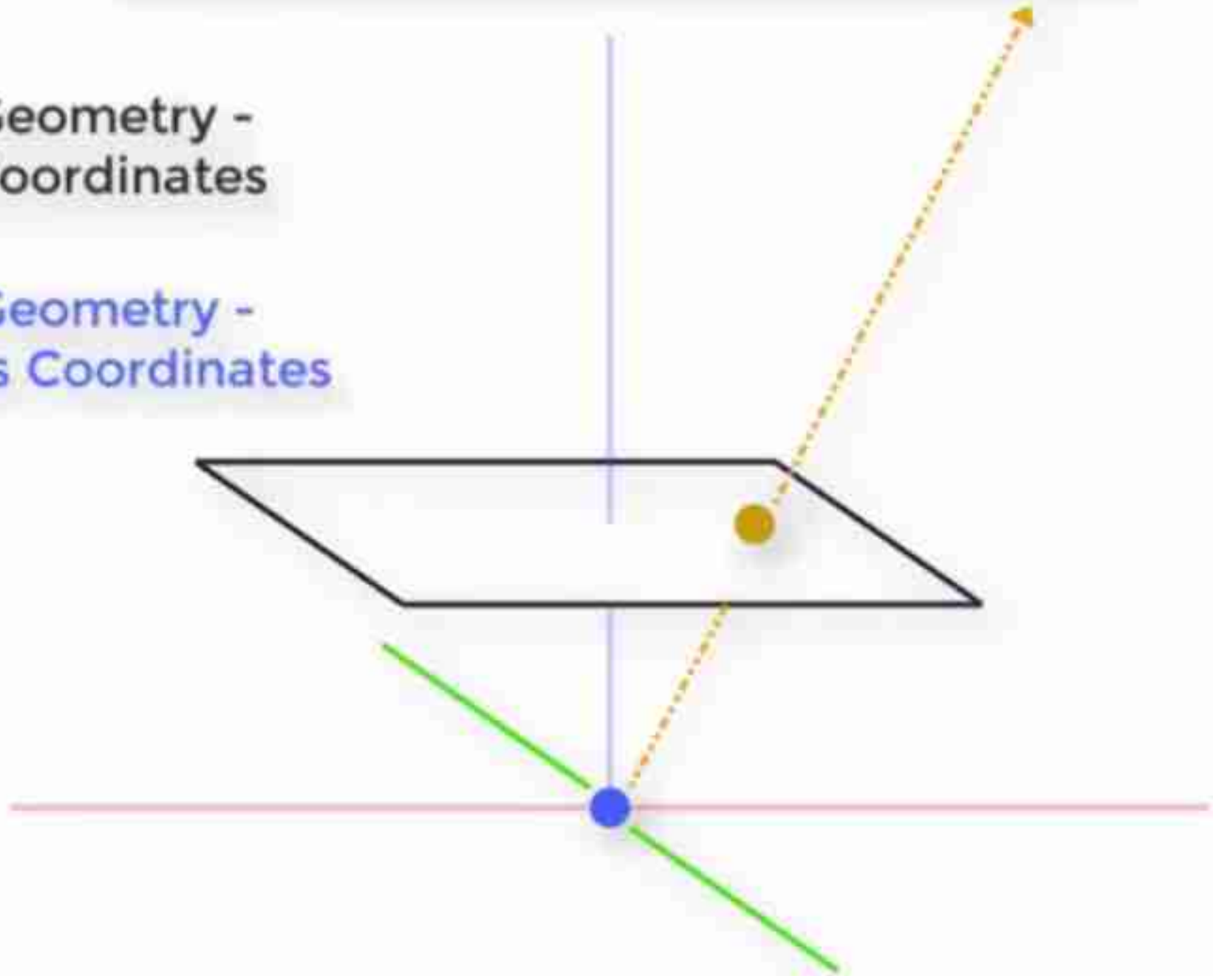



Vanishing
Point

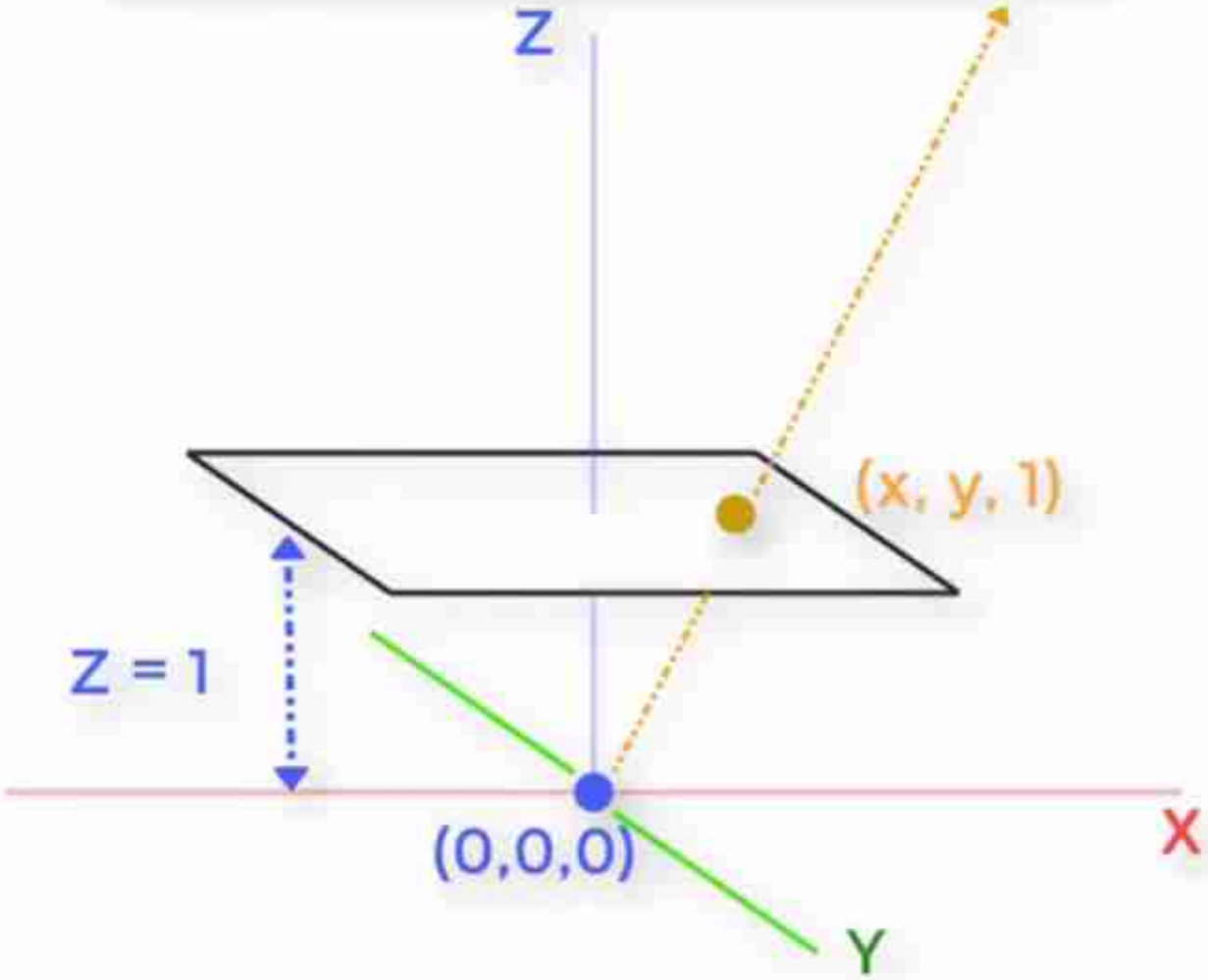
# Homogeneous Coordinates

Euclidean Geometry -
Cartesian Coordinates

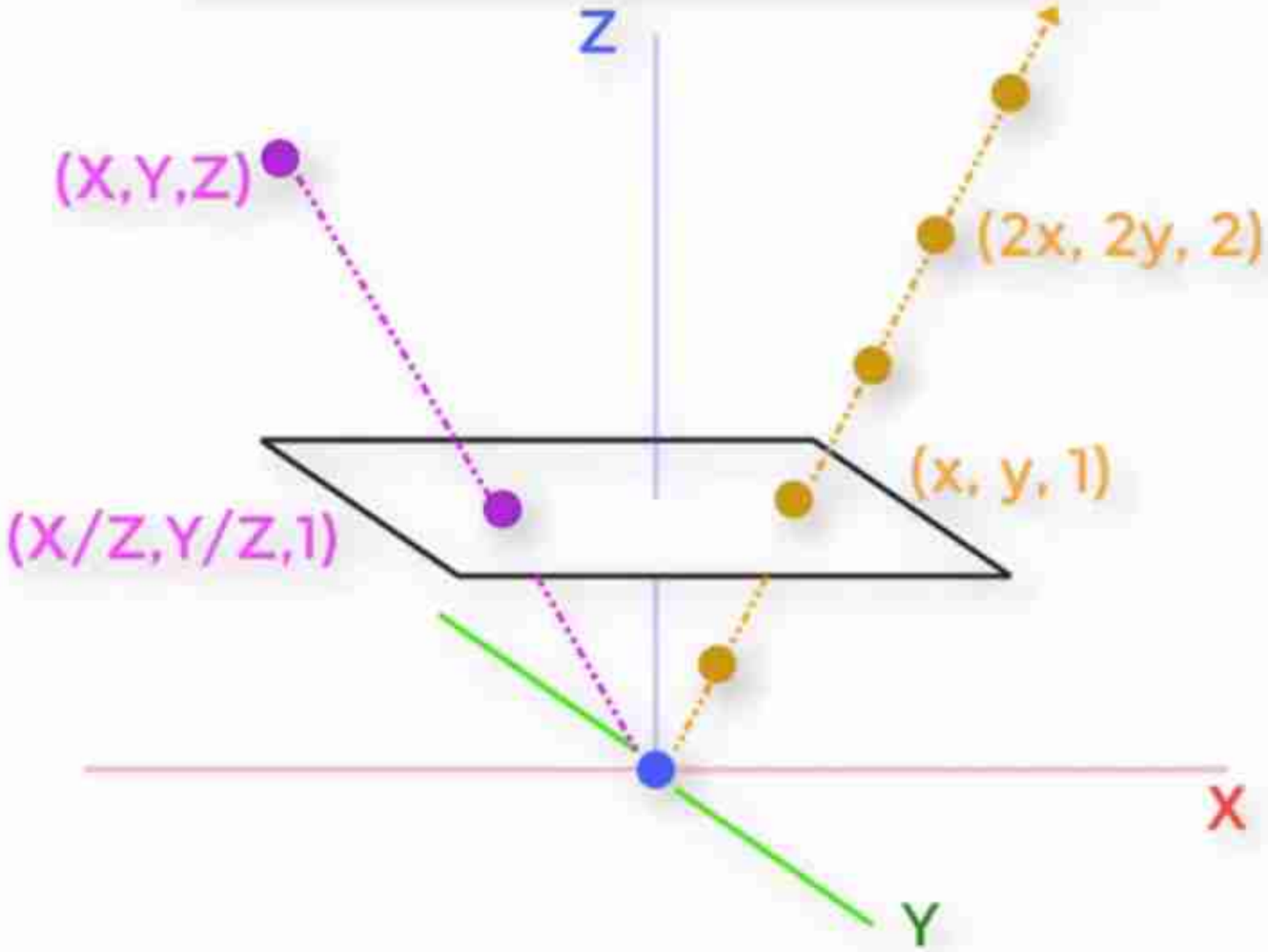
Projective Geometry -
Homogeneous Coordinates



# Homogeneous Coordinates

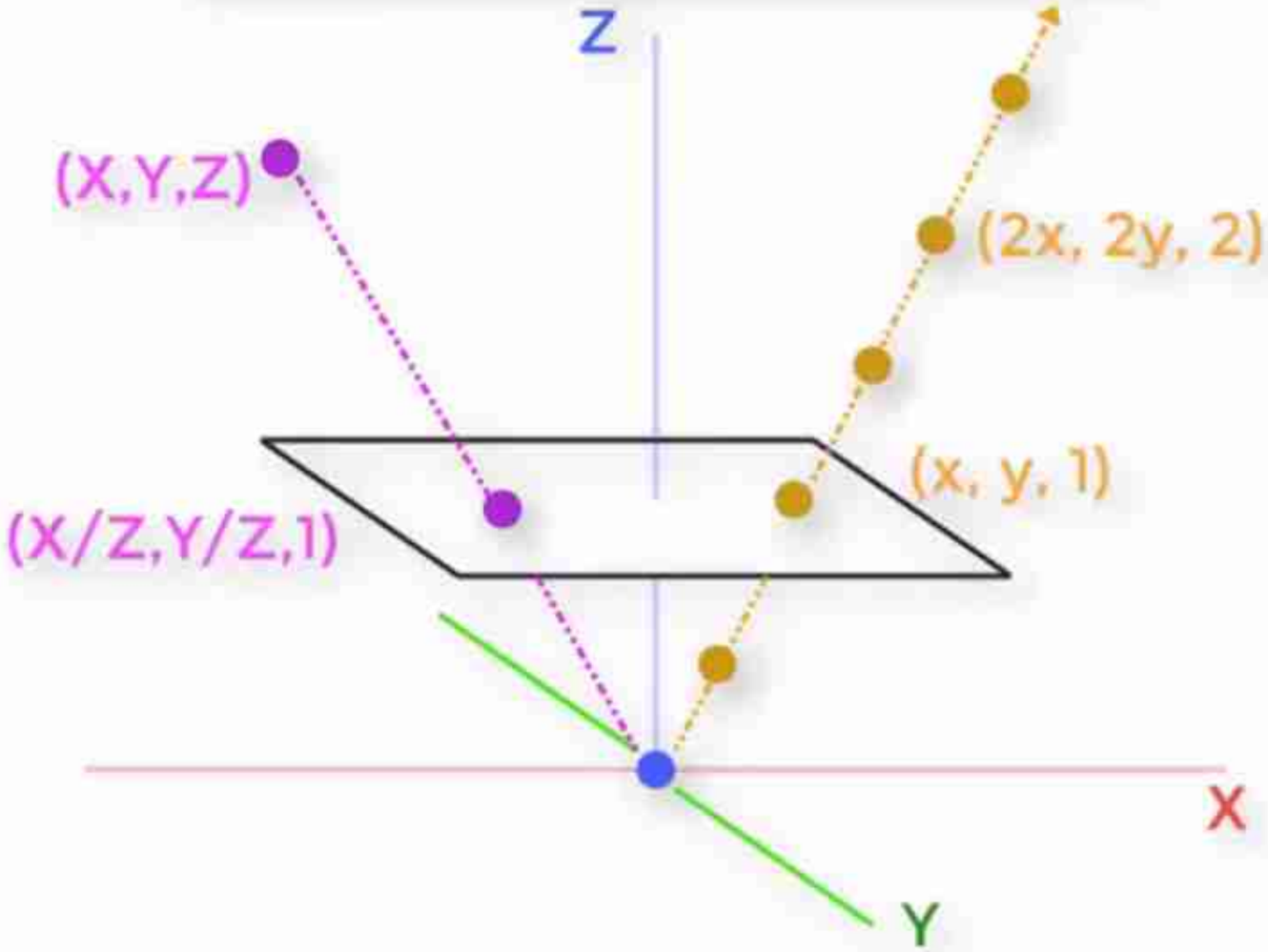


# Homogeneous Coordinates



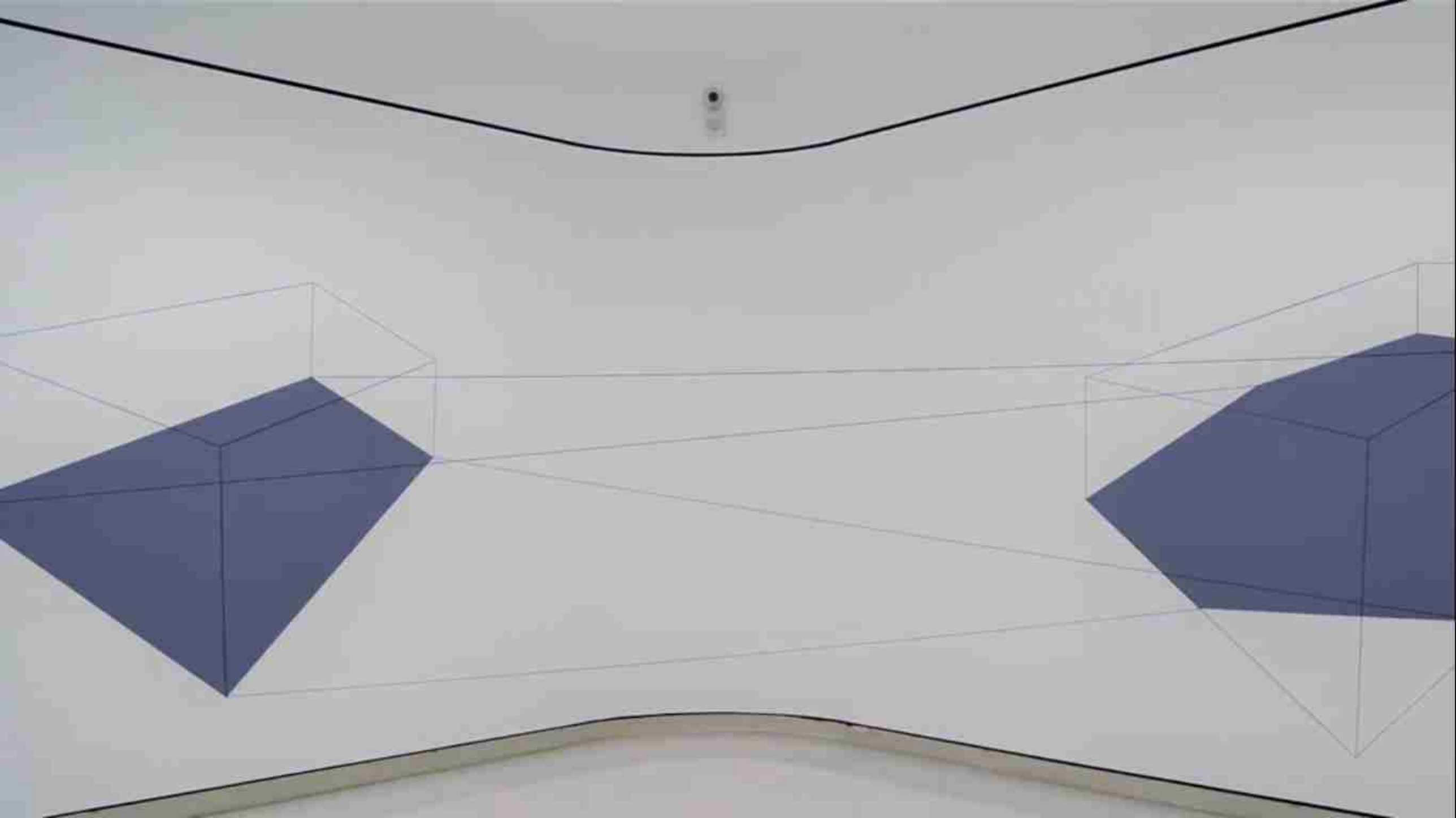

Example: Cartesian point (1,3)can be represented in homogeneous co-ordinates as (1,3,1) or (2,6,2) and so on...

# Homogeneous Coordinates

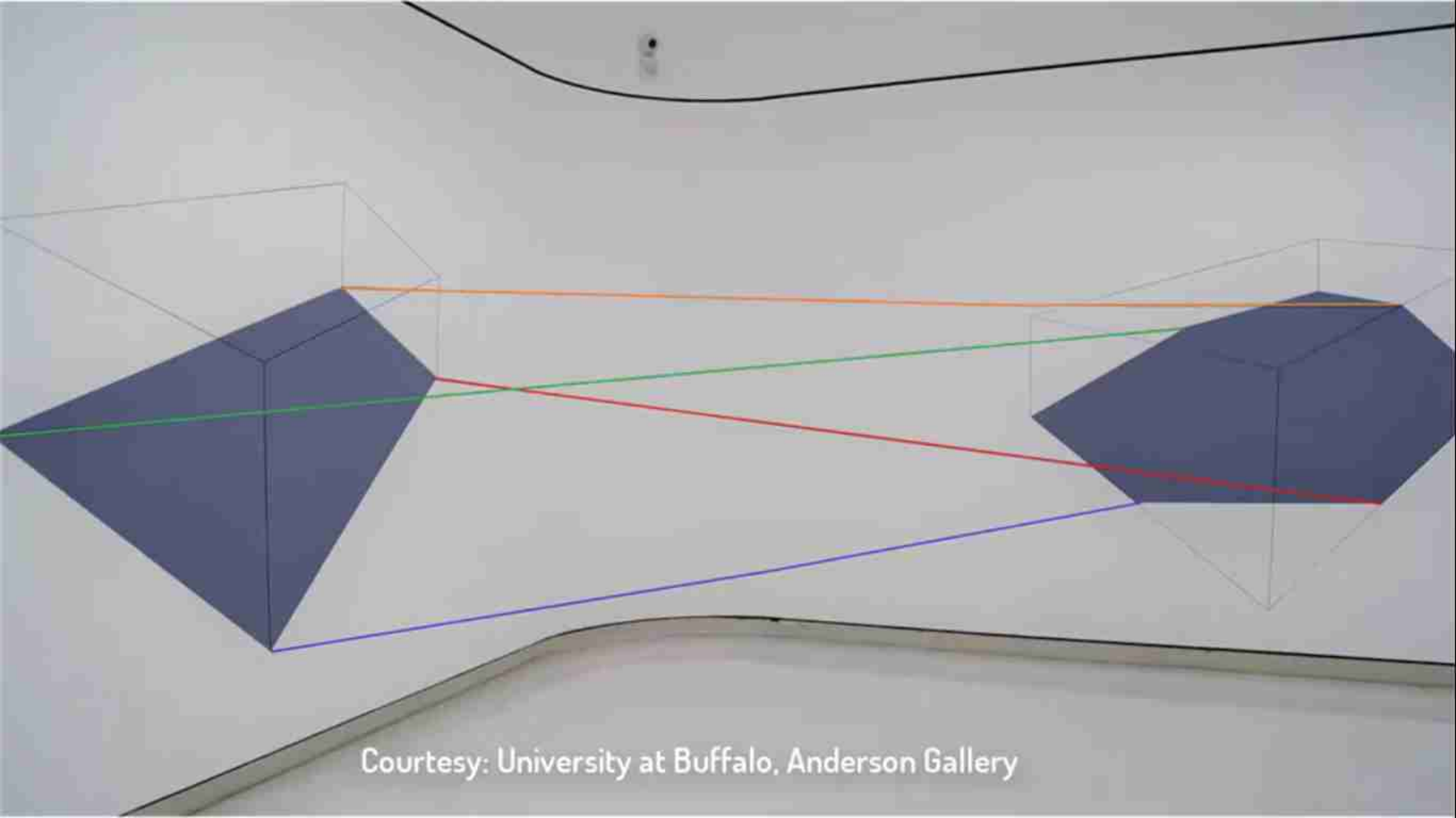



For cartesian co-ordinates (x,y) on the plane,
(x,y,1) is the set of homogeneous co-ordinates

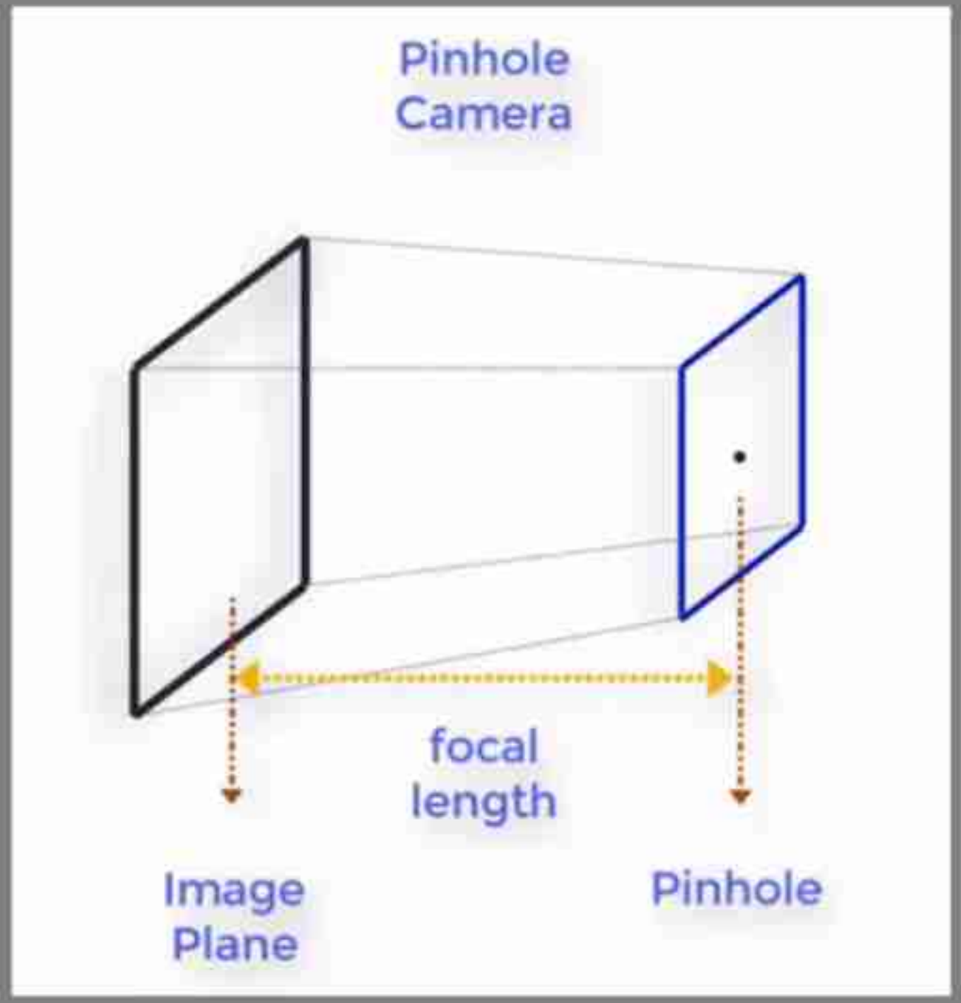
NEW



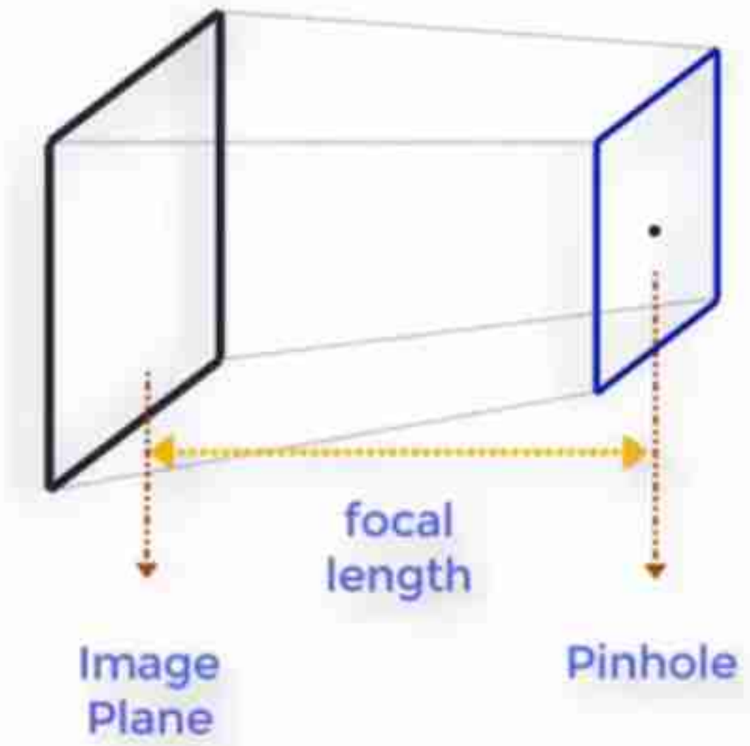
Courtesy: University at Buffalo, Anderson Gallery


Courtesy: University at Buffalo, Anderson Gallery

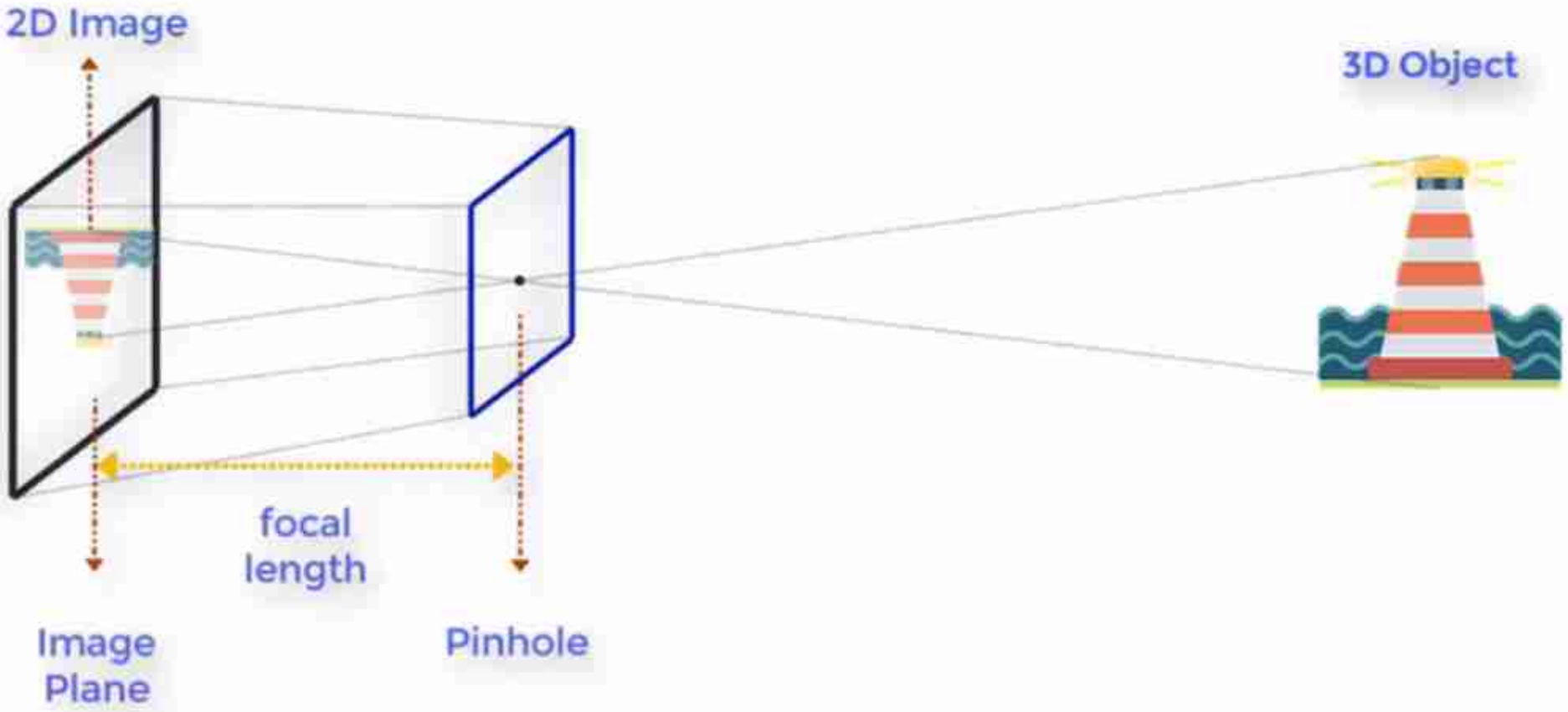
# Pinhole Camera Model



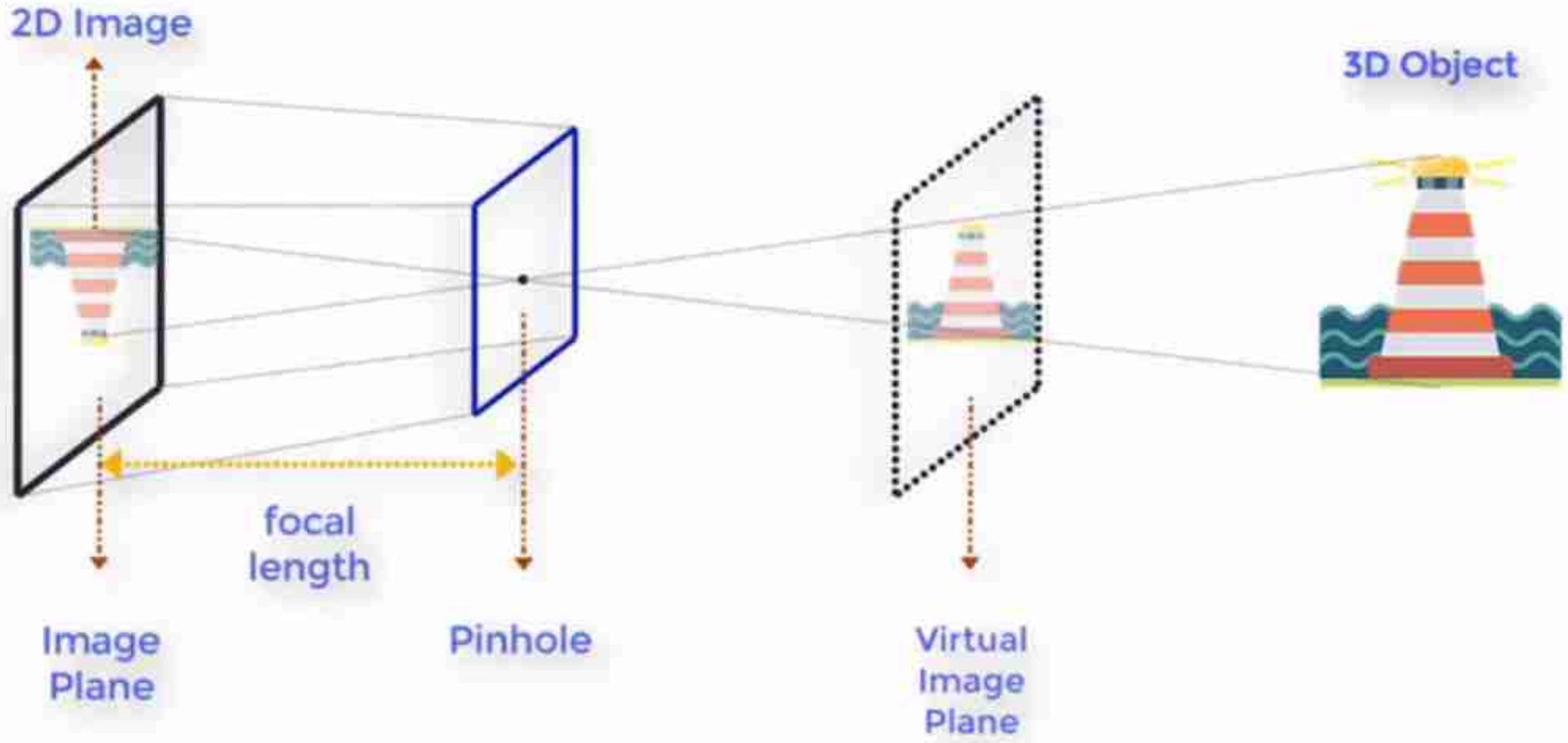
# Pinhole Camera Model

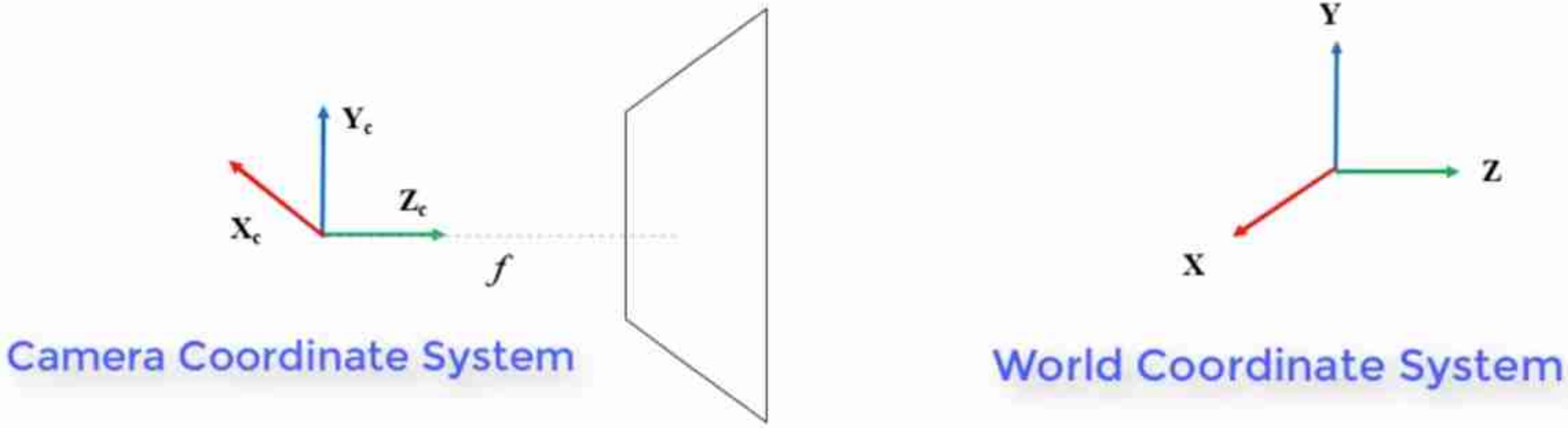


# Pinhole Camera Model



# Pinhole Camera Model




$Y_c$
$Z_c$
$X_c$
$f$
Camera Coordinate System
Y
Z
X
World Coordinate System

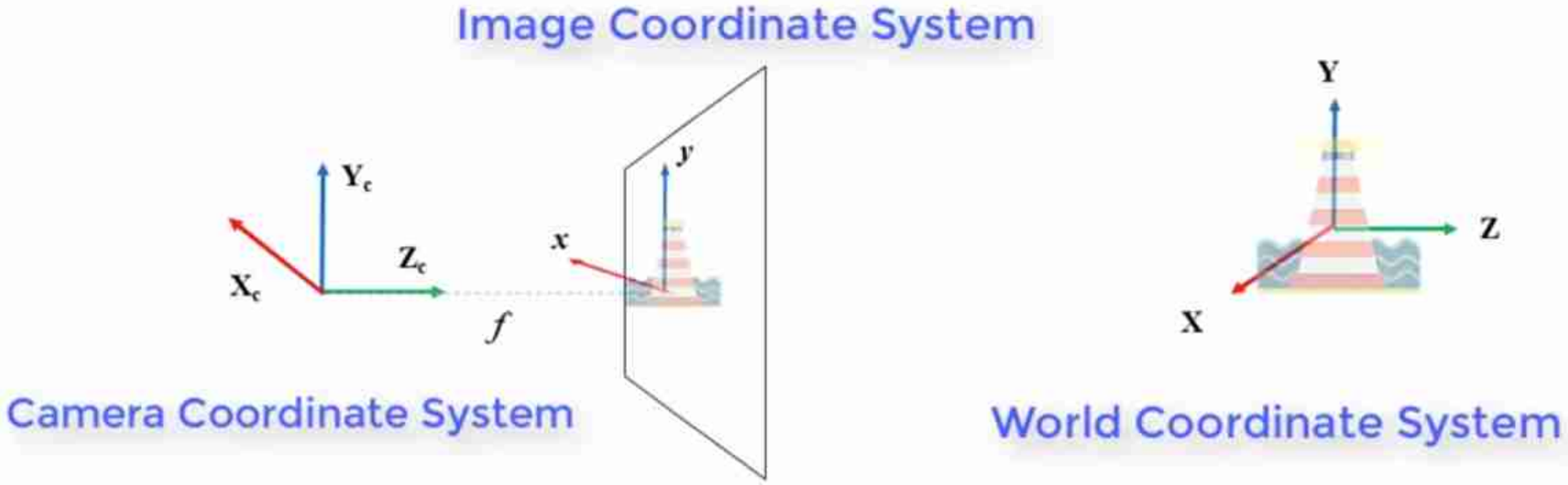


Image Coordinate System
$Y_c$
$Z_c$
$X_c$
$f$
$y$
$x$
Camera Coordinate System
Y
Z
X
World Coordinate System

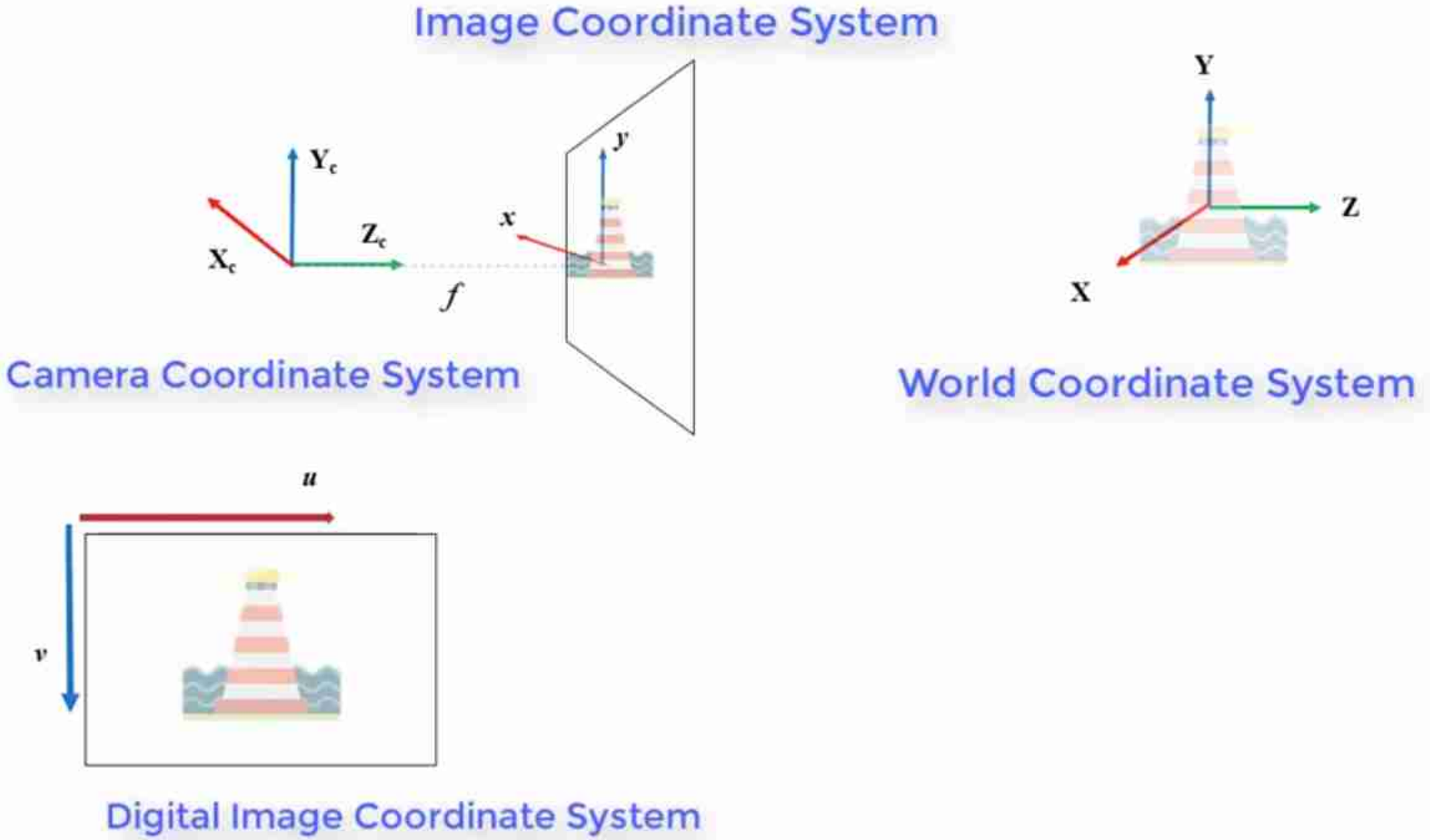


Image Coordinate System
$Y_c$
$Z_c$
$X_c$
$f$
$y$
$x$
Camera Coordinate System
Y
Z
X
World Coordinate System
$u$
$v$
Digital Image Coordinate System

# Forward Projection

| World Coordinates | → | Camera Coordinates | → | Image Plane Coordinates | → | Pixel Coordinates |
|---|---|---|---|---|---|---|
| (X,Y,Z) | | (Xc,Yc,Zc) | | (x,y) | | (u,v) |

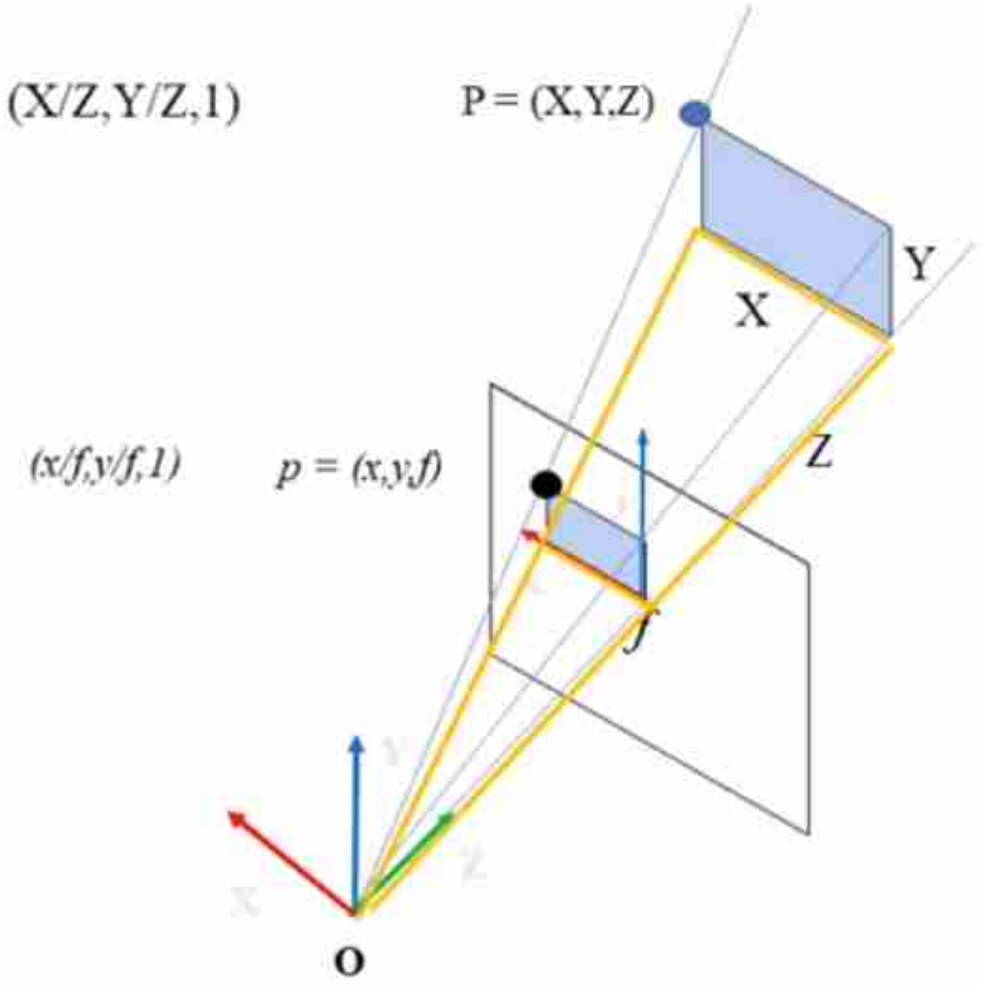
# Backward Projection

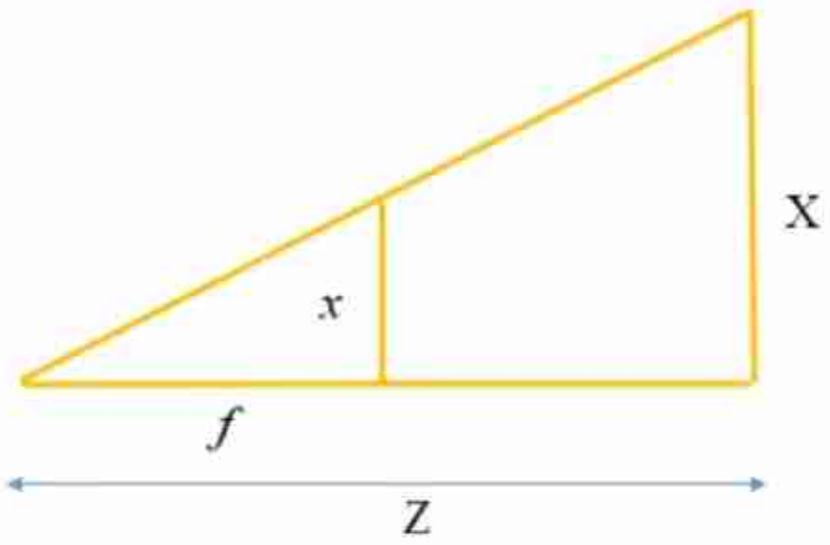
World Coordinates ← Camera Coordinates ← Image Plane Coordinates ← Pixel Coordinates

(X,Y,Z) (Xc,Yc,Zc) (x,y) (u,v)

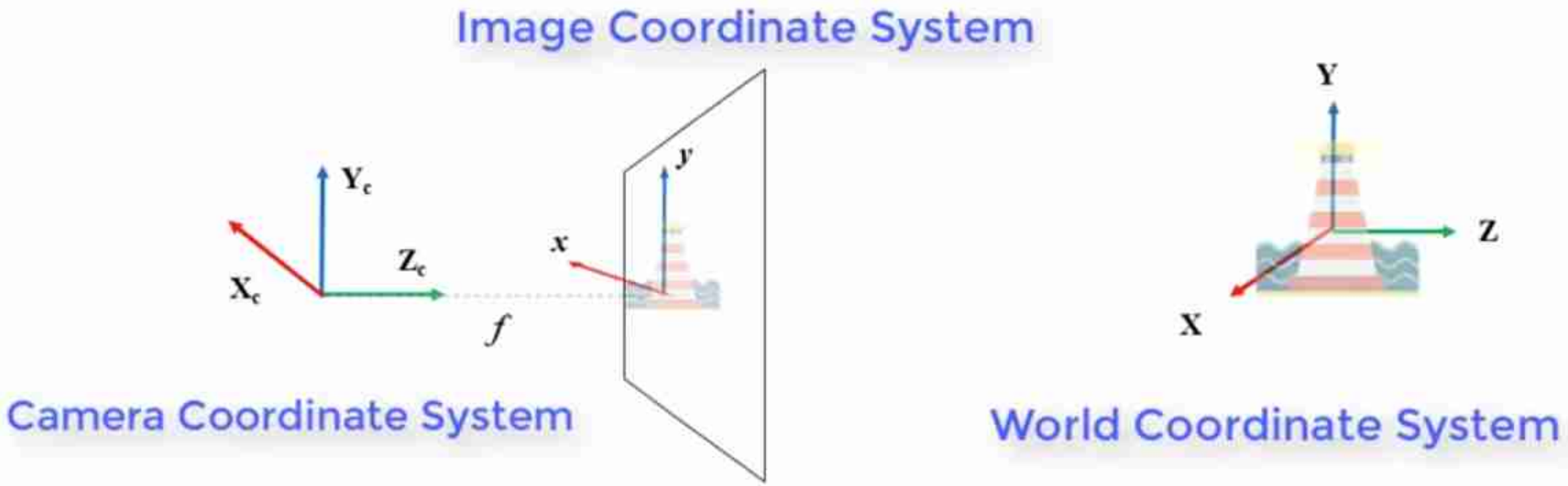
$$\boxed{\frac{x}{f} = \frac{X}{Z}}$$

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



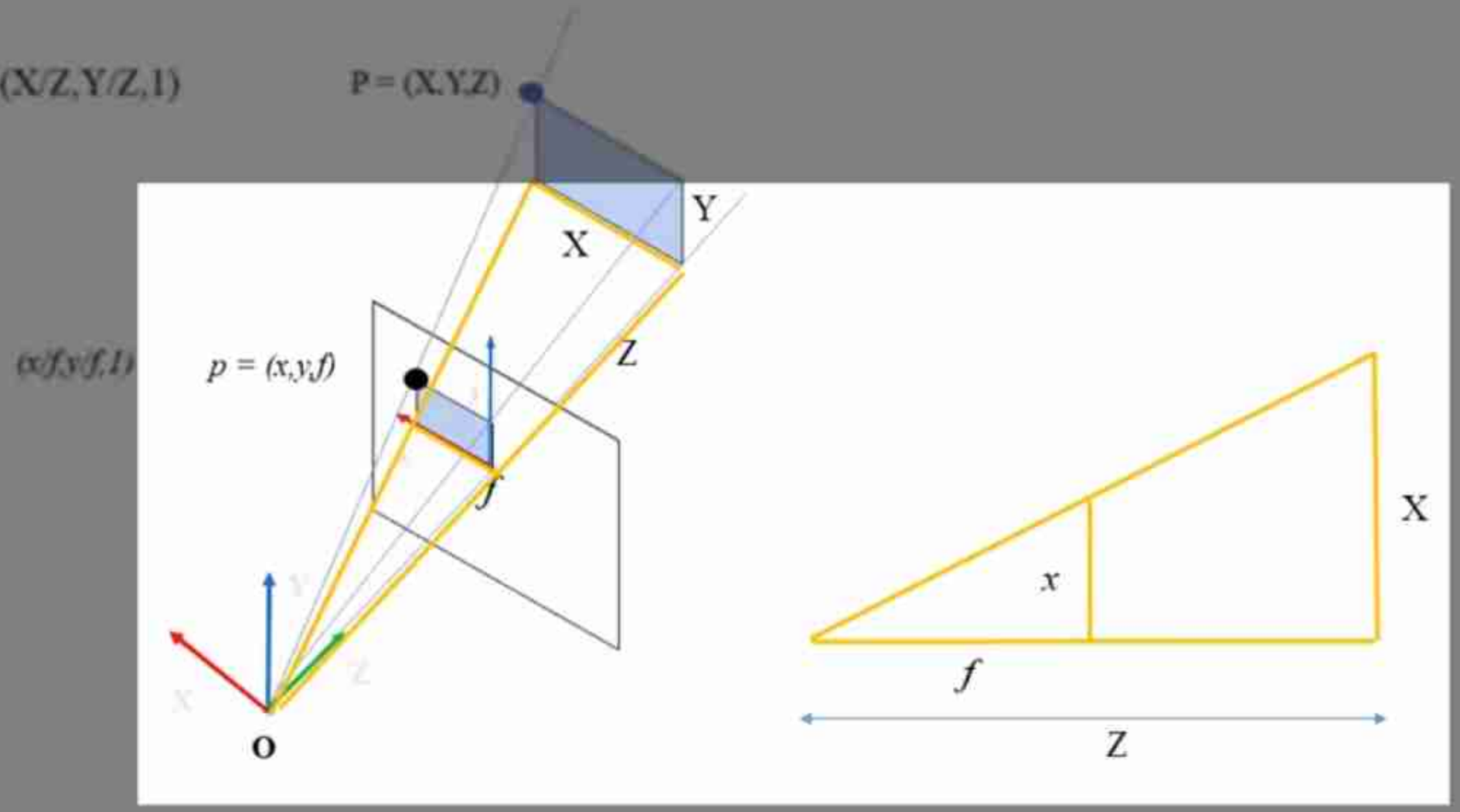
Share




4:37 / 9:03

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

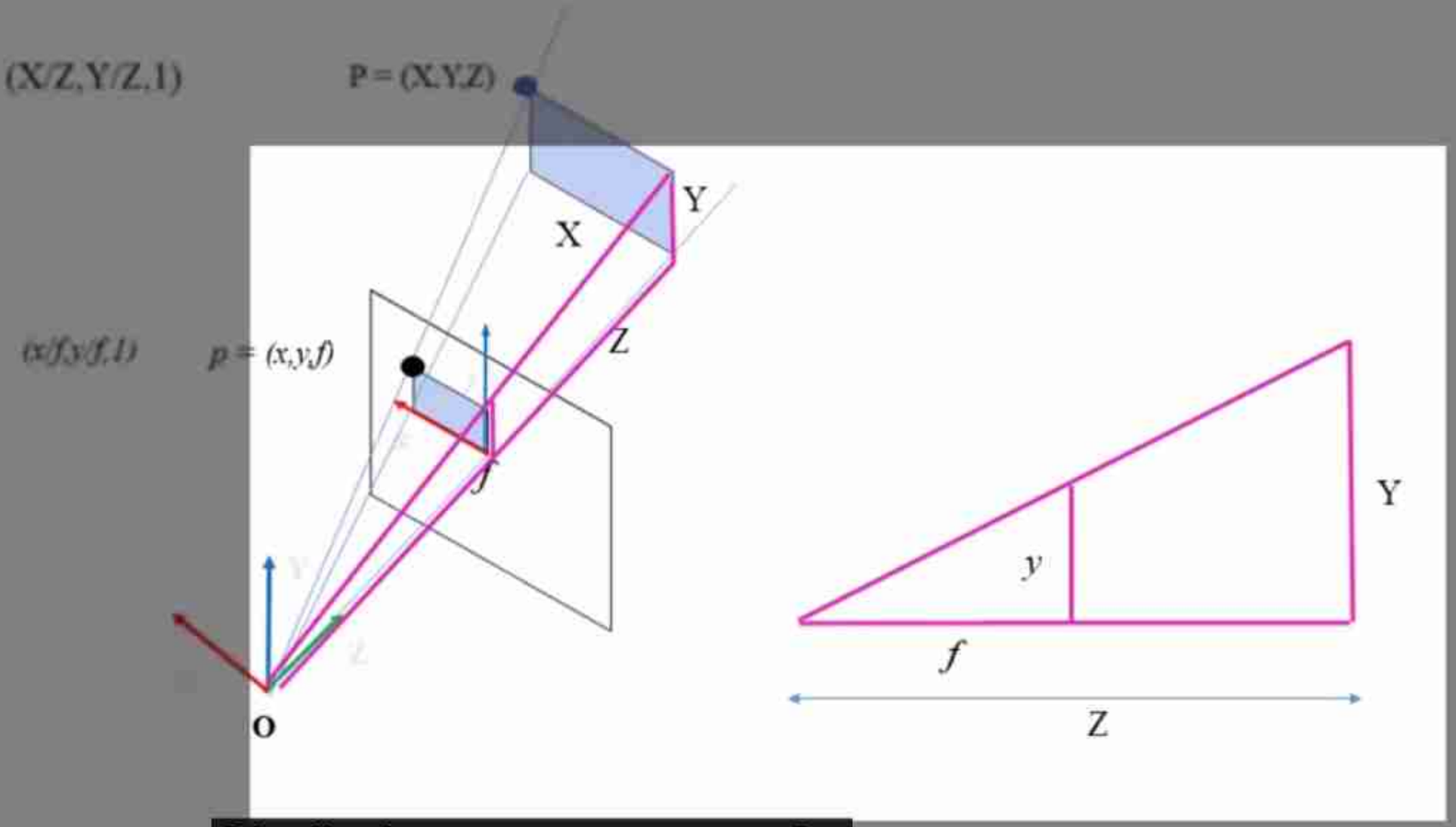
$$\frac{y}{f} = \frac{Y}{Z}$$




The highlighted right angle triangle is similar to this bigger

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

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



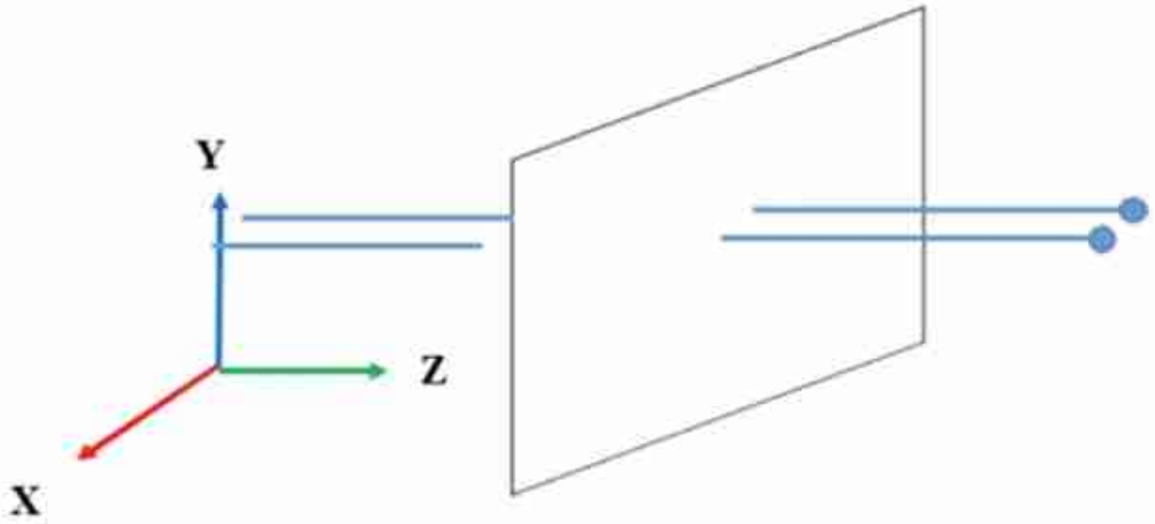

Similarly, we can prove for Y by f as well.

Share
7:29 / 9:03

# Orthographic Projection

$$x = X$$

$$y = Y$$



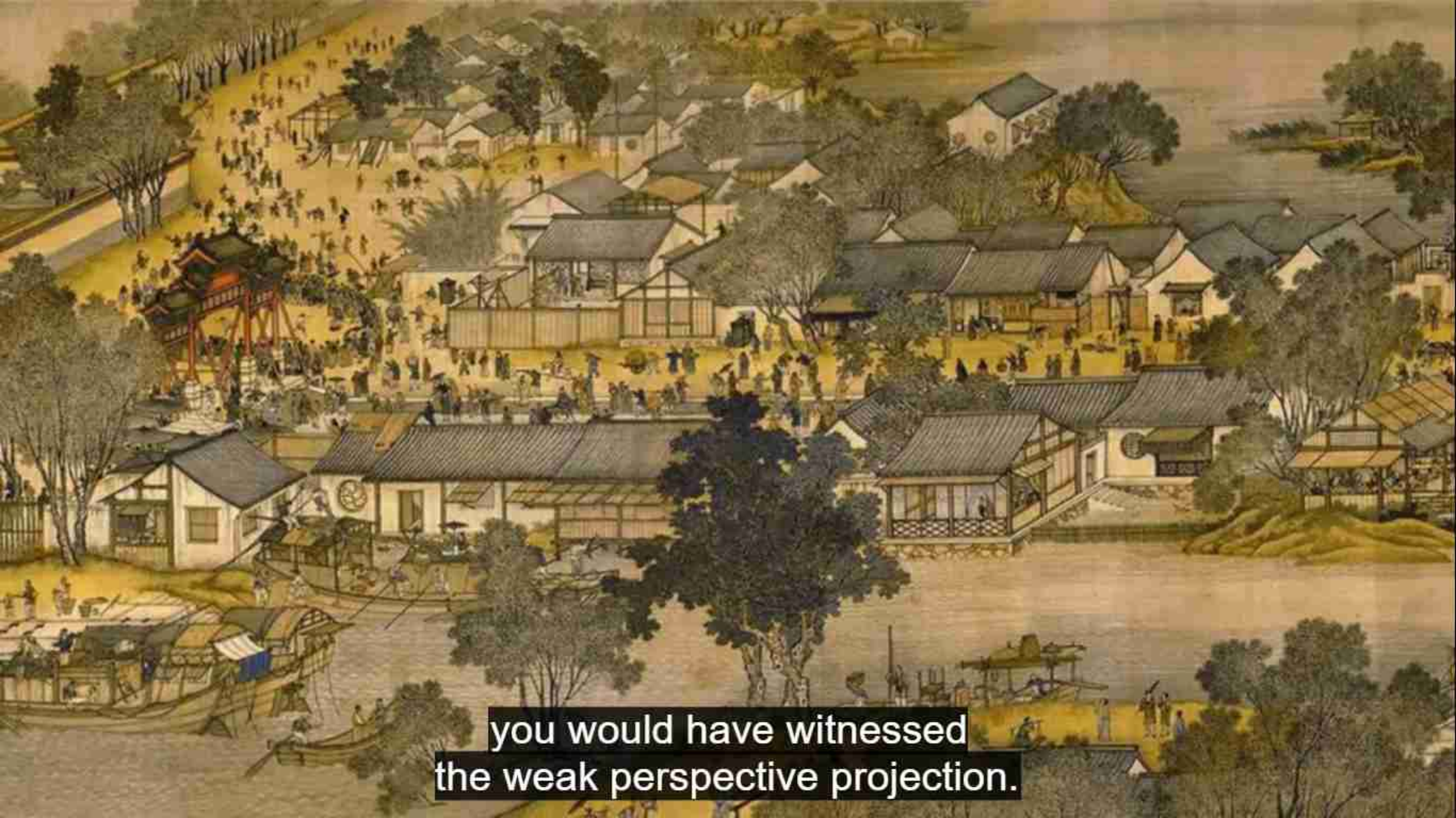

The next model we're paying attention
to is weak perspective projection.

# Weak Perspective Projection

$$x = f\frac{X}{Z_0}$$

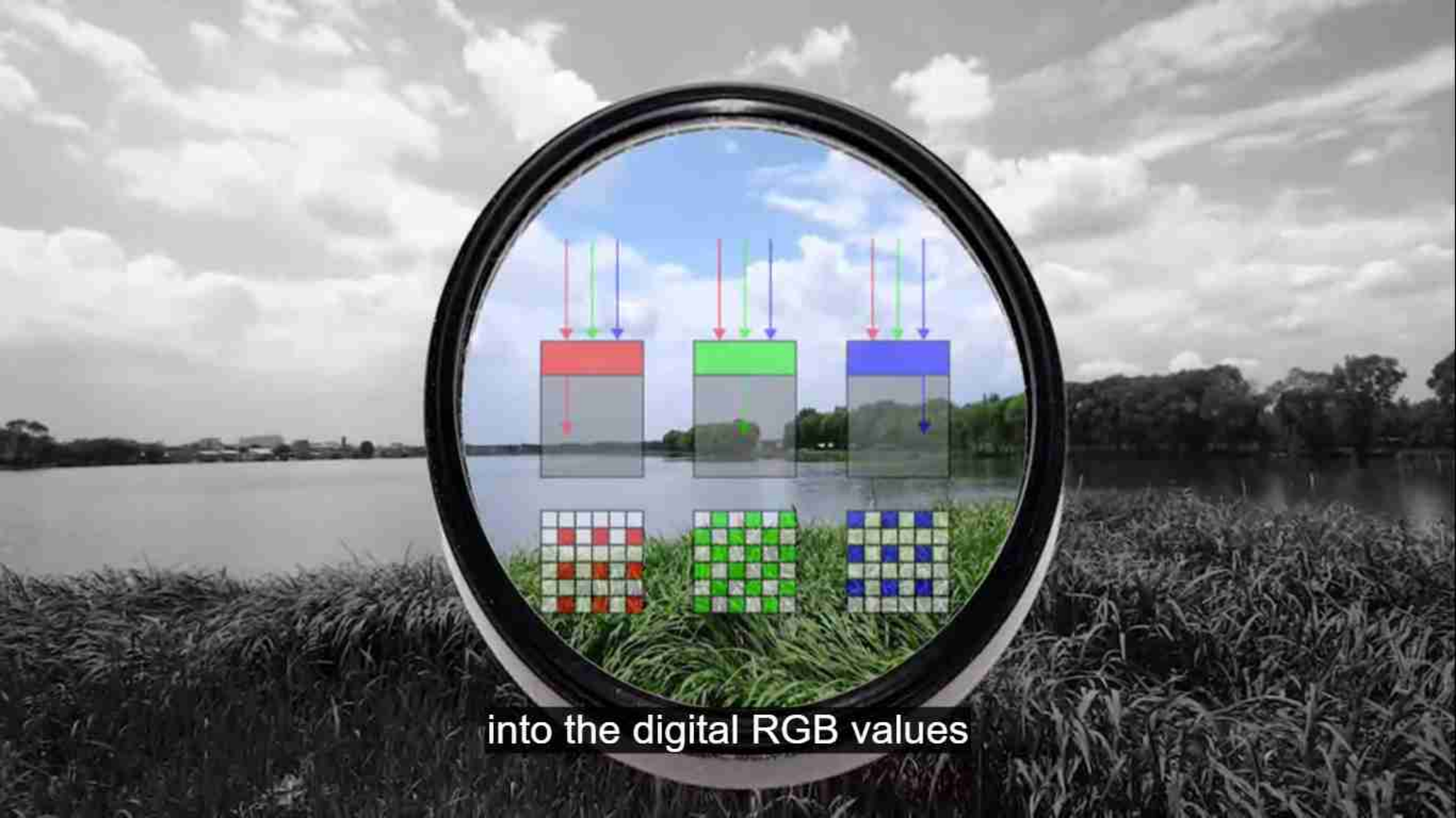
$$y = f\frac{Y}{Z_0}$$

If you ever tried to take a picture
by zooming your camera way too far,


you would have witnessed
the weak perspective projection.

Share
Digital Camer
0:01 / 5:53

finally reaches
the imaging sensor.


into the digital RGB values

# Simple Model - Important Effects

exposure

nonlinear mappings

sampling

aliasing

noise

aliasing and noise
associated with the cameras.

# Sophisticated Model - Important Effects

focus

exposure

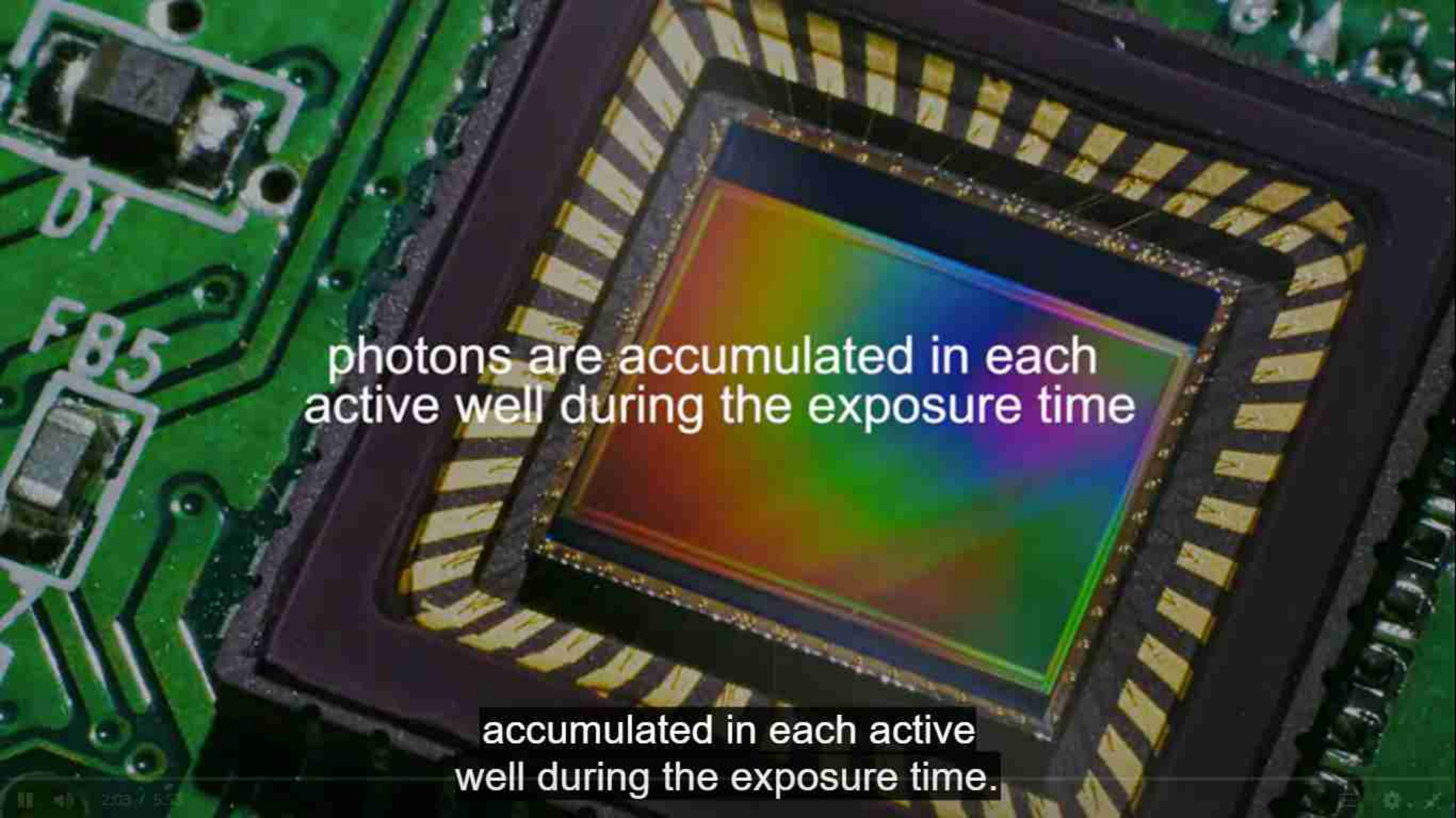
vignetting

aberration

focus exposure,
vignetting or aberration.

# Shutter Speed

light falling on an imaging sensor,
for the duration of the exposure

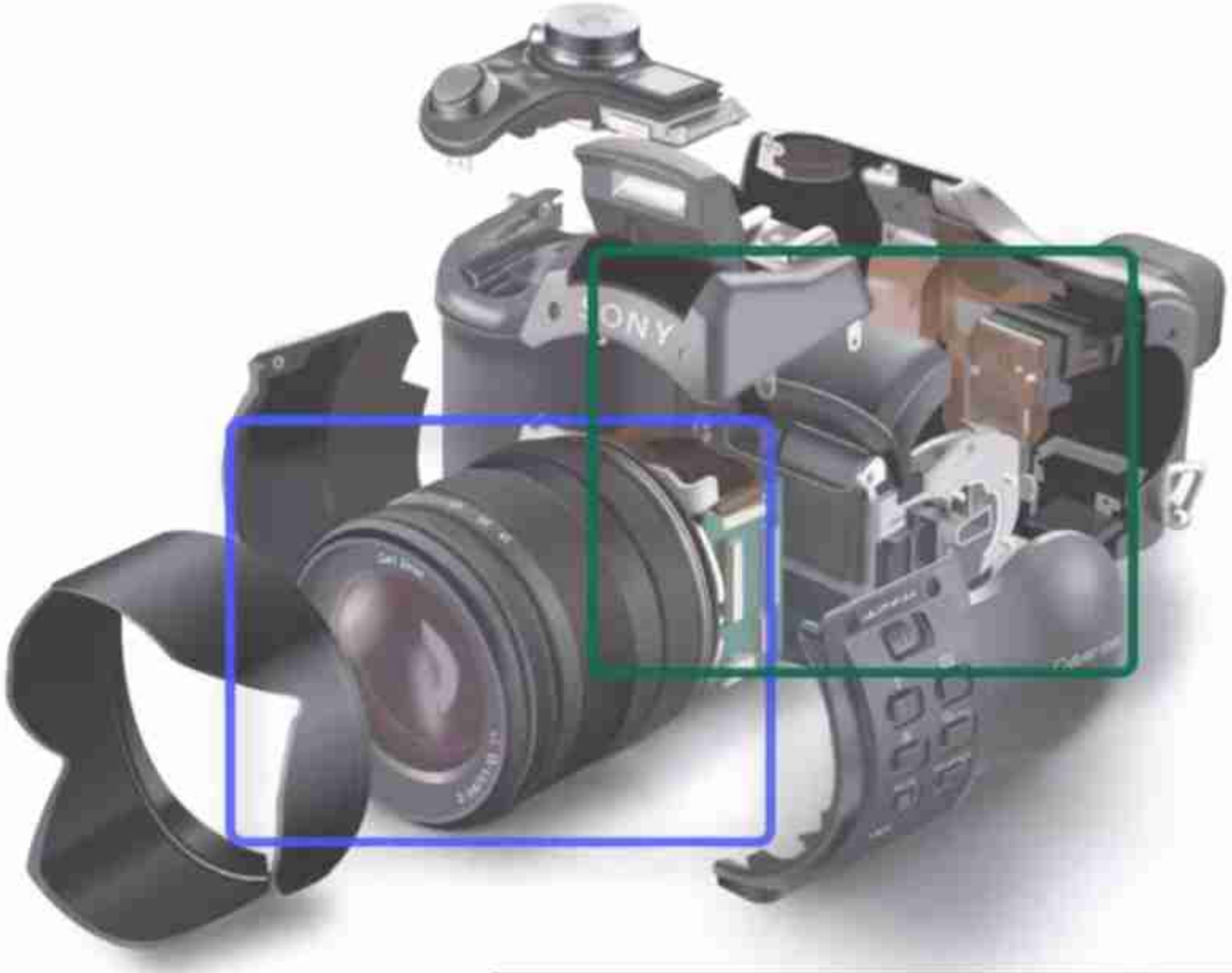

photons are accumulated in each
active well during the exposure time
accumulated in each active
well during the exposure time.

# Factors Affecting Performance

shutter speed
sampling pitch
chip size
analog gain
sensor noise
analog-to-digital converter

sensor noise, and the quality

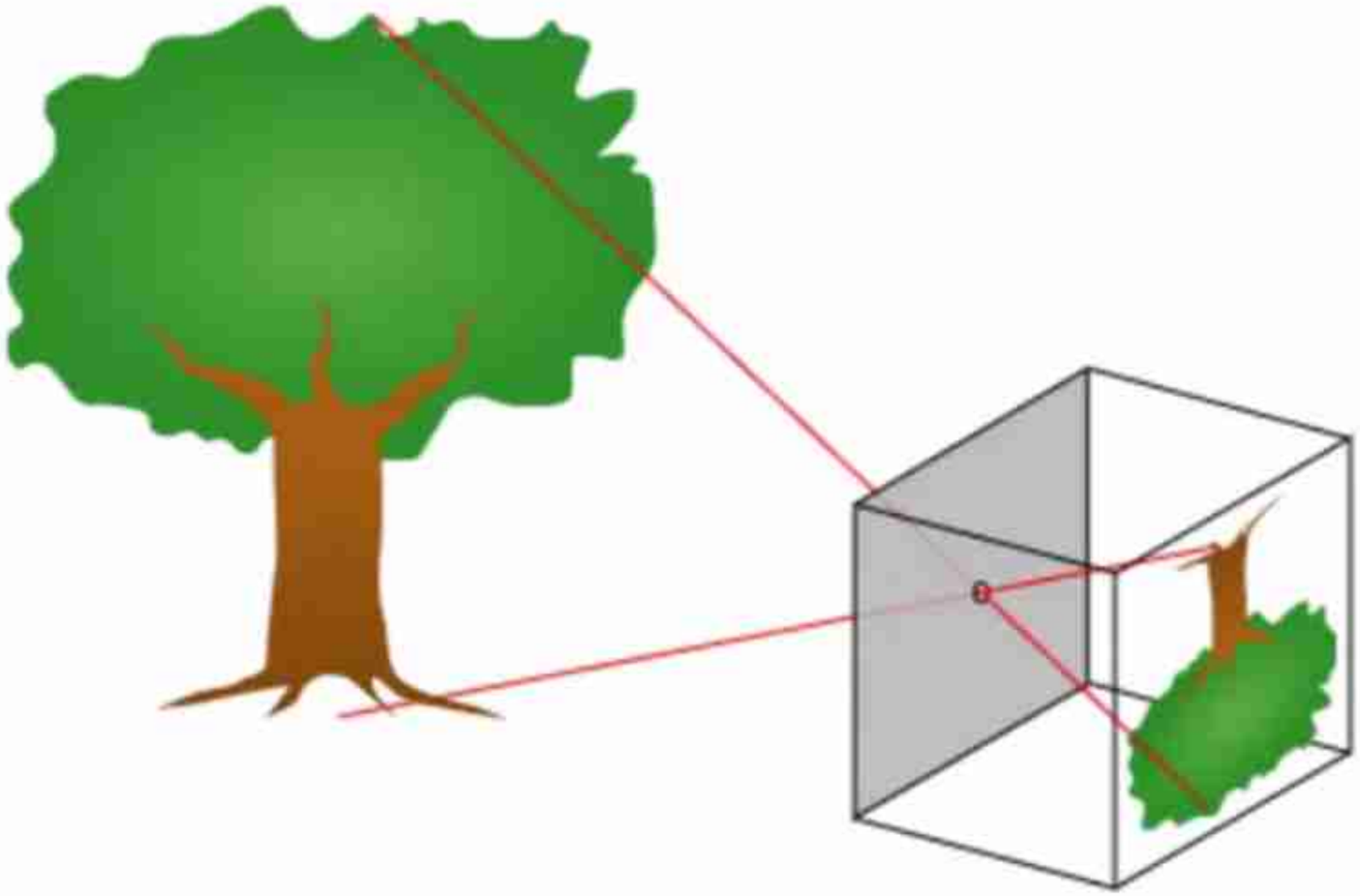
captured using a digital camera



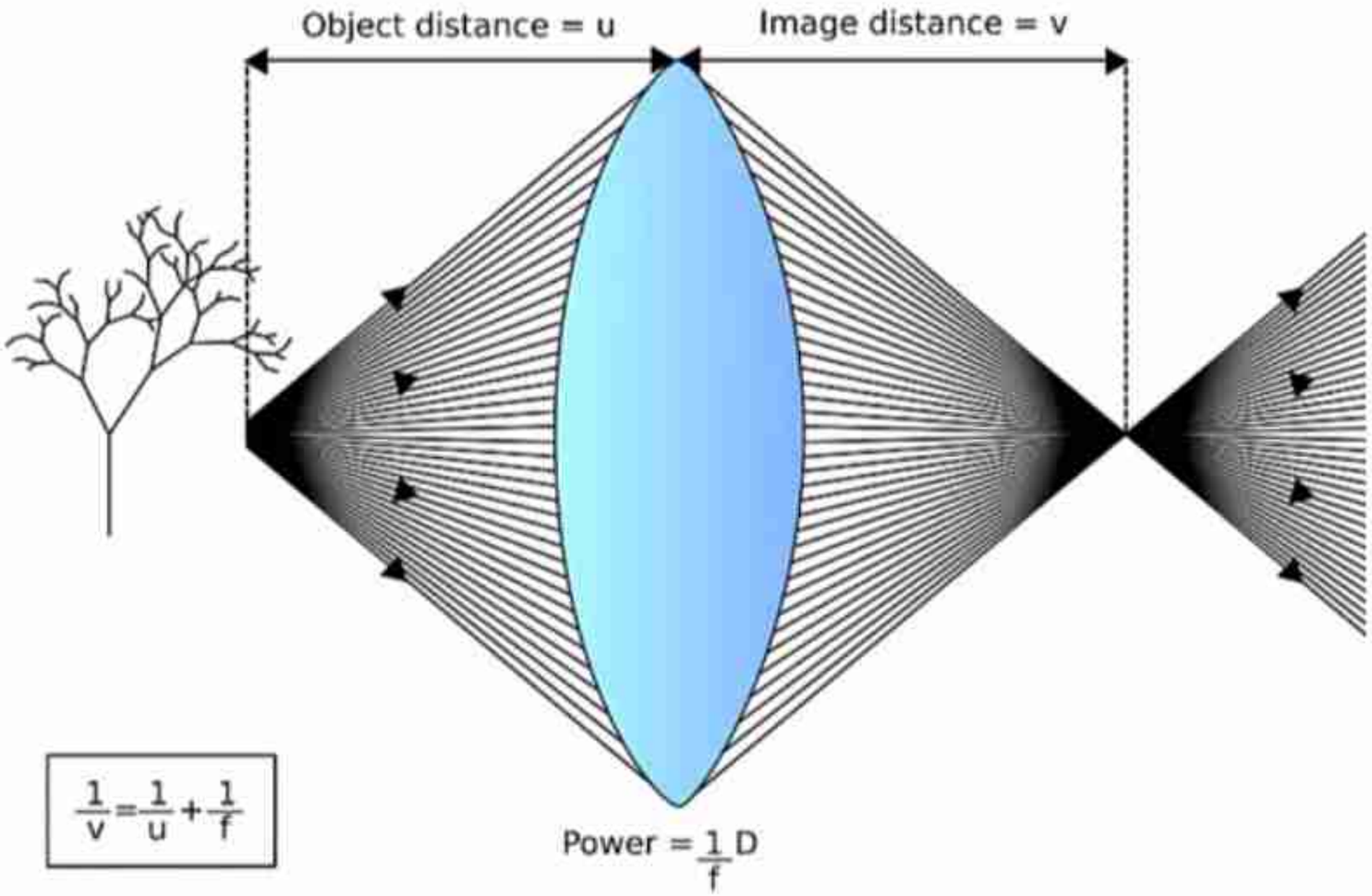

Optics

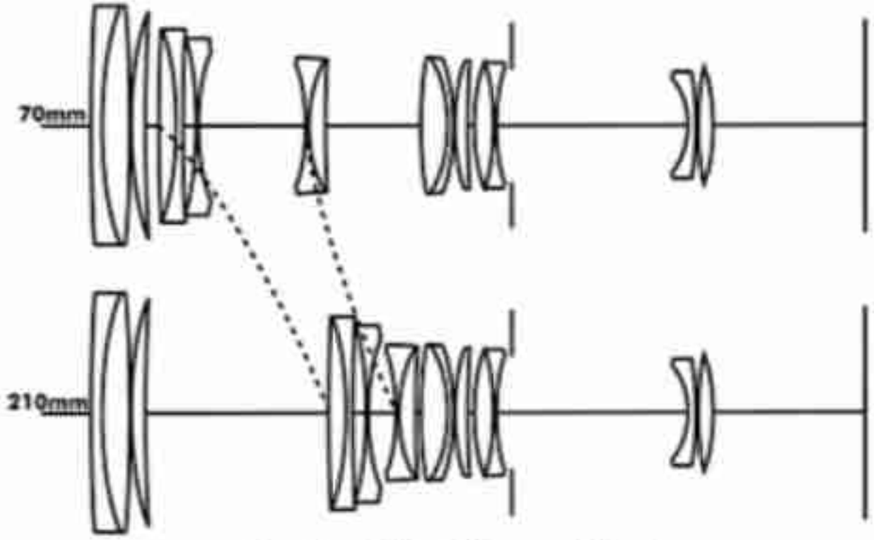
Sensor and Image Processing

as well as the image processors.

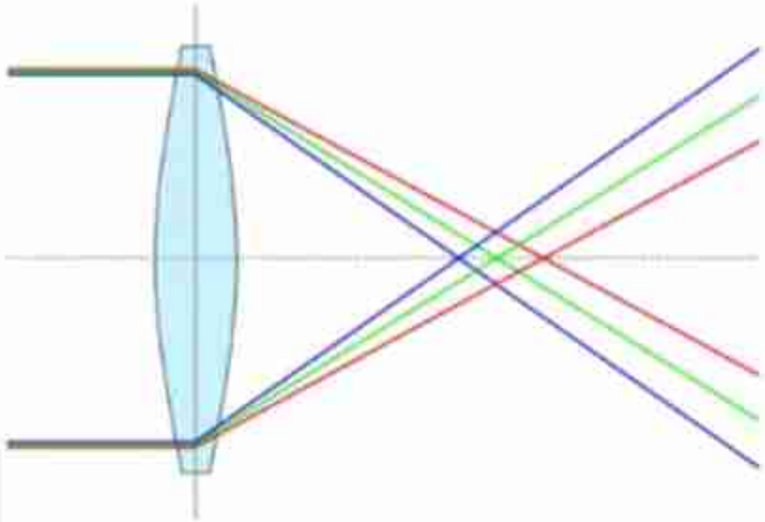


Moving away from
the simple pinhole camera model,



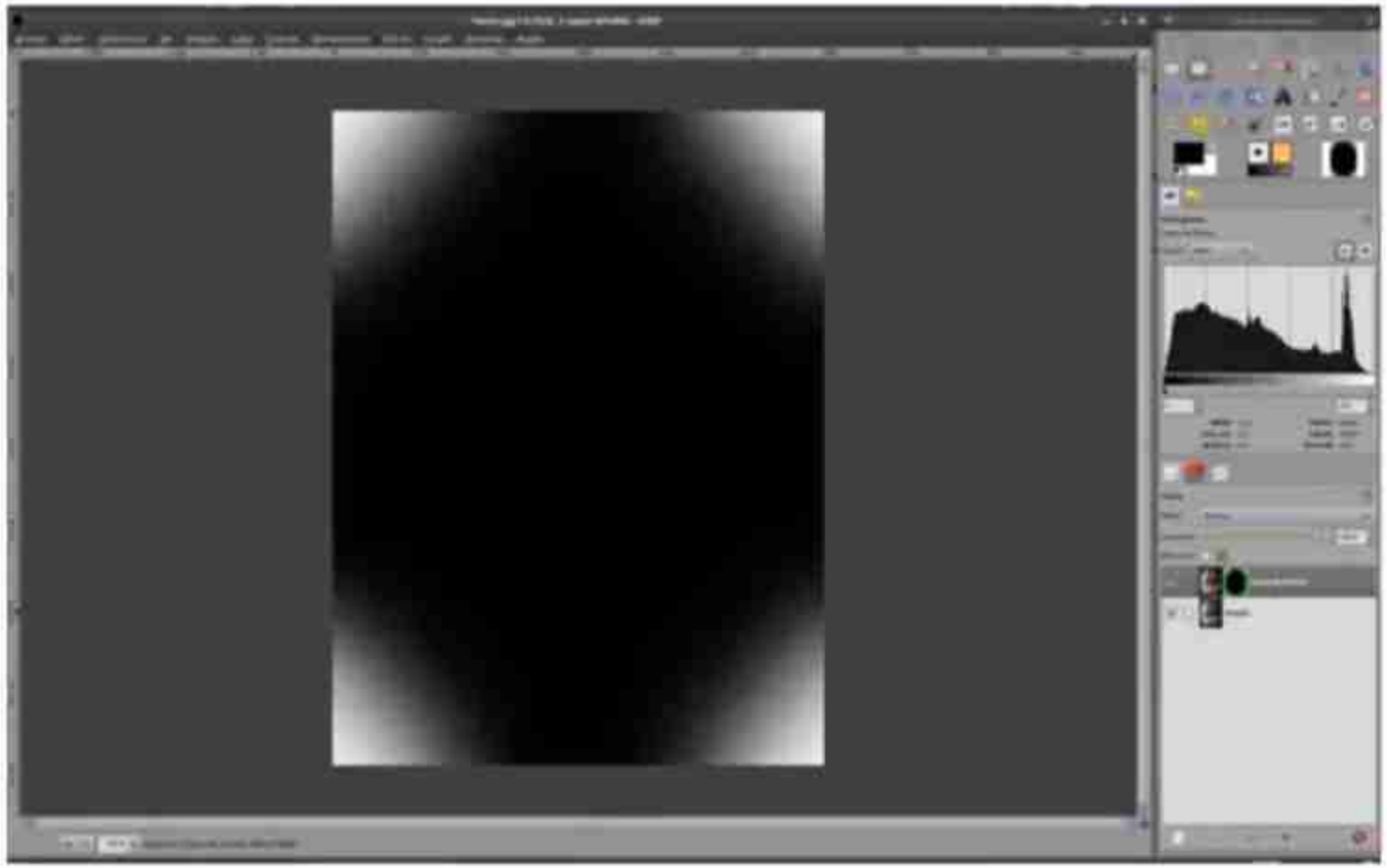




Ponder & Best (Opcon/Kino)
Vivitar Series 1 70-210mm f/3.5, 1974

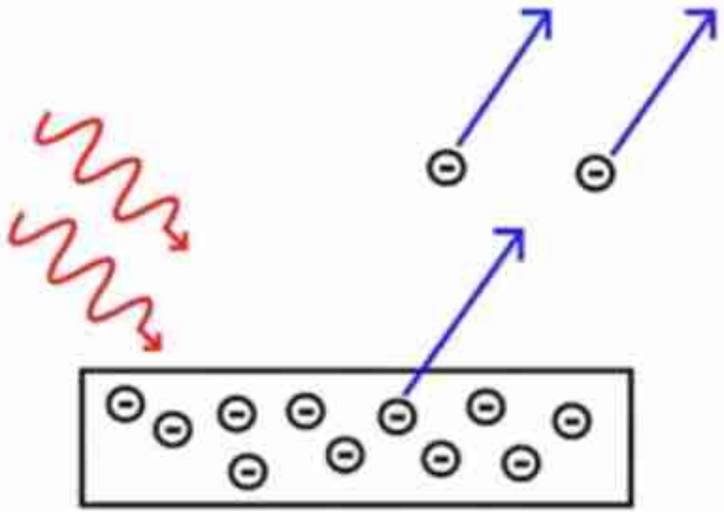


take them into account in
understanding digital images.

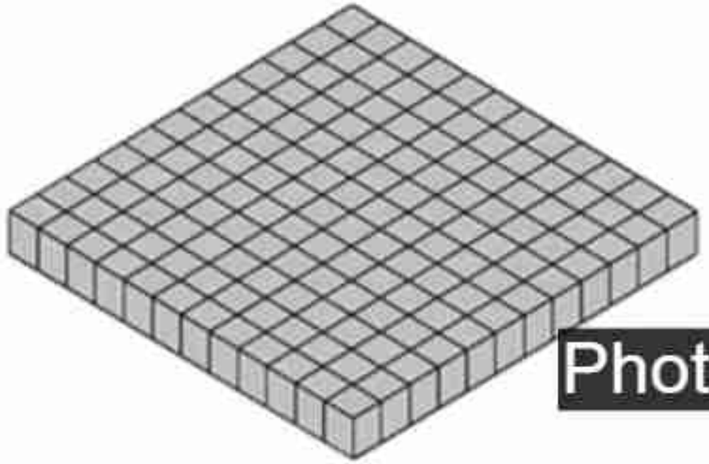
# Vignetting



the lens optics which needs
to be compensated for.



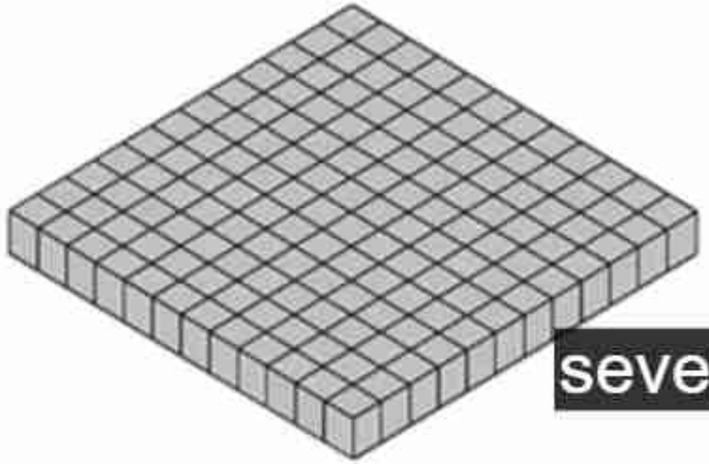
Photoelectric Effect



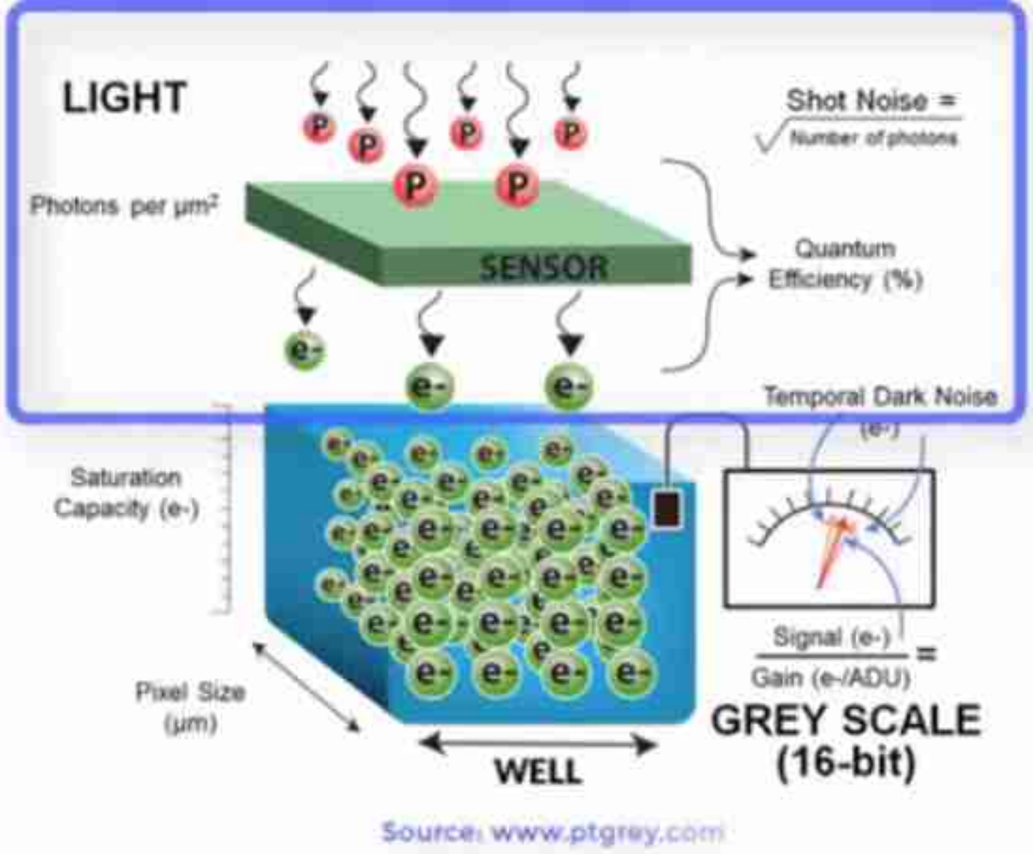
Photoelectric effect is

Intensity at an image pixel

Exposure time/Shutter Speed

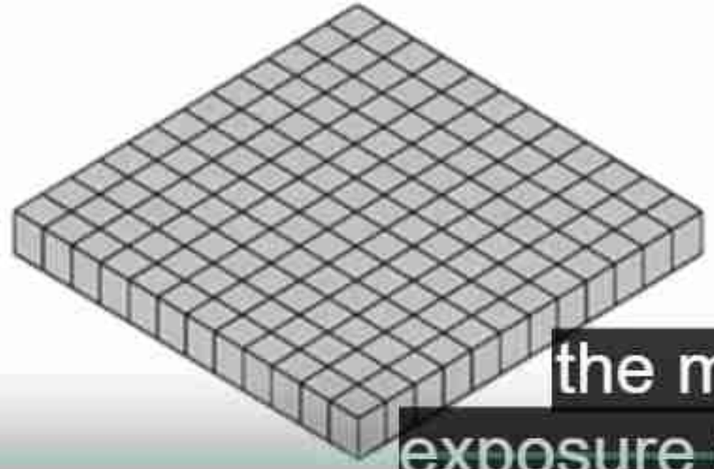


several factors one of



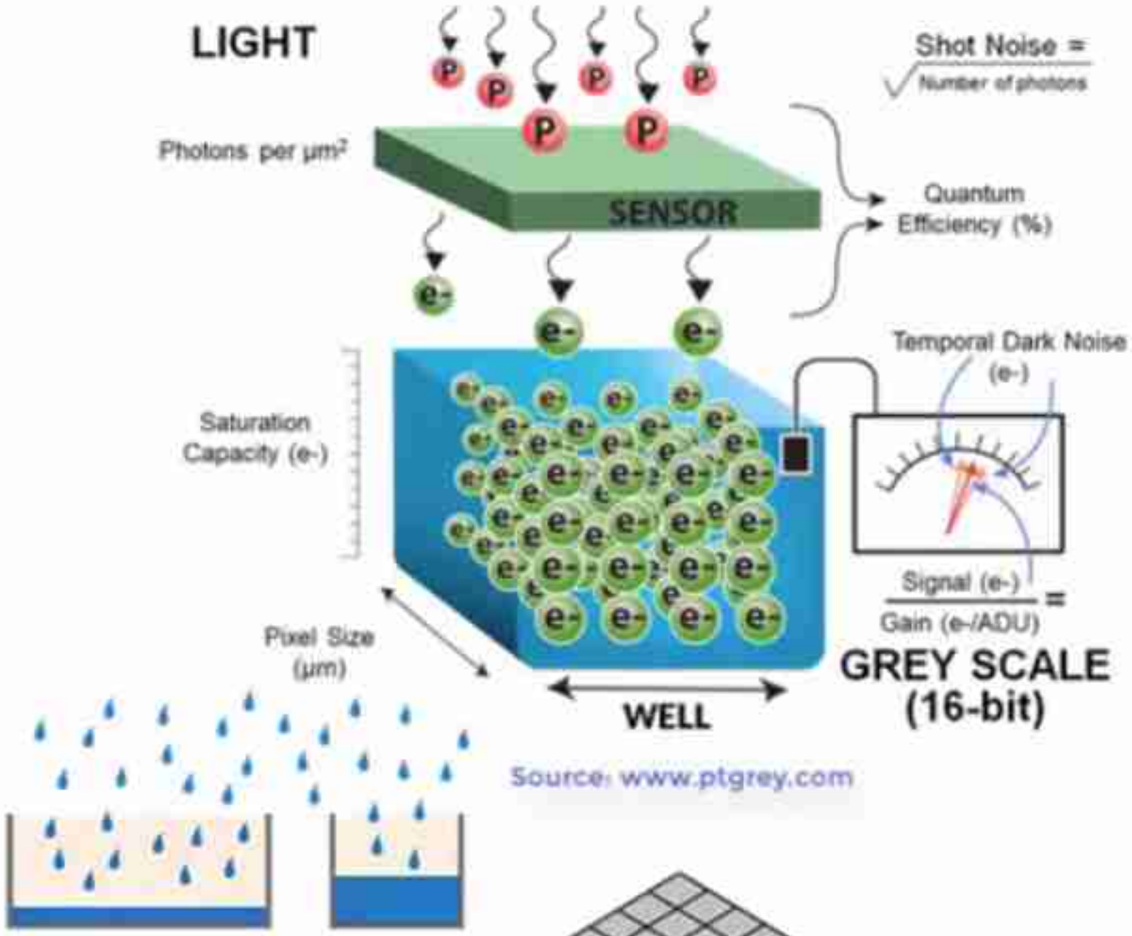
Intensity at an image pixel

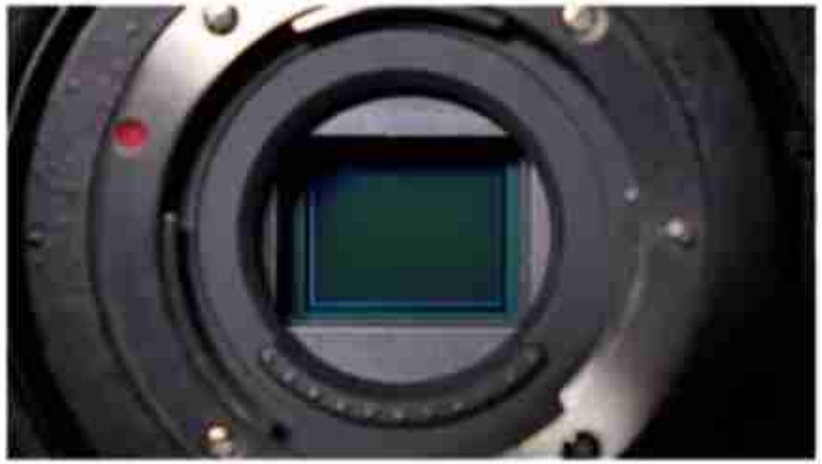
Exposure time/Shutter Speed



the main factors being
exposure time or shutter speed,

3:50 / 5:53



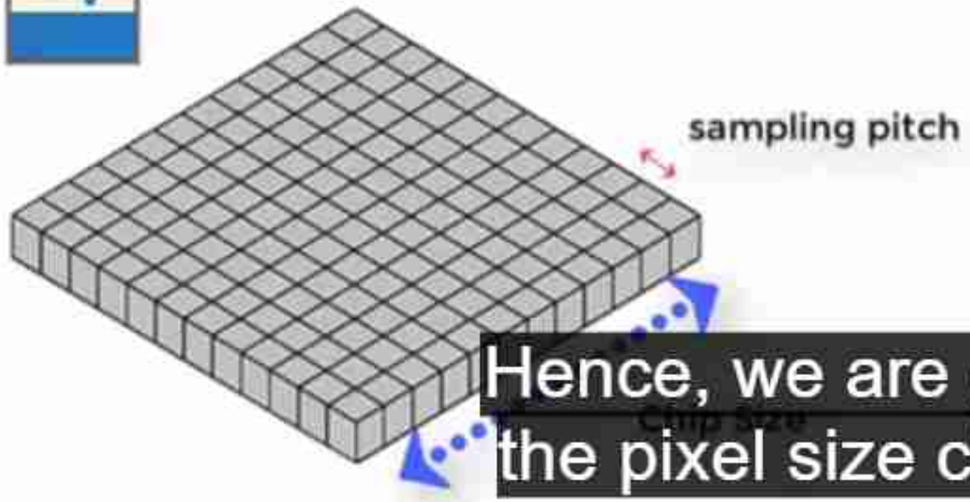



## Intensity at an image pixel

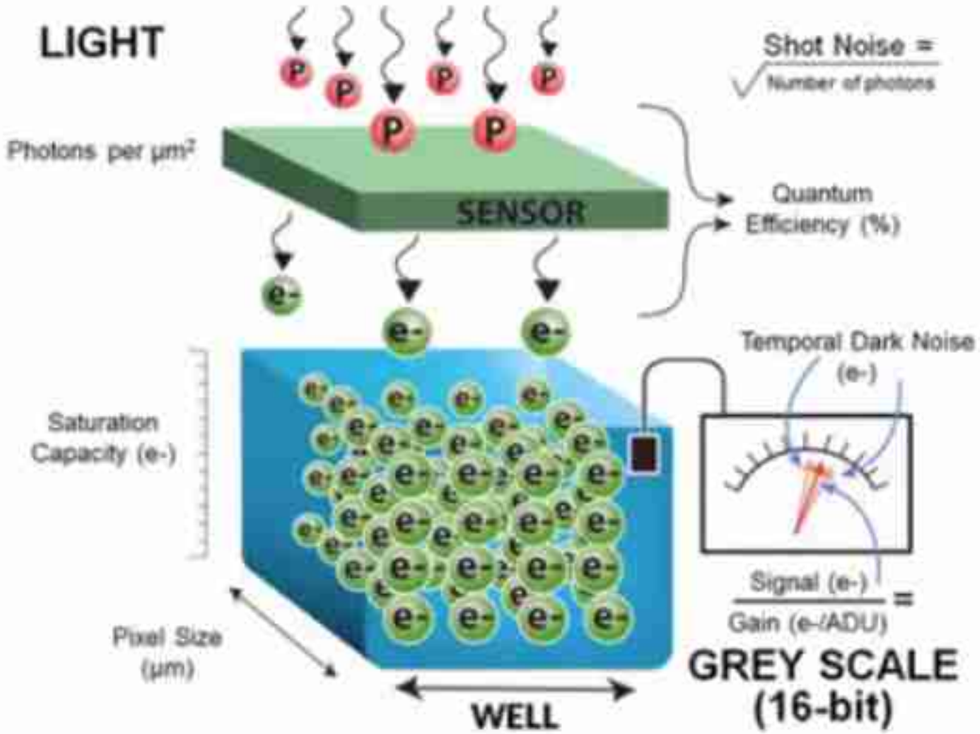
Exposure time/Shutter Speed

Photodiode Quantum Efficiency

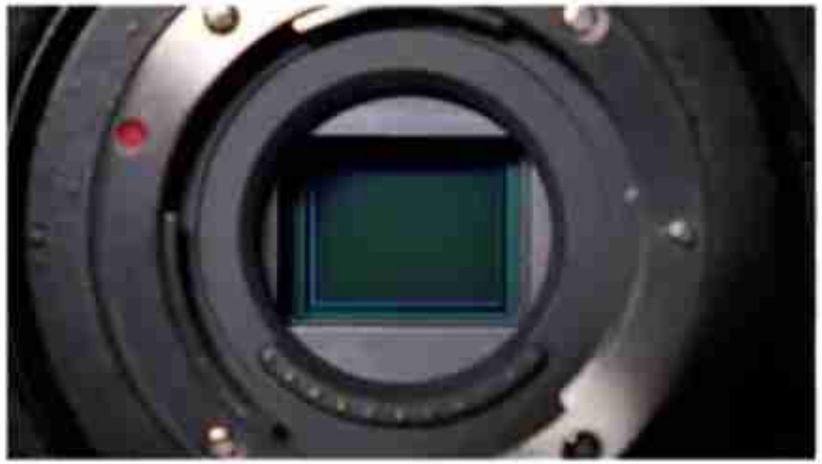
pixel size/chip size/sampling pitch




Hence, we are going to list the pixel size chip size as




Source: www.ptgrey.com



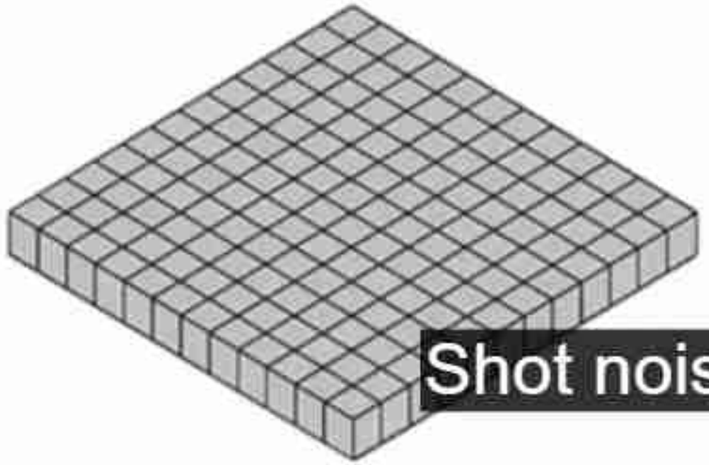
## Intensity at an image pixel

Exposure time/Shutter Speed

Photodiode Quantum Efficiency

pixel size/chip size/sampling pitch

(Photon) Shot Noise/ Dark Noise

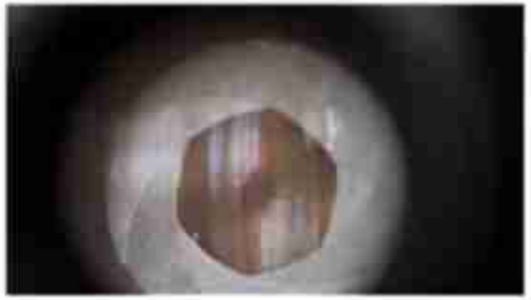


Shot noise and dark noise are

# Digital Camera Pipeline

Optics

Aperture



Shutter

Sensor

Finally, light
reaches the sensor.

Share

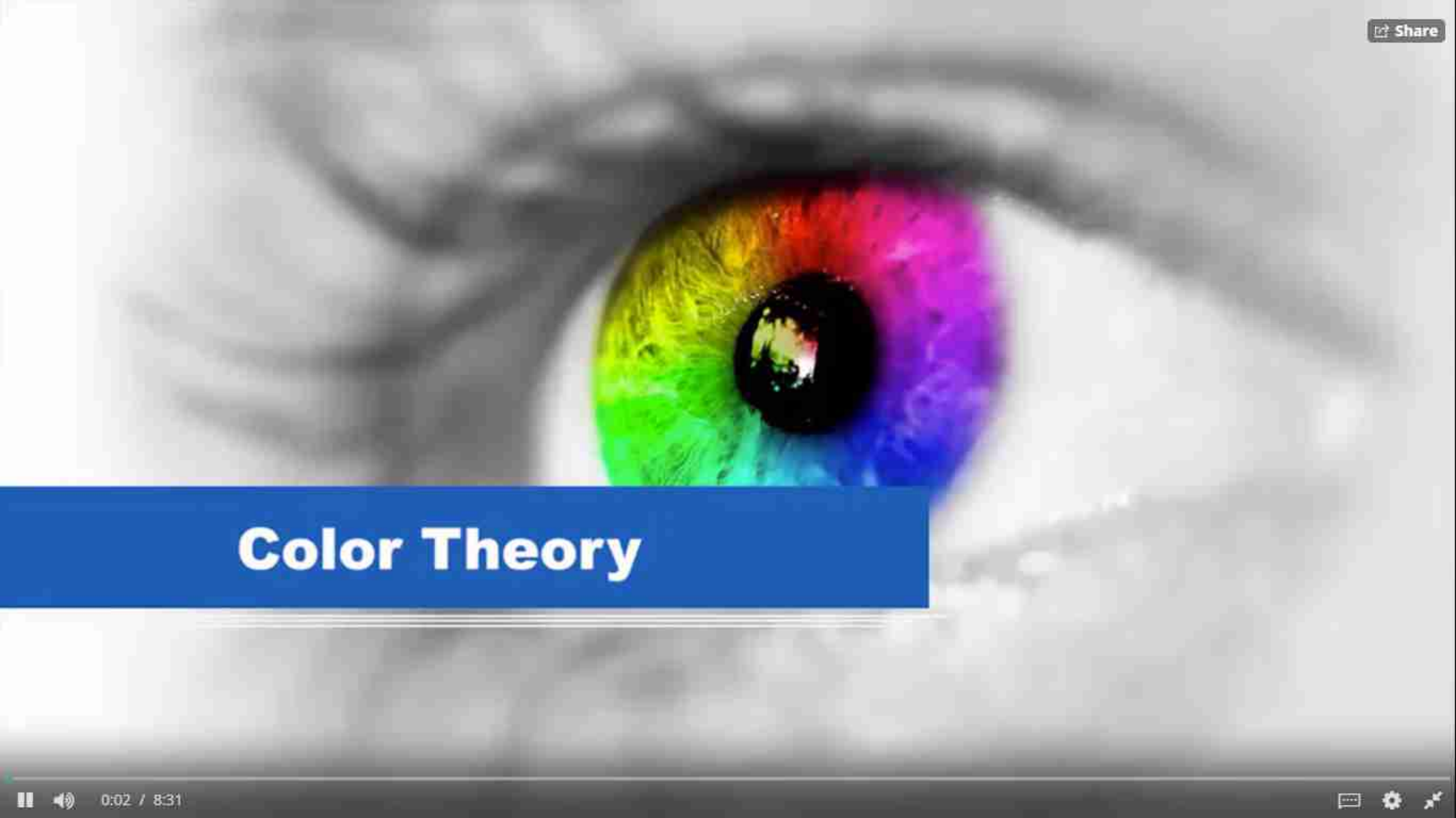
# Digital Camera Pipeline

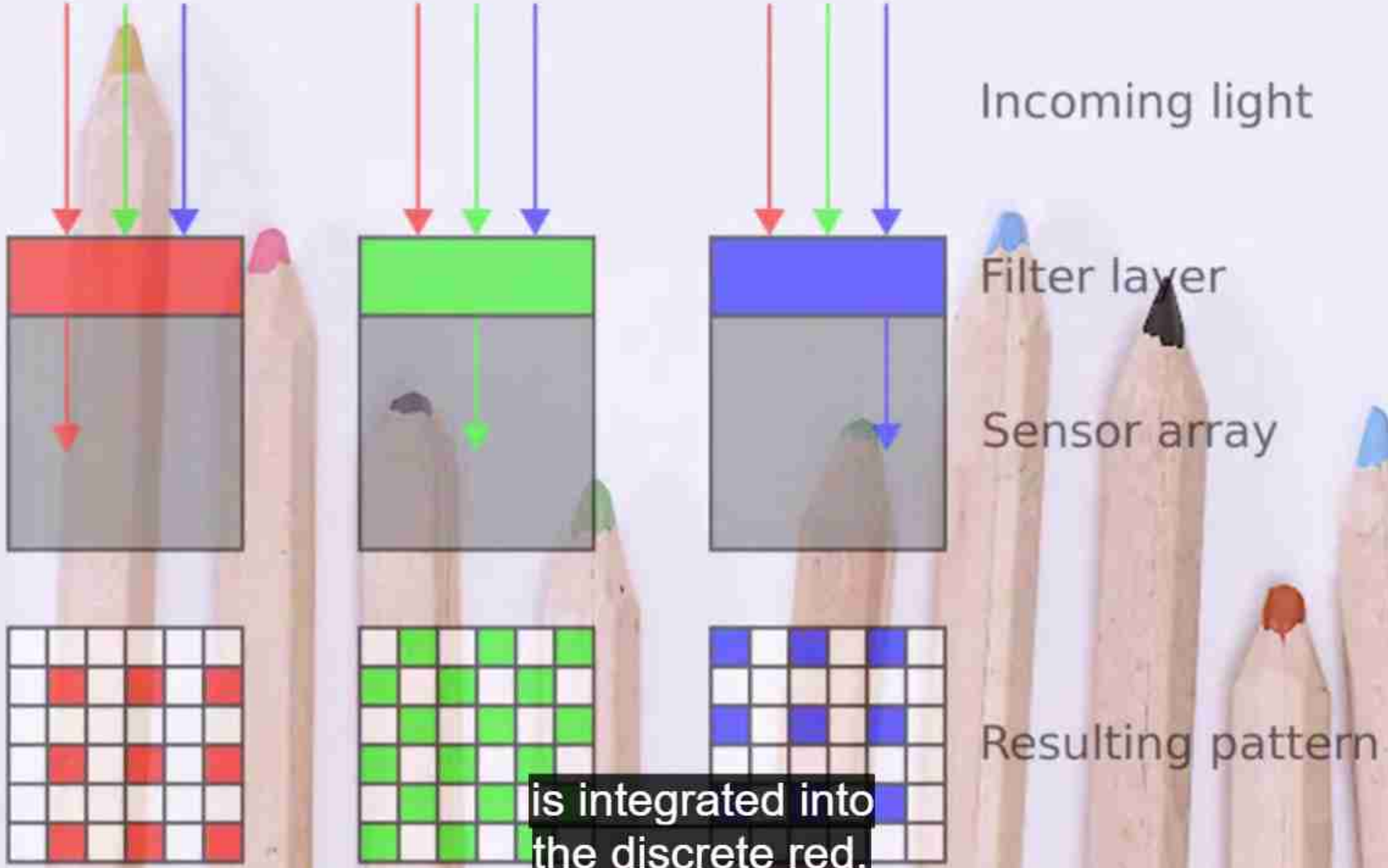
Optics → Aperture → Shutter → Sensor

Gain → Demosaic → Denoise → White Balance → Gamma Correction → Compress

has been made in digital imaging.

5:40 / 5:53


Share
Color Theory
0:02 / 8:31




is integrated into
the discrete red,

Blue

Yello

For example, adding blue
with yellow makes green.

Share

Press Esc to exit full screen

The additive primary colors
red, green, and blue,

0:52 / 8:31

Share

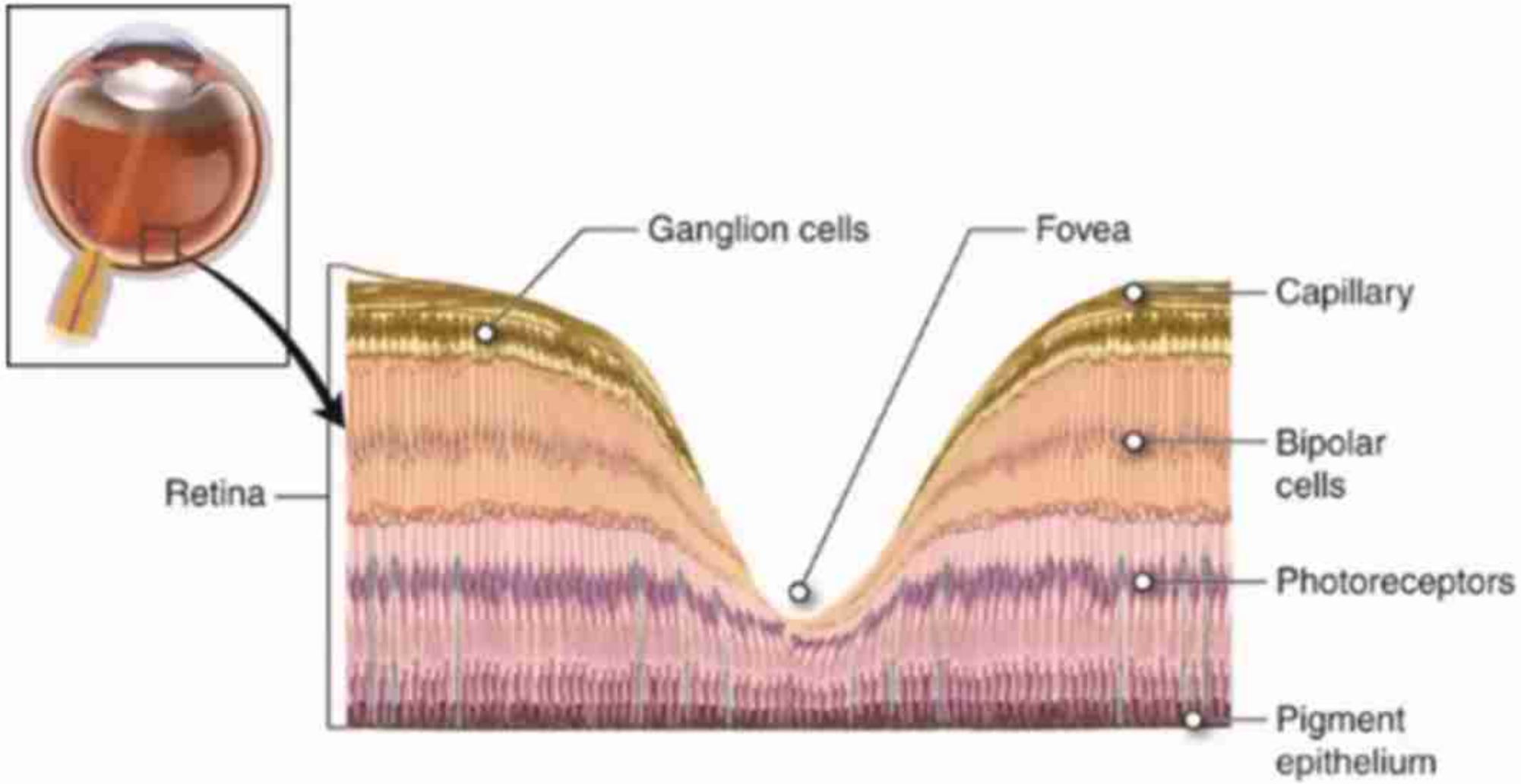
Press Esc to exit full screen

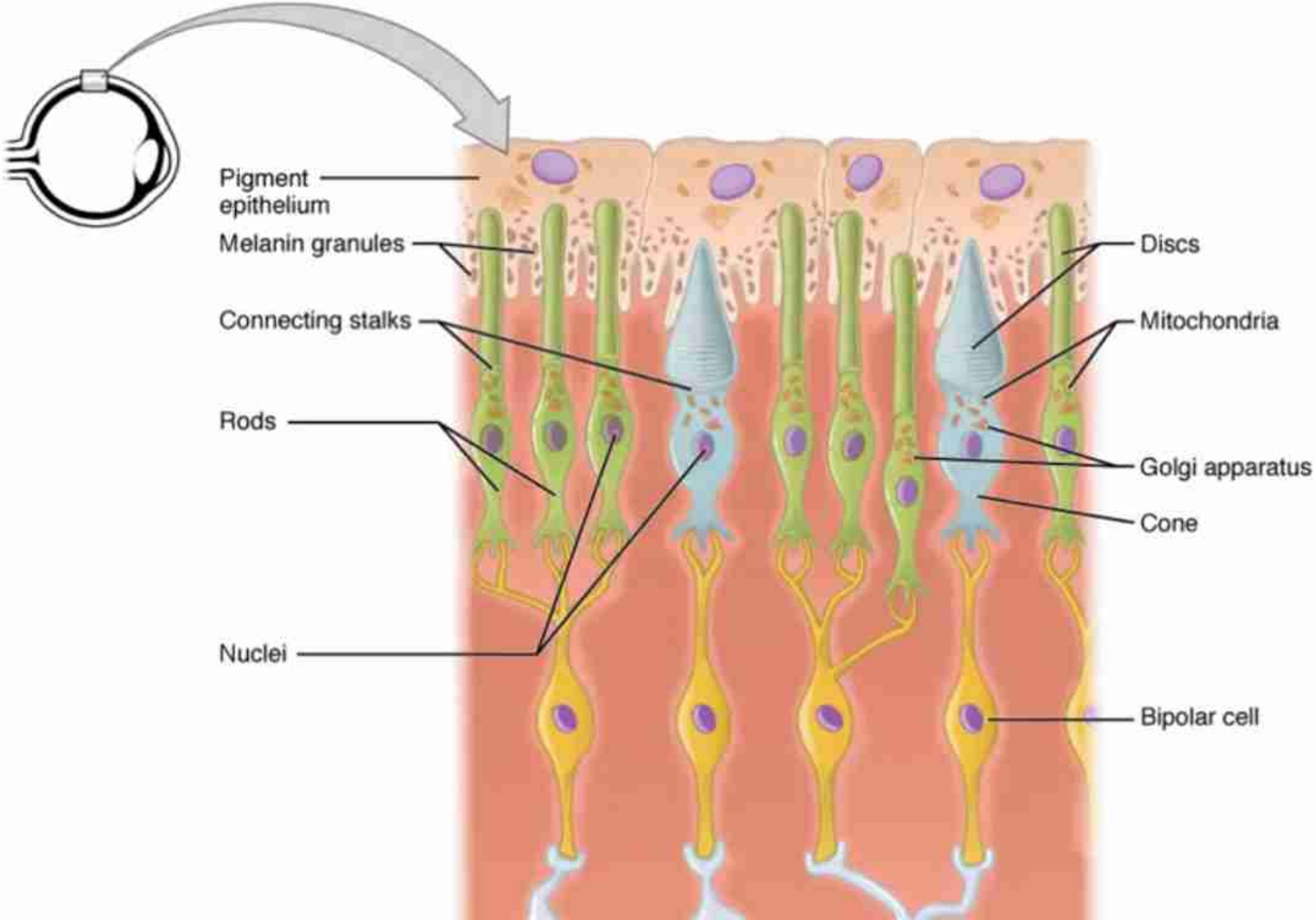
# Additive Primary Colors

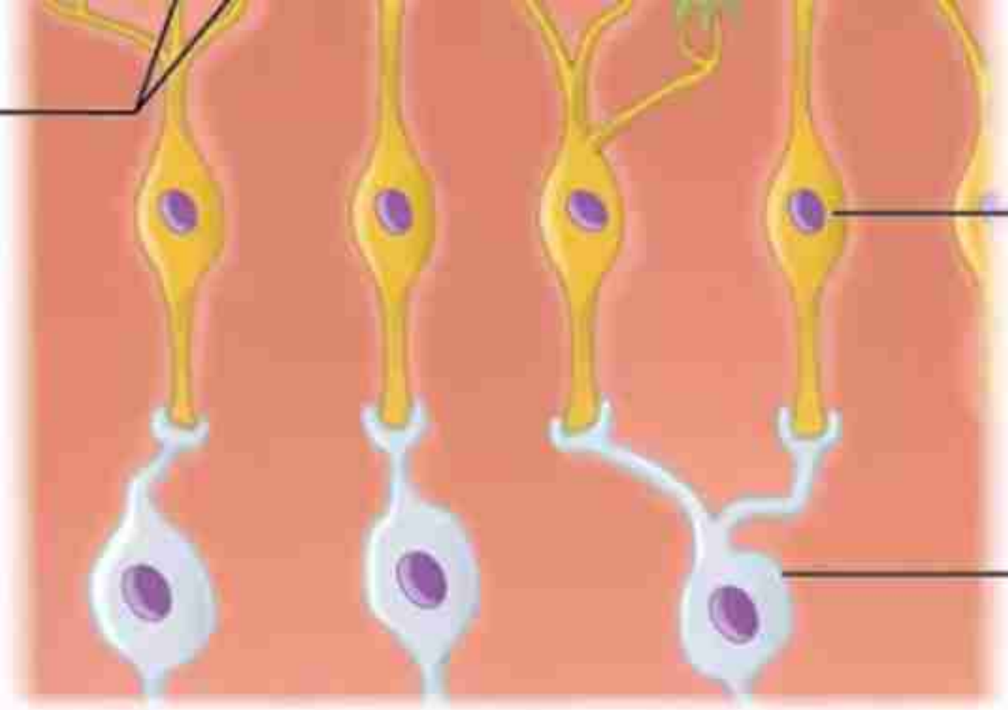
Red Green Blue

The additive primary colors
red, green, and blue,

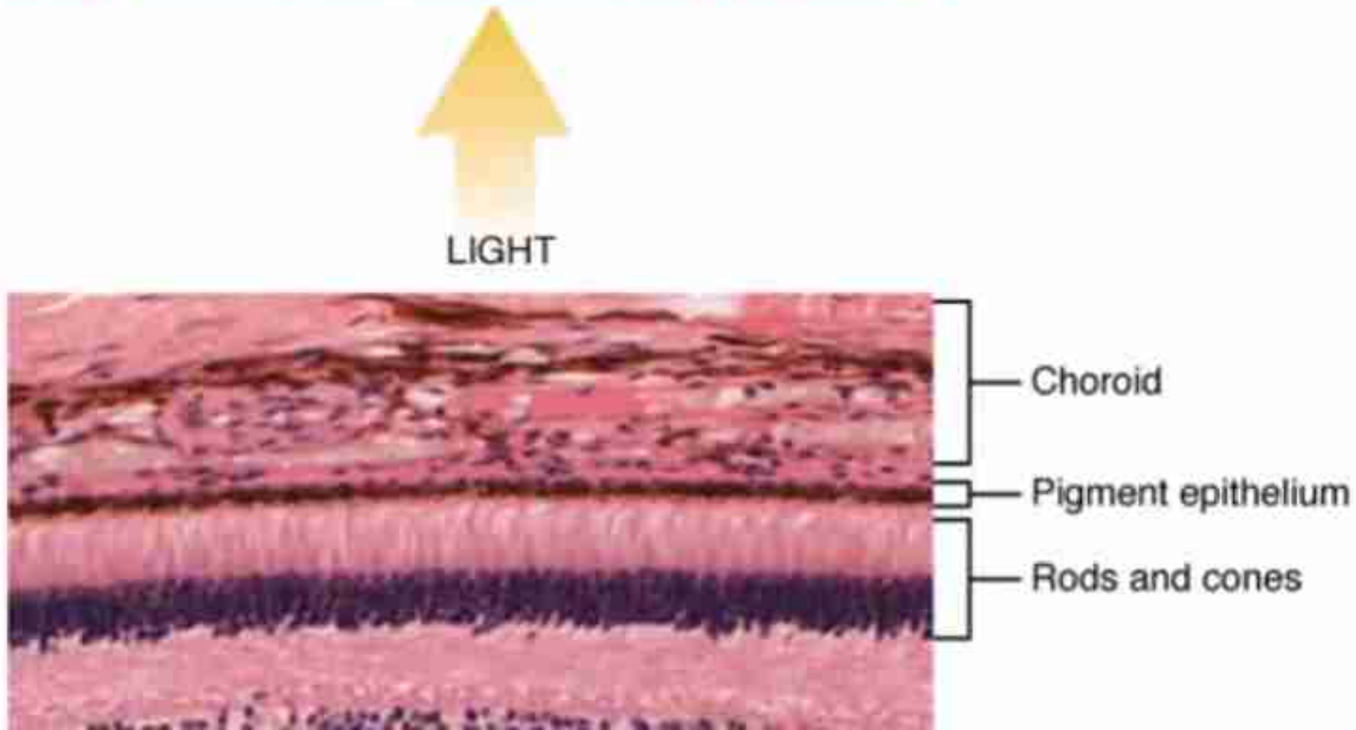
0:54 / 8:31

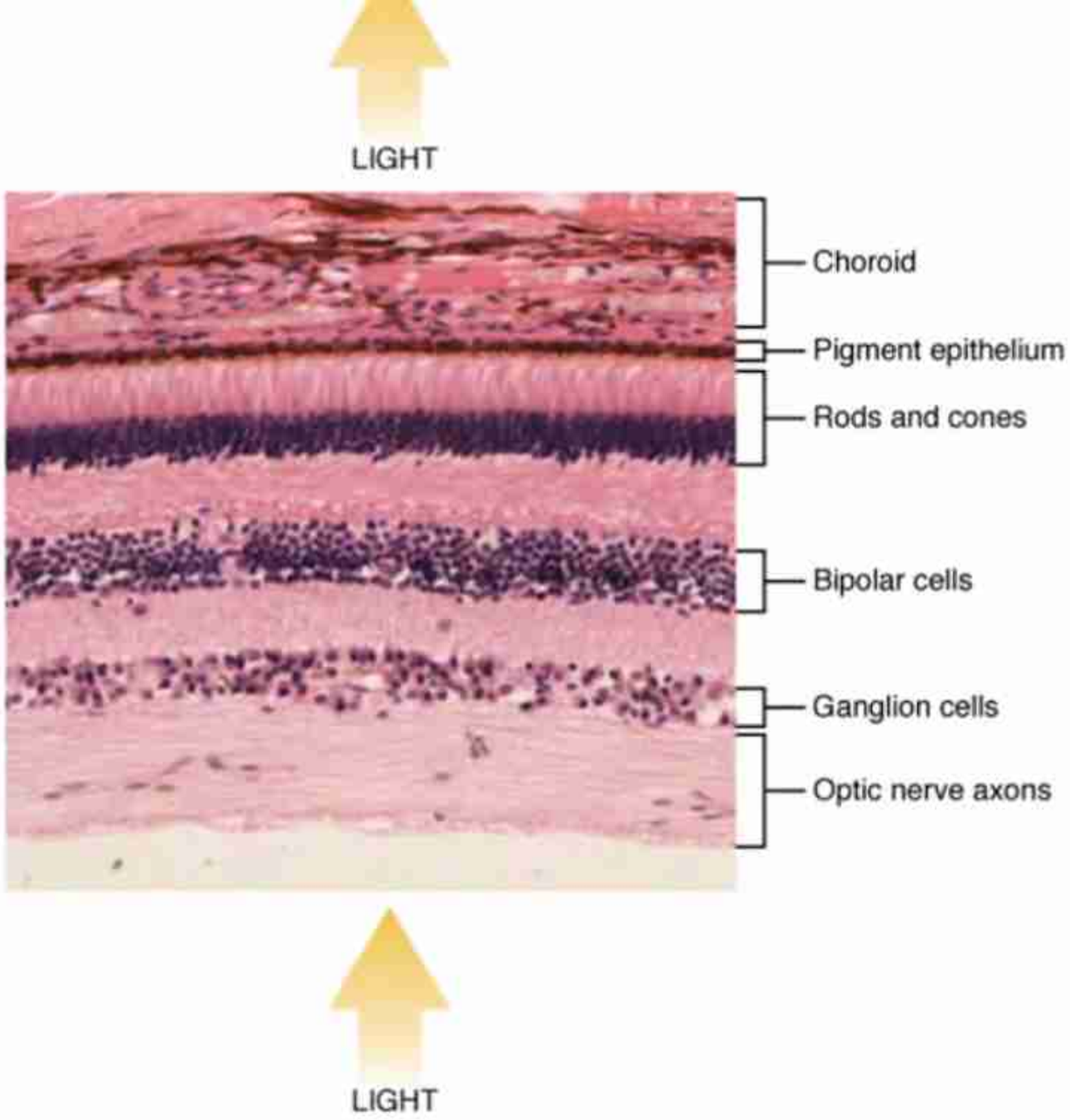

Ganglion cells
Fovea
Capillary
Retina
Bipolar cells
Photoreceptors
Pigment epithelium


Pigment epithelium
Melanin granules
Connecting stalks
Rods
Nuclei
Discs
Mitochondria
Golgi apparatus
Cone
Bipolar cell

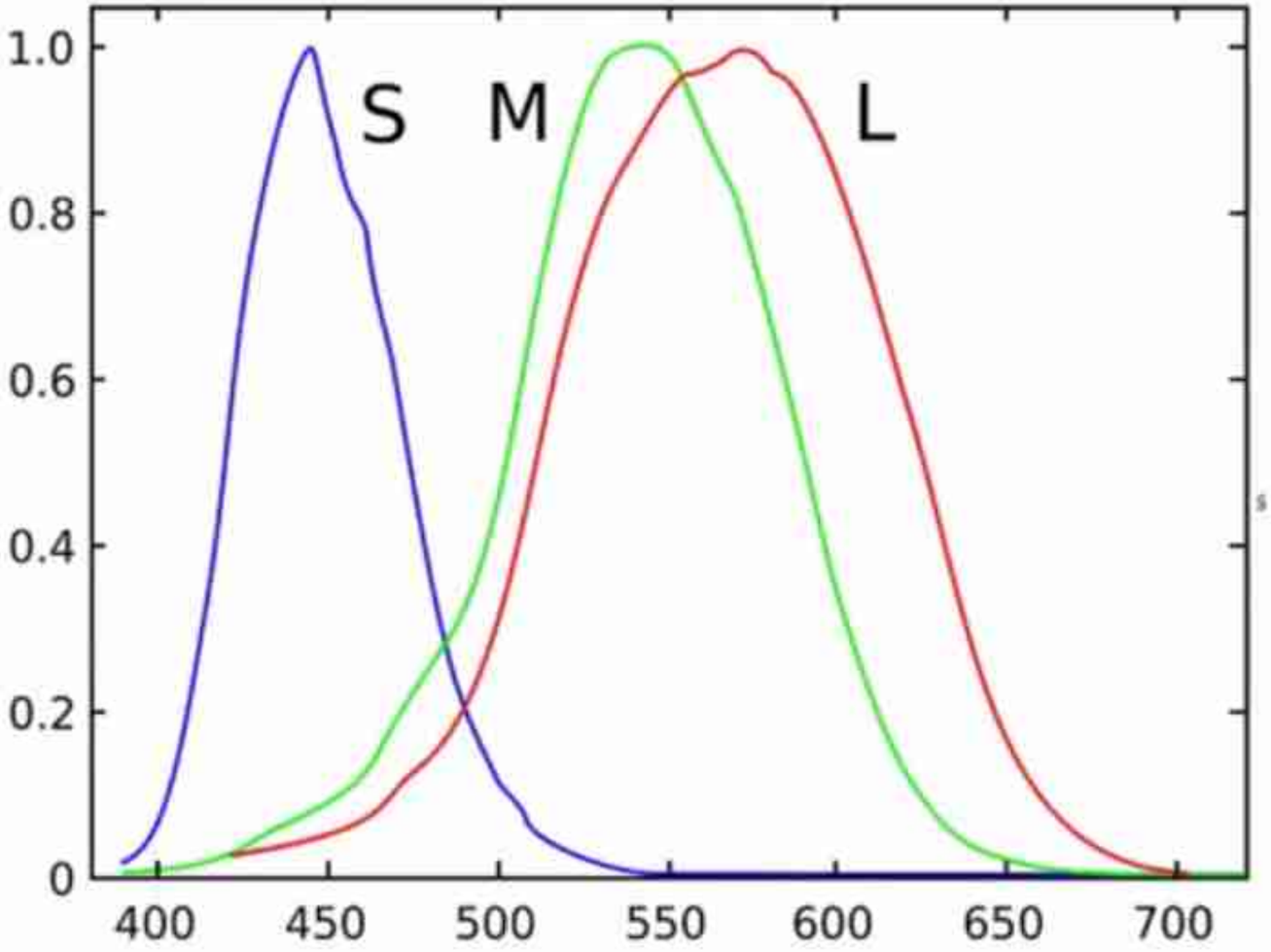

Nuclei
Bipolar cell
Ganglion cell

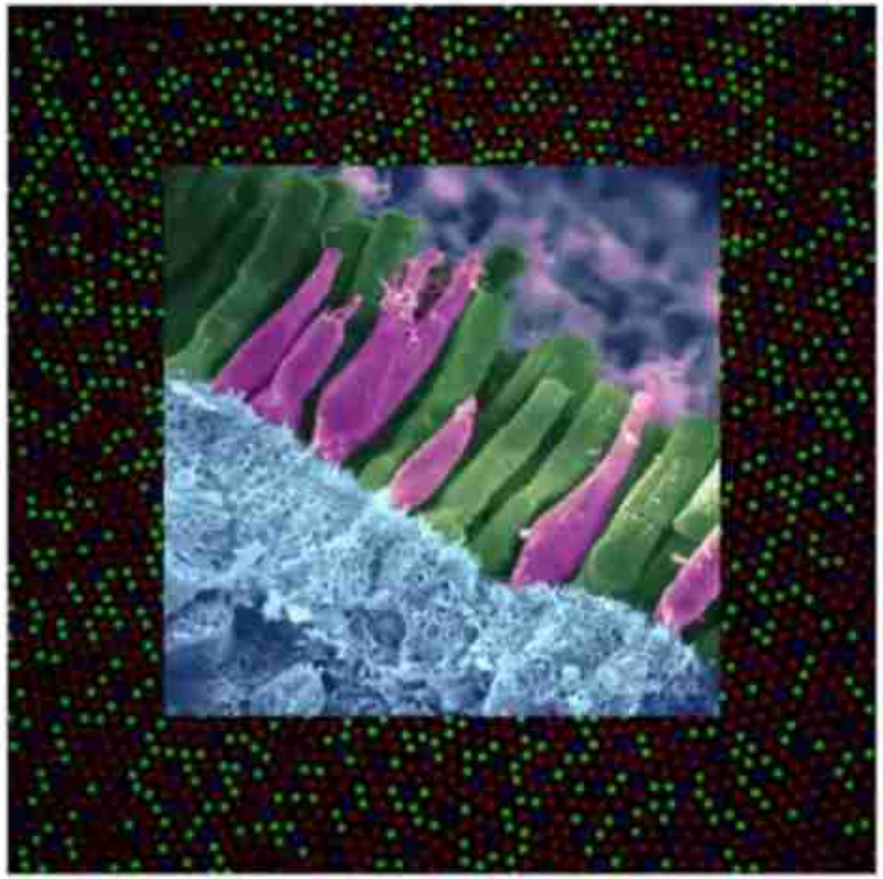
(a)

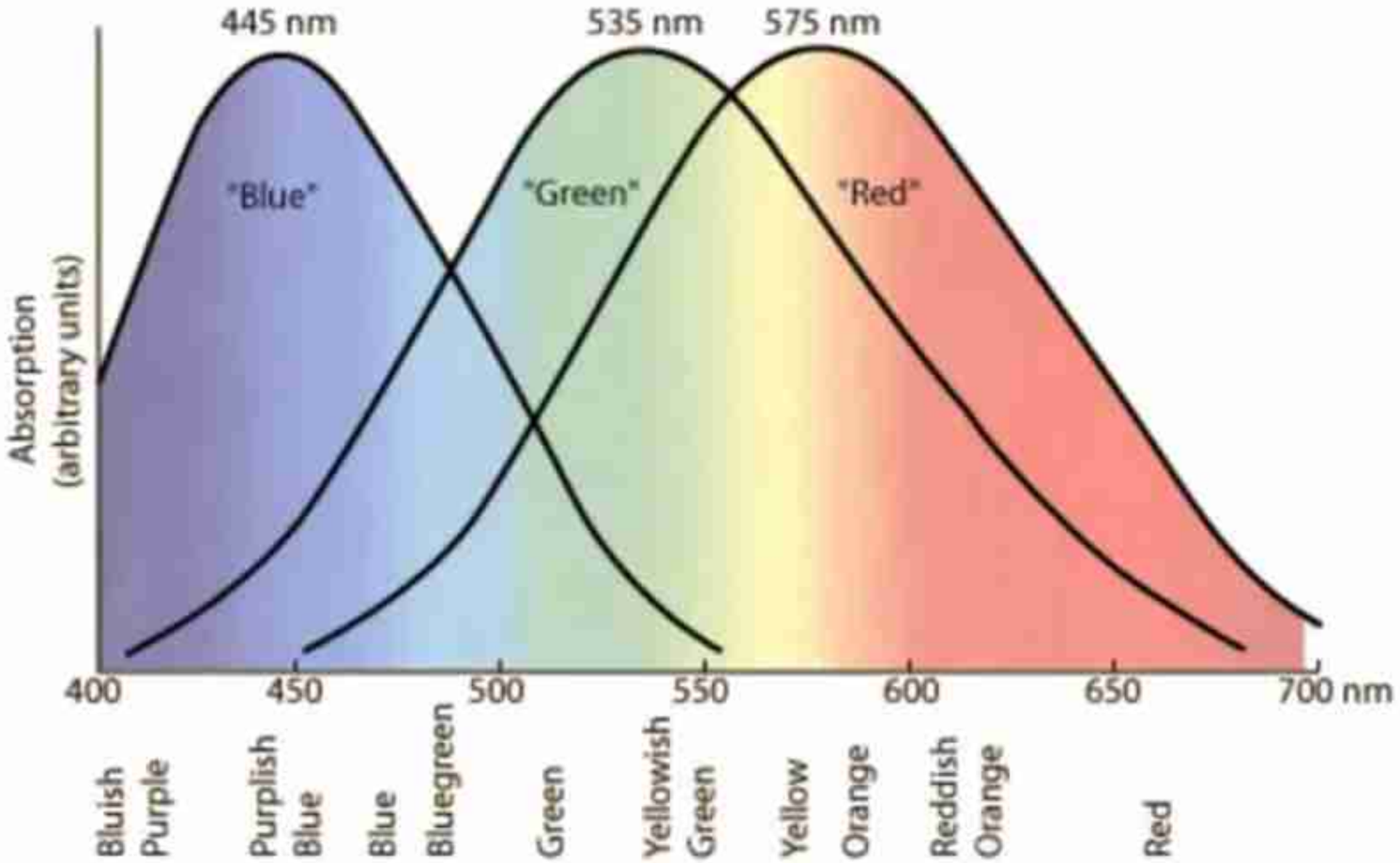

LIGHT
Choroid
Pigment epithelium
Rods and cones

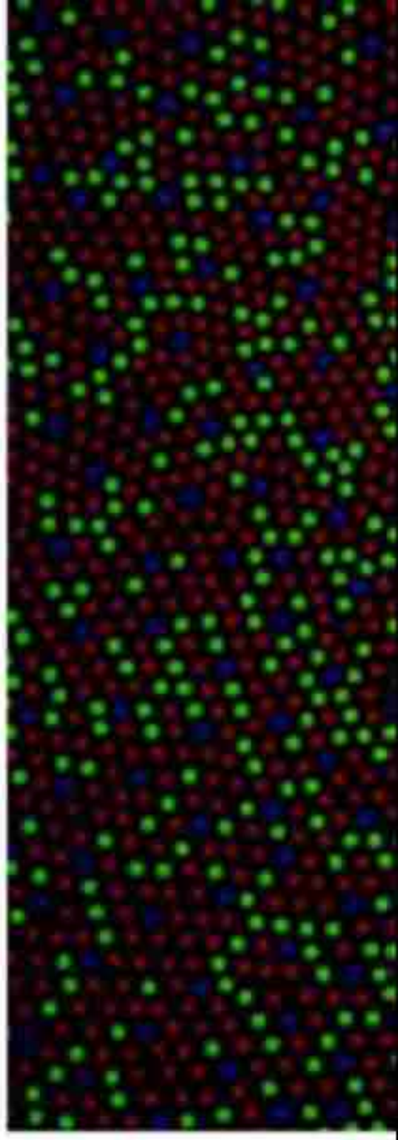

LIGHT
Choroid
Pigment epithelium
Rods and cones
Bipolar cells
Ganglion cells
Optic nerve axons
LIGHT

(b)

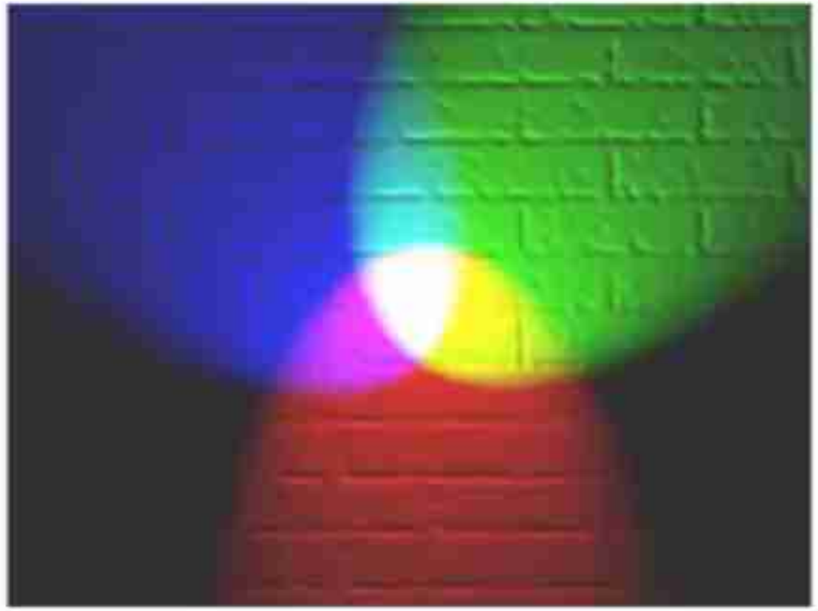

S
M
L
1.0
0.8
0.6
0.4
0.2
0
400
450
500
550
600
650
700

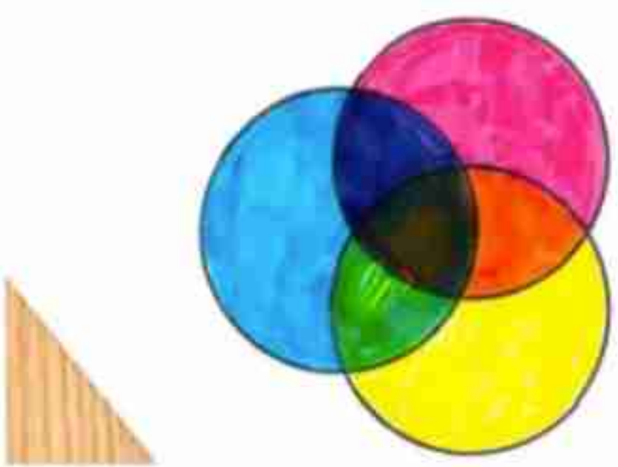


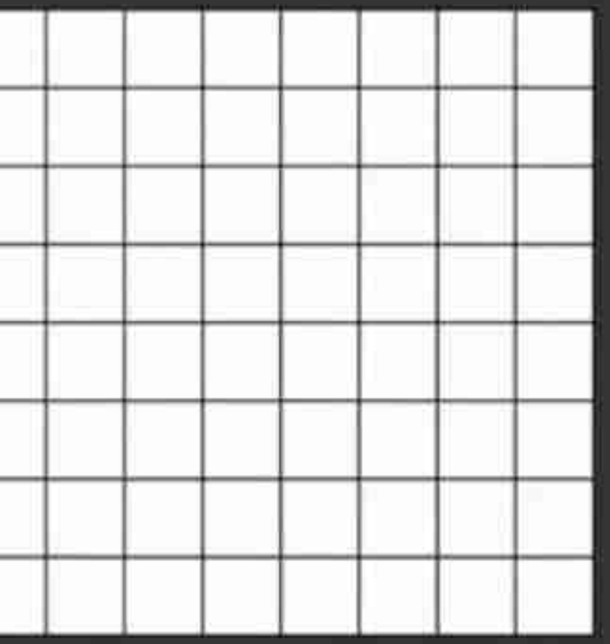

445 nm
535 nm
575 nm
"Blue"
"Green"
"Red"
Absorption
(arbitrary units)
400
450
500
550
600
650
700 nm
Bluish Purple
Purplish Blue
Blue
Bluegreen
Green
Yellowish Green
Yellow
Orange
Reddish Orange
Red

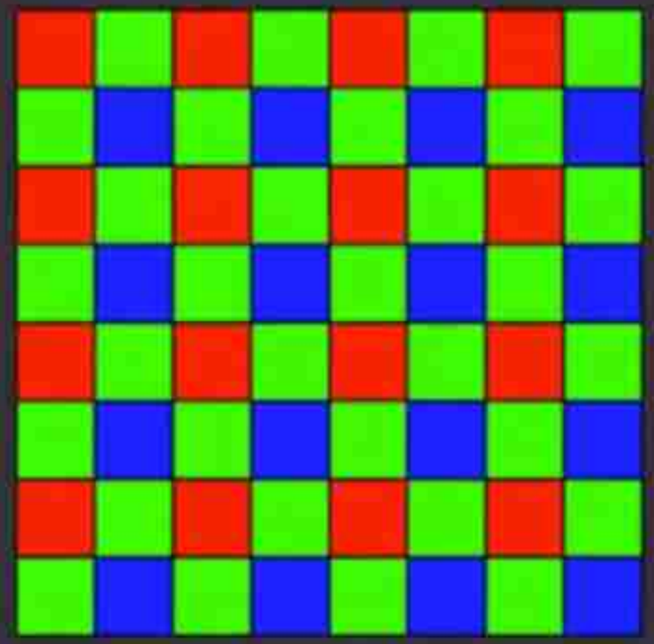


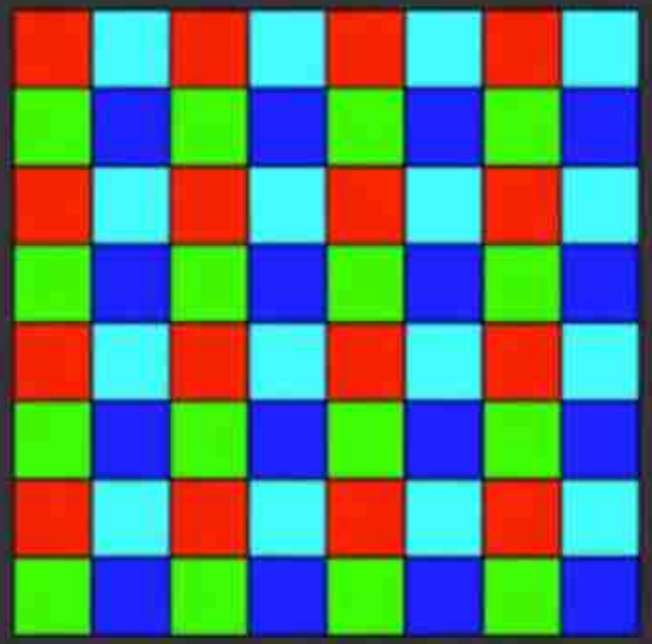
# Trichromatic Theory

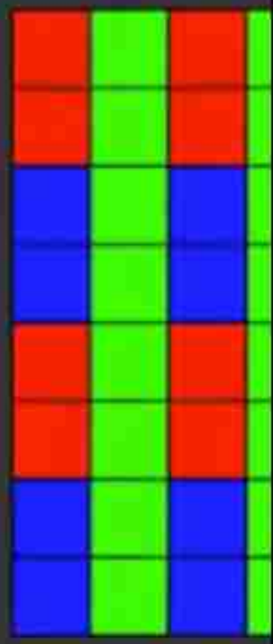










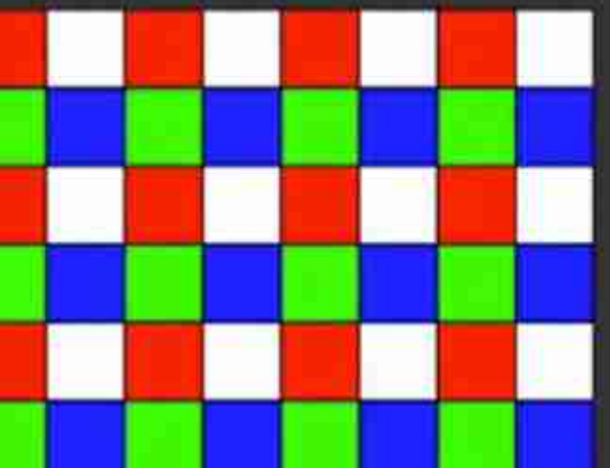


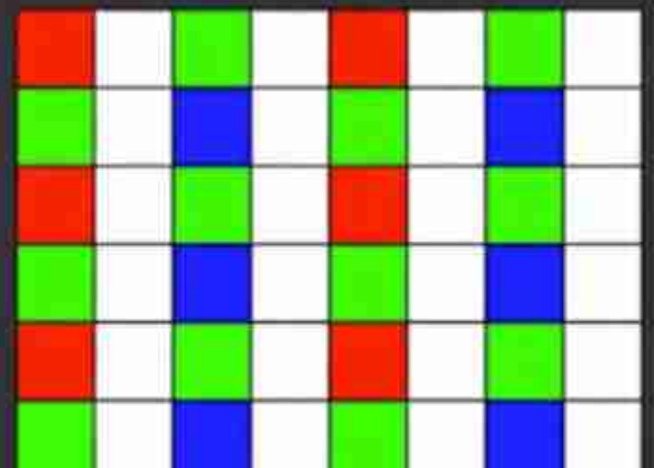
RGBW

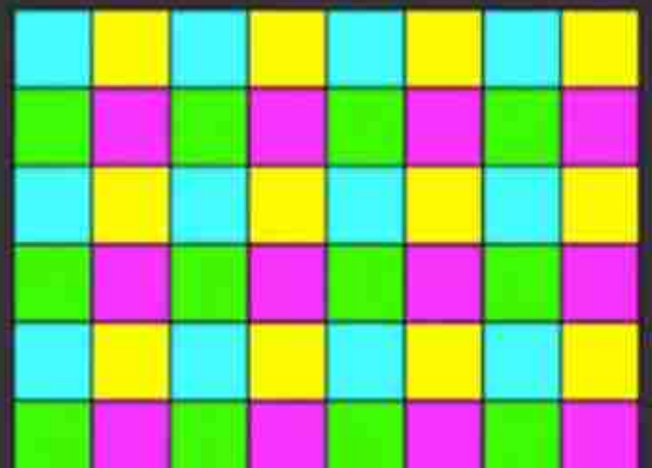
RGBW

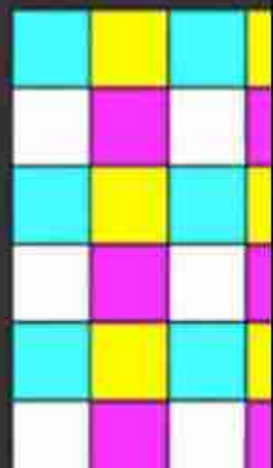
CYGM

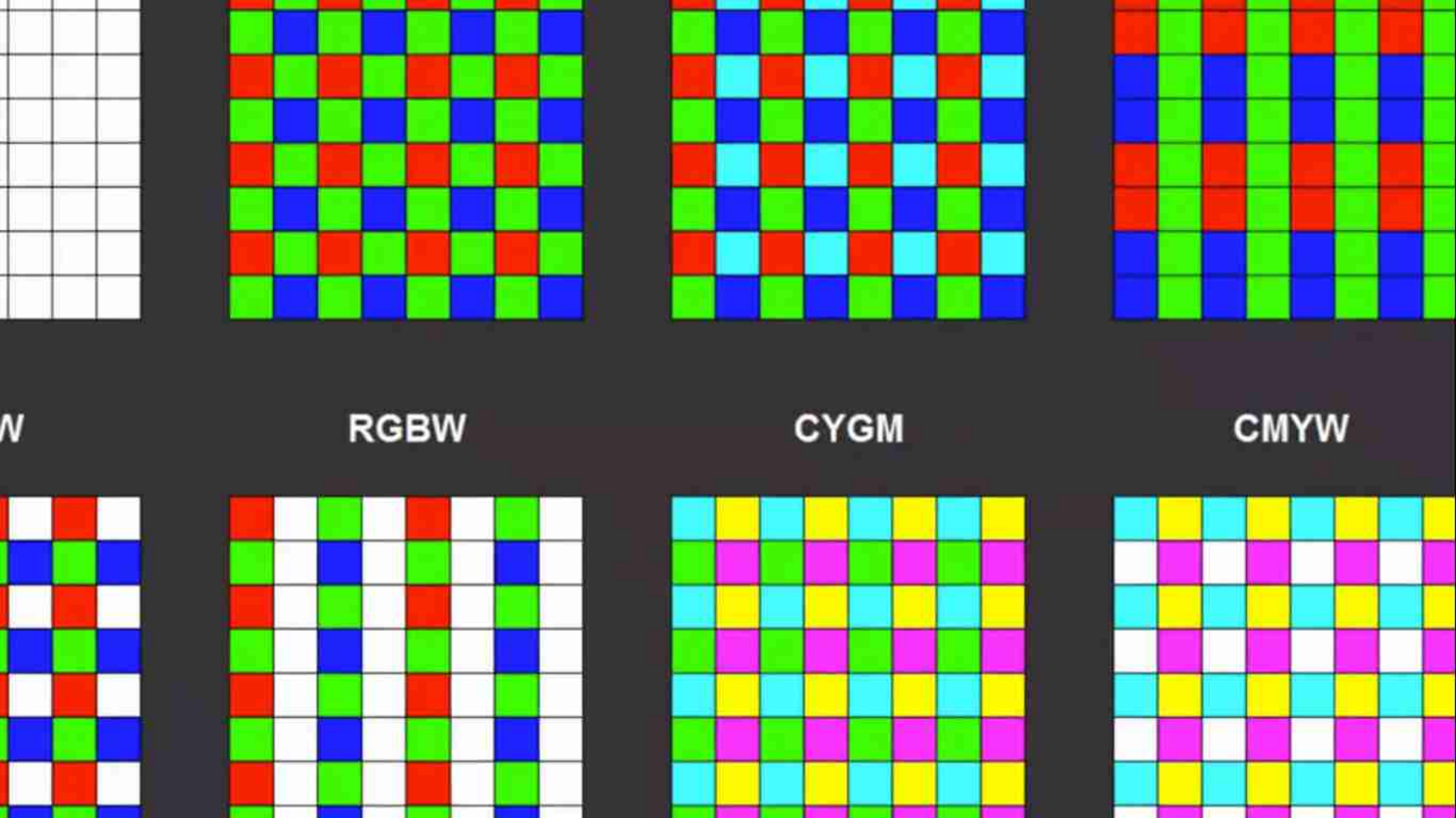
C

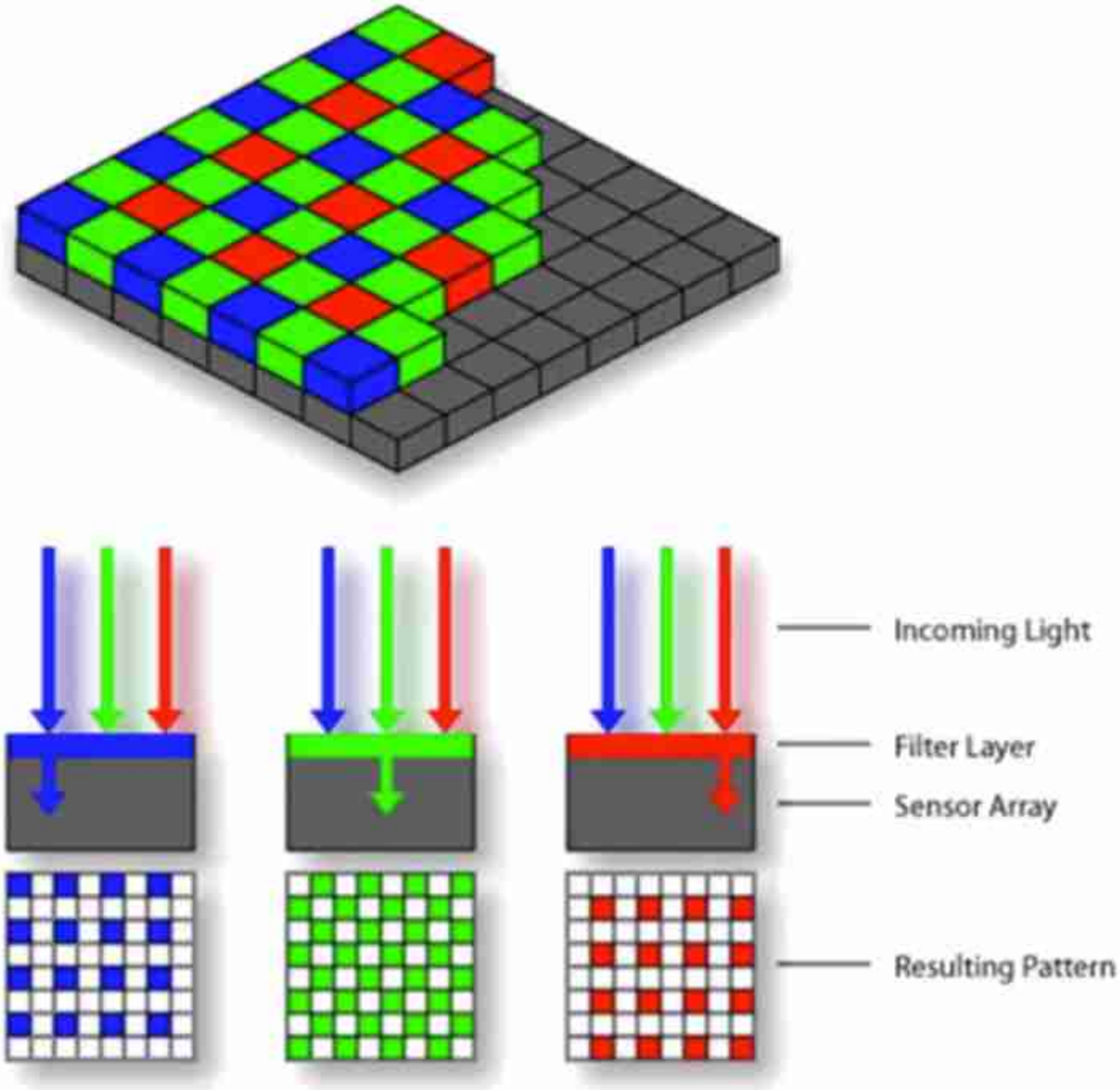


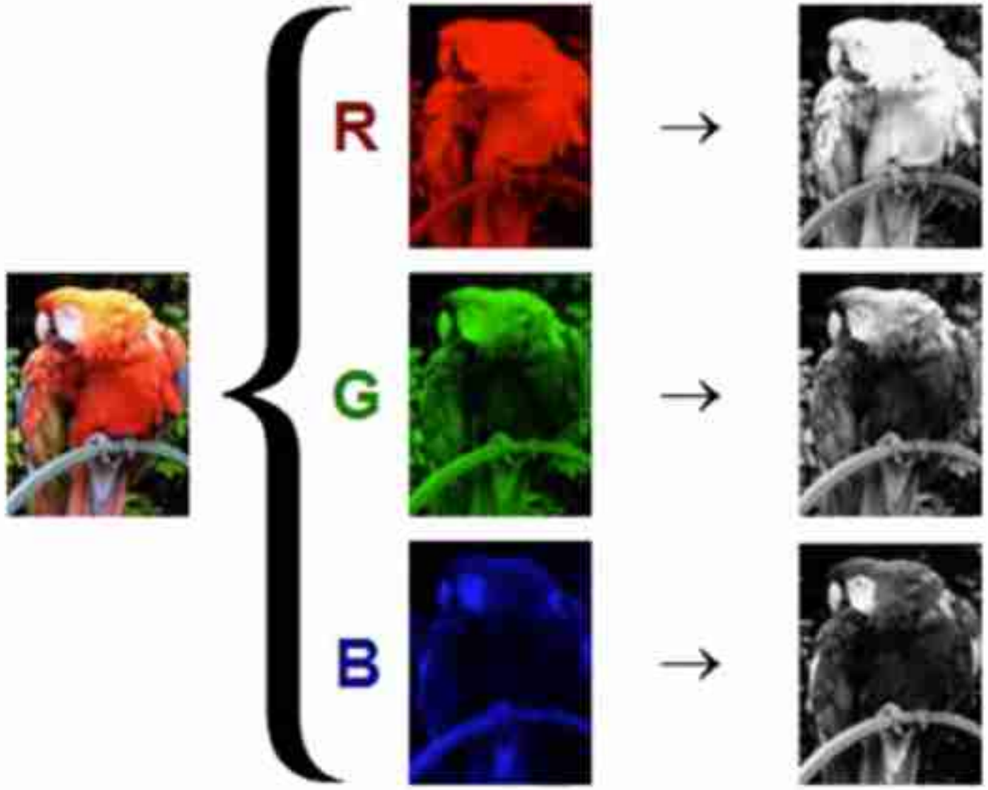


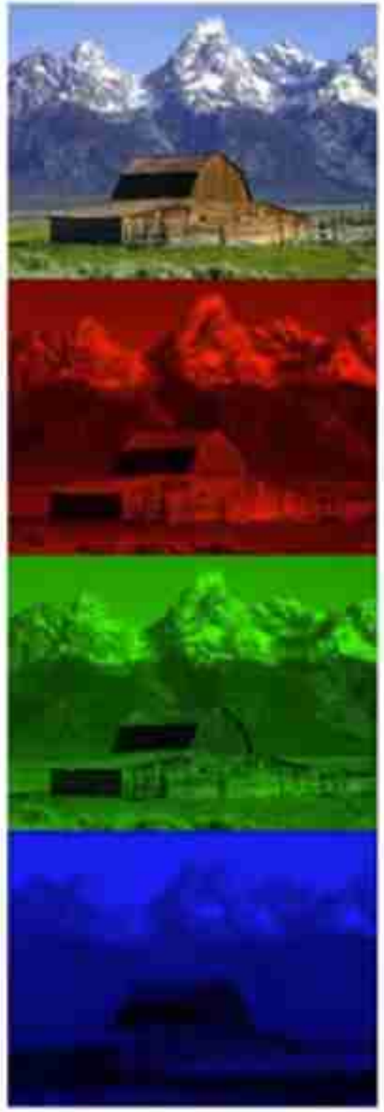





W
RGBW
CYGM
CMYW

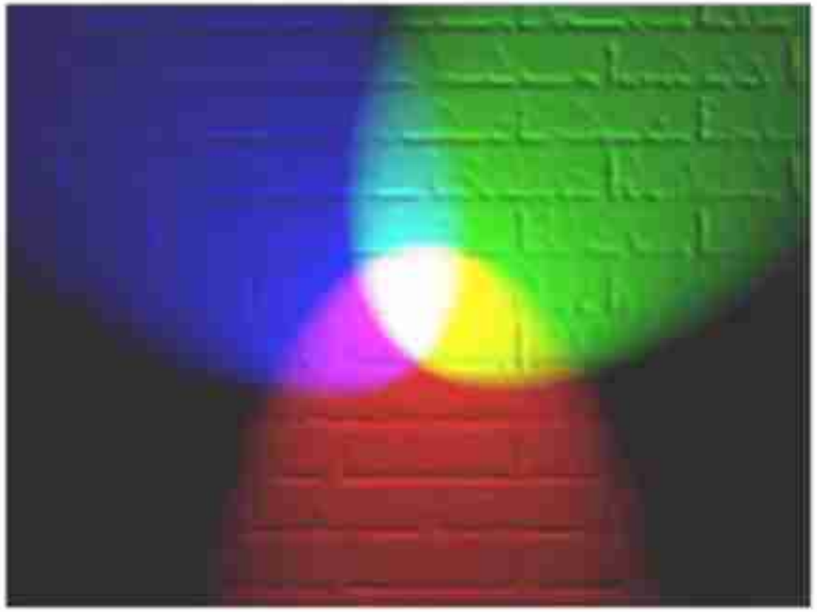

Incoming Light
Filter Layer
Sensor Array
Resulting Pattern


R
G
B

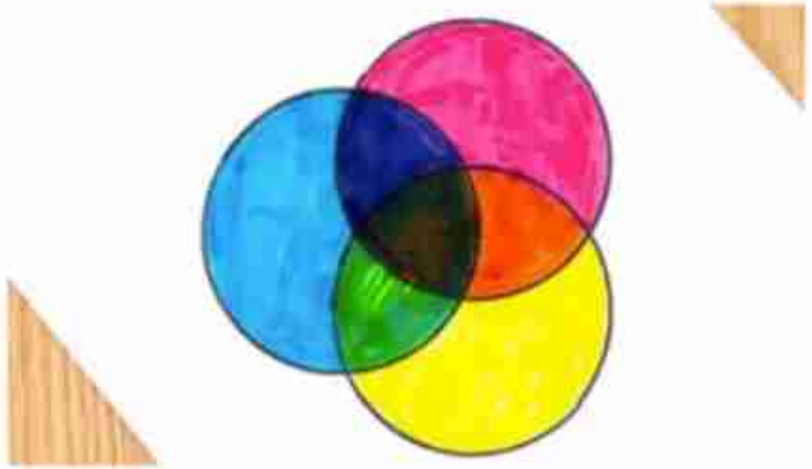


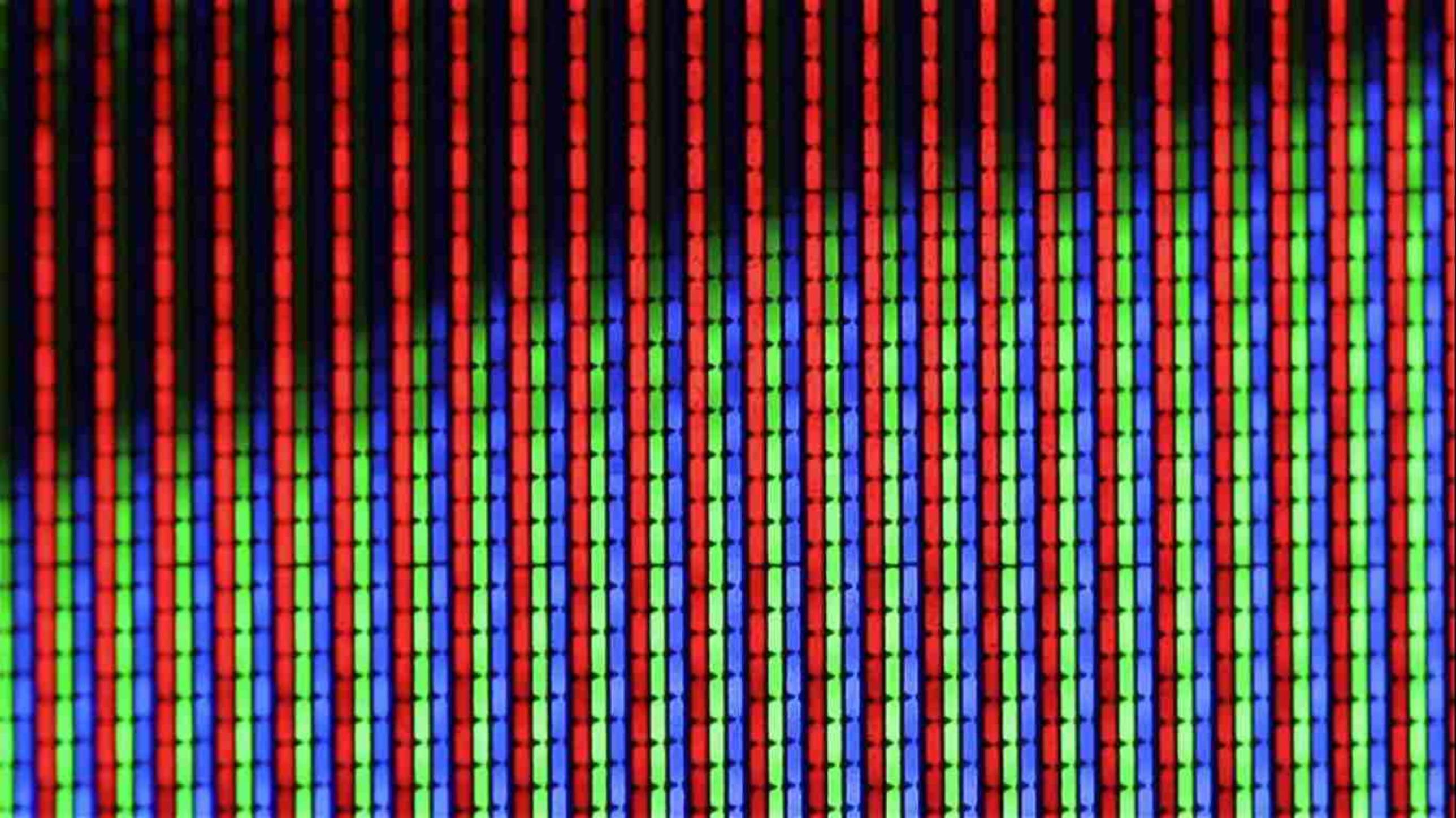
| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 39 | 79 | 74 | 44 | 42 | 23 | 43 | 54 | 48 | 61 | 255 | 255 | 69 | 60 | 34 | 30 | 49 | 39 | [illegible] | [illegible] |
| 50 | 73 | 68 | 34 | 24 | 0 | 15 | 19 | 8 | 16 | 255 | 255 | 17 | 13 | 0 | 0 | 27 | 25 | 29 | [illegible] |
| [illegible] | 77 | 72 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 255 | 255 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 51 | 79 | 67 | 47 | 38 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 34 | 51 | 46 | 44 | [illegible] |
| 45 | 61 | 58 | 33 | 10 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 0 | 21 | 29 | 35 | 47 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 48 | 54 | 57 | 50 | 49 | 255 | 152 | 218 | 237 | 243 | 244 | 246 | 216 | 169 | 255 | 42 | 52 | 56 | 52 | [illegible] |
| 42 | 49 | 48 | 30 | 5 | 255 | 90 | 147 | 157 | 150 | 154 | 156 | 133 | 92 | 255 | 0 | 20 | 36 | 44 | 48 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0 | 70 | 119 | 120 | 117 | 111 | 119 | 102 | 72 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 49 | 49 | 47 | 47 | 59 | 255 | 211 | 249 | 234 | 231 | 241 | 253 | 243 | 200 | 255 | 54 | 55 | 57 | 52 | [illegible] |
| 43 | 43 | 38 | 23 | 0 | 255 | 126 | 118 | 136 | 148 | 134 | 147 | 143 | 113 | 255 | 0 | 16 | 36 | 45 | 44 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0 | 106 | 129 | 97 | 83 | 88 | 105 | 110 | 88 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 53 | 48 | 39 | 41 | 67 | 255 | 238 | 255 | 230 | 236 | 250 | 255 | 252 | 224 | 255 | 68 | 48 | 48 | 51 | [illegible] |
| 47 | 42 | 33 | 13 | 0 | 255 | 139 | 143 | 144 | 143 | 134 | 137 | 140 | 128 | 255 | 6 | 7 | 27 | 42 | 43 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0 | 116 | 118 | 73 | 68 | 79 | 92 | 102 | 94 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 57 | 57 | 40 | 39 | 92 | 255 | 238 | 255 | 245 | 255 | 255 | 250 | 255 | 240 | 255 | 72 | 41 | 47 | [illegible] | [illegible] |
| 51 | 46 | 34 | 9 | 14 | 255 | 139 | 143 | 149 | 139 | 138 | 139 | 142 | 135 | 255 | 8 | 0 | 23 | 47 | 50 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0 | 111 | 114 | 81 | 82 | 79 | 76 | 102 | 104 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 57 | 54 | 49 | 47 | 93 | 255 | 206 | 245 | 249 | 255 | 252 | 244 | 253 | 237 | 255 | 71 | 46 | 41 | [illegible] | [illegible] |
| 58 | 55 | 45 | 19 | 22 | 255 | 128 | 136 | 135 | 136 | 128 | 133 | 140 | 132 | 255 | 9 | 5 | 40 | 65 | 62 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0 | 86 | 107 | 91 | 88 | 77 | 73 | 98 | 105 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 68 | 67 | 60 | 54 | 61 | 255 | 180 | 222 | 236 | 244 | 249 | 255 | 232 | 218 | 255 | 67 | 57 | 53 | [illegible] | [illegible] |
| 64 | 63 | 56 | 33 | 1 | 255 | 74 | 100 | 132 | 136 | 137 | 148 | 128 | 123 | 255 | 10 | 19 | 61 | 84 | 73 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 0 | 58 | 99 | 93 | 90 | 89 | 103 | 89 | 98 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 89 | 77 | 68 | 60 | 57 | 255 | 180 | 193 | 252 | 253 | 245 | 250 | 223 | 180 | 255 | 65 | 57 | 63 | [illegible] | [illegible] |
| 98 | 76 | 57 | 44 | 15 | 255 | 41 | 100 | 172 | 168 | 159 | 161 | 138 | 68 | 255 | 16 | 25 | 47 | 59 | 47 |
| 96 | 74 | [illegible] | [illegible] | [illegible] | 0 | 36 | 96 | 137 | 127 | 114 | 121 | 104 | 45 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 74 | 64 | 51 | 54 | 38 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 46 | 35 | 34 | 33 | [illegible] |
| 74 | 63 | 51 | 44 | 14 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 255 | 6 | 10 | 21 | 30 | 23 |
| 75 | 61 | [illegible] | [illegible] | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 51 | 49 | 49 | 45 | 28 | 24 | 19 | 25 | 42 | 58 | 61 | 53 | 50 | 34 | 31 | 33 | 22 | 17 | 21 | [illegible] |
| 53 | 49 | 45 | 41 | 18 | 10 | 1 | 0 | 10 | 21 | 17 | 7 | 4 | 0 | 0 | 4 | 3 | 7 | 18 | 23 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |

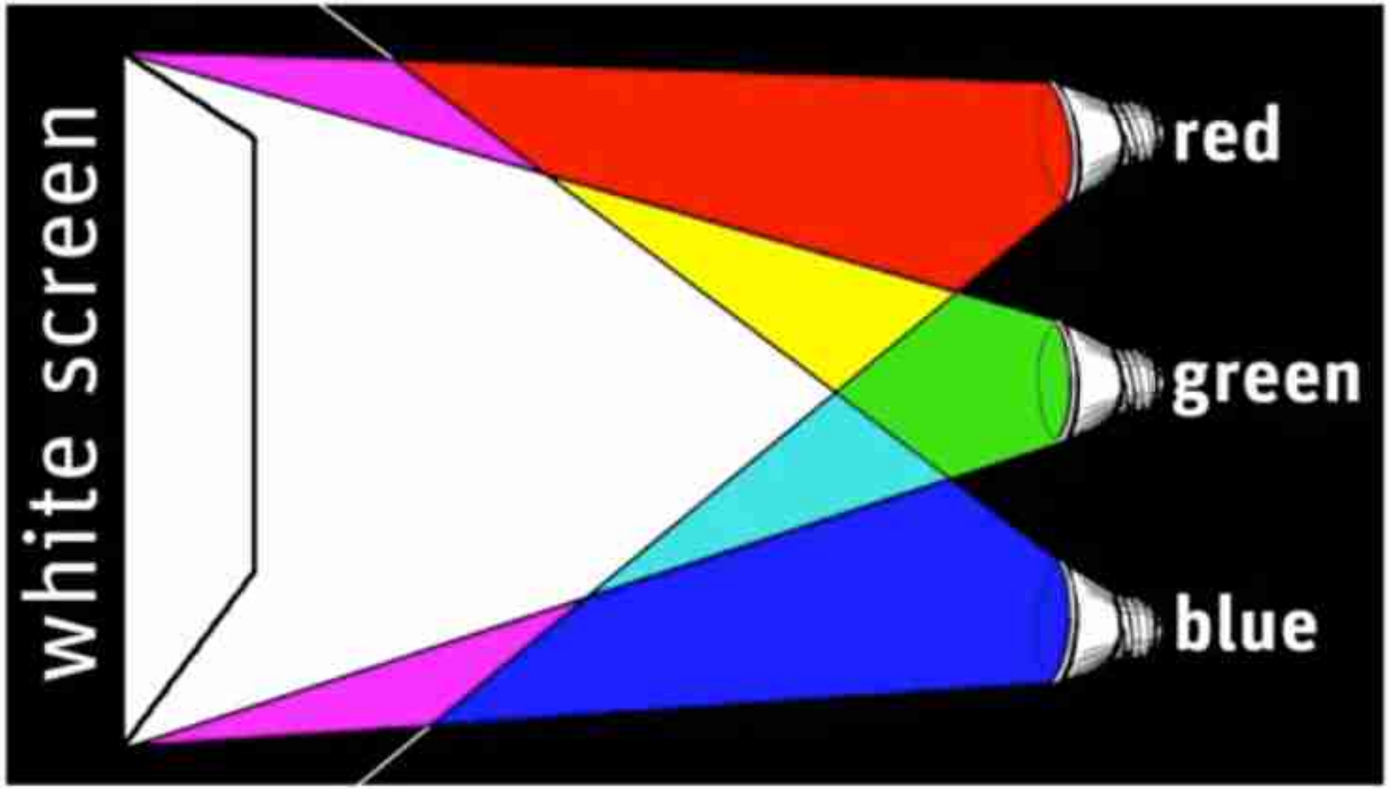
## Additive Mixing

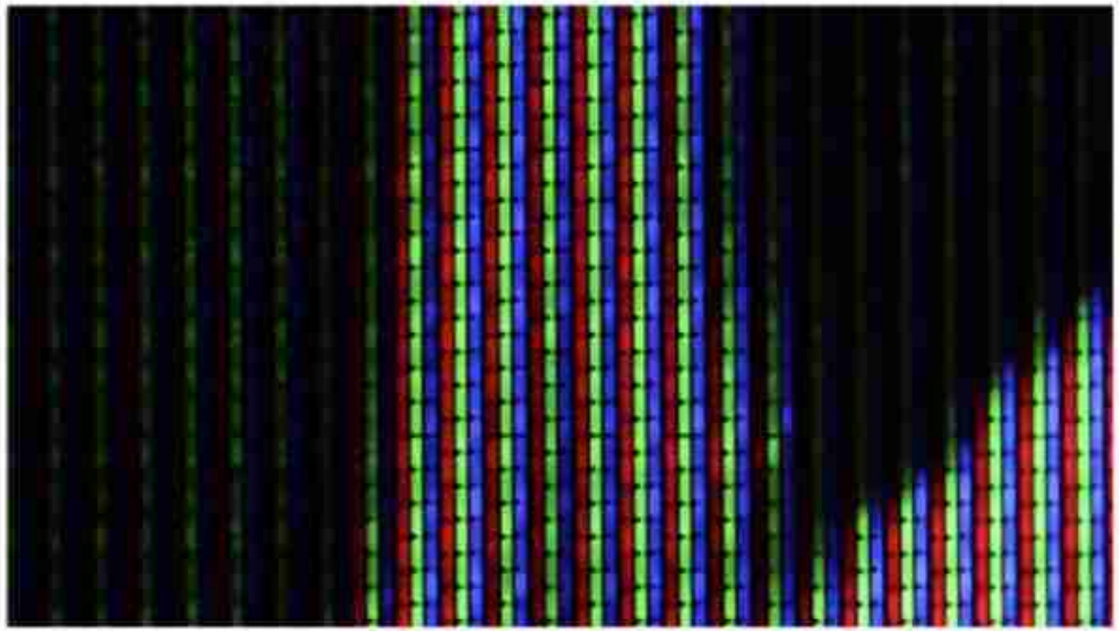


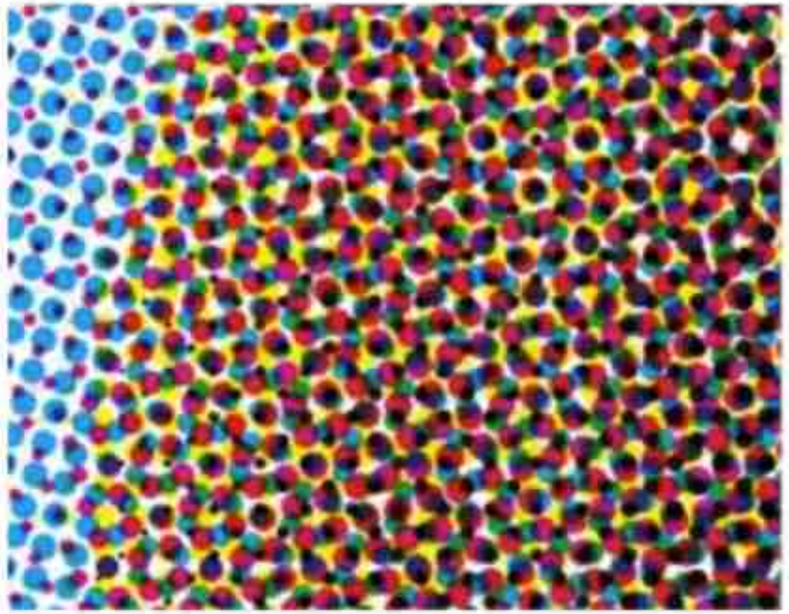
## Subtractive Mixing

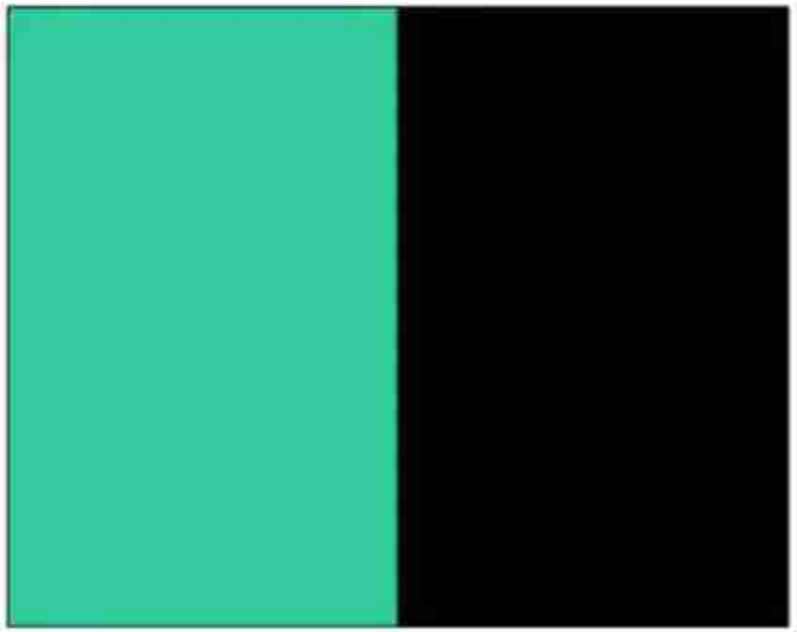


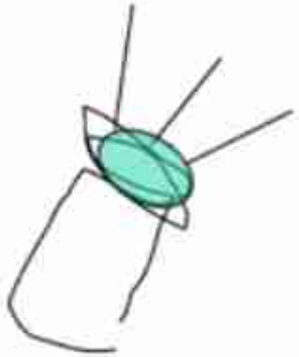


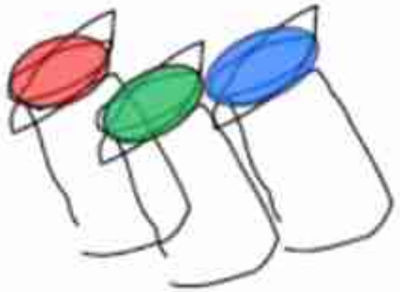

white screen
red
green
blue



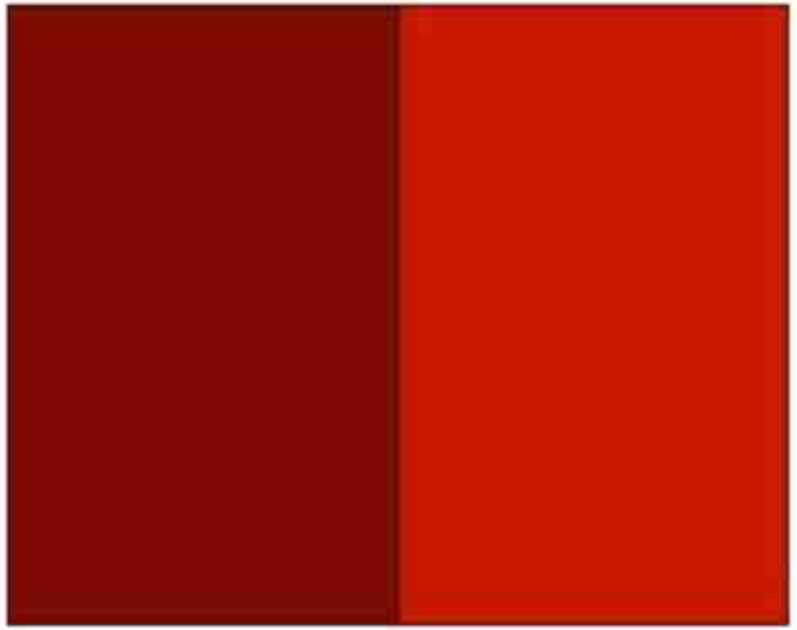


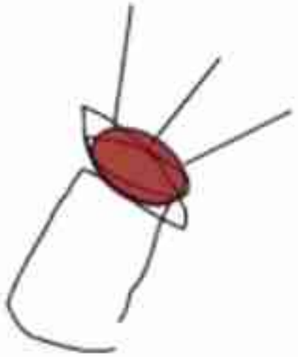


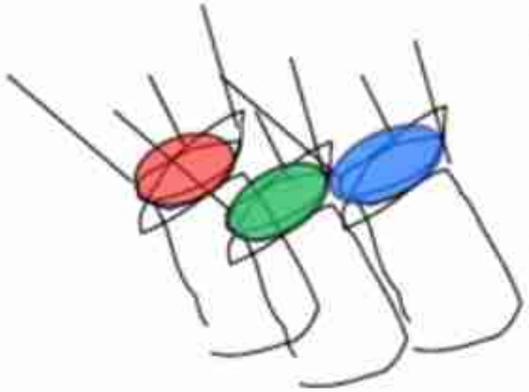


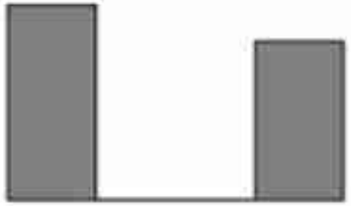


Source: W. Freeman

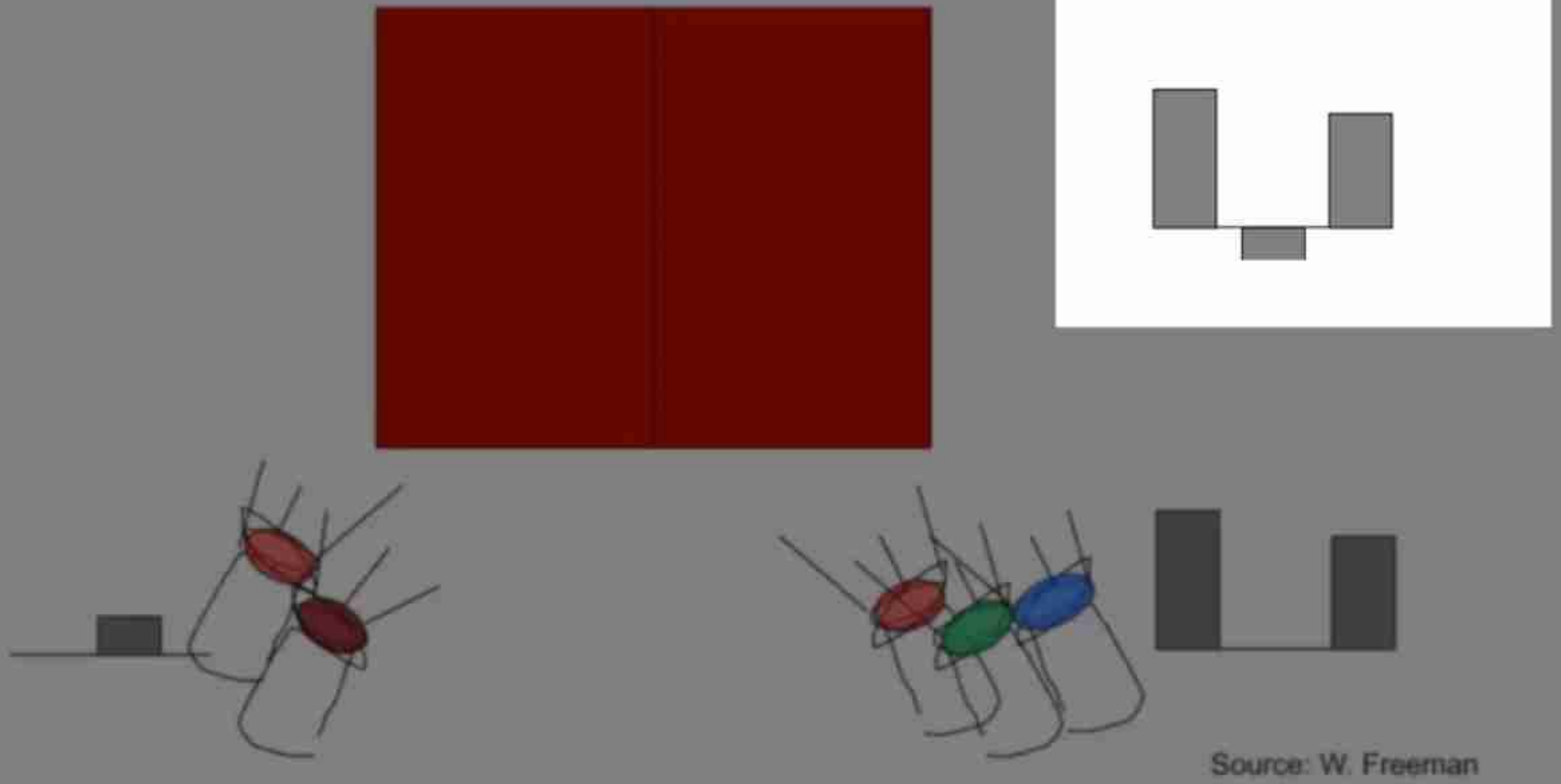


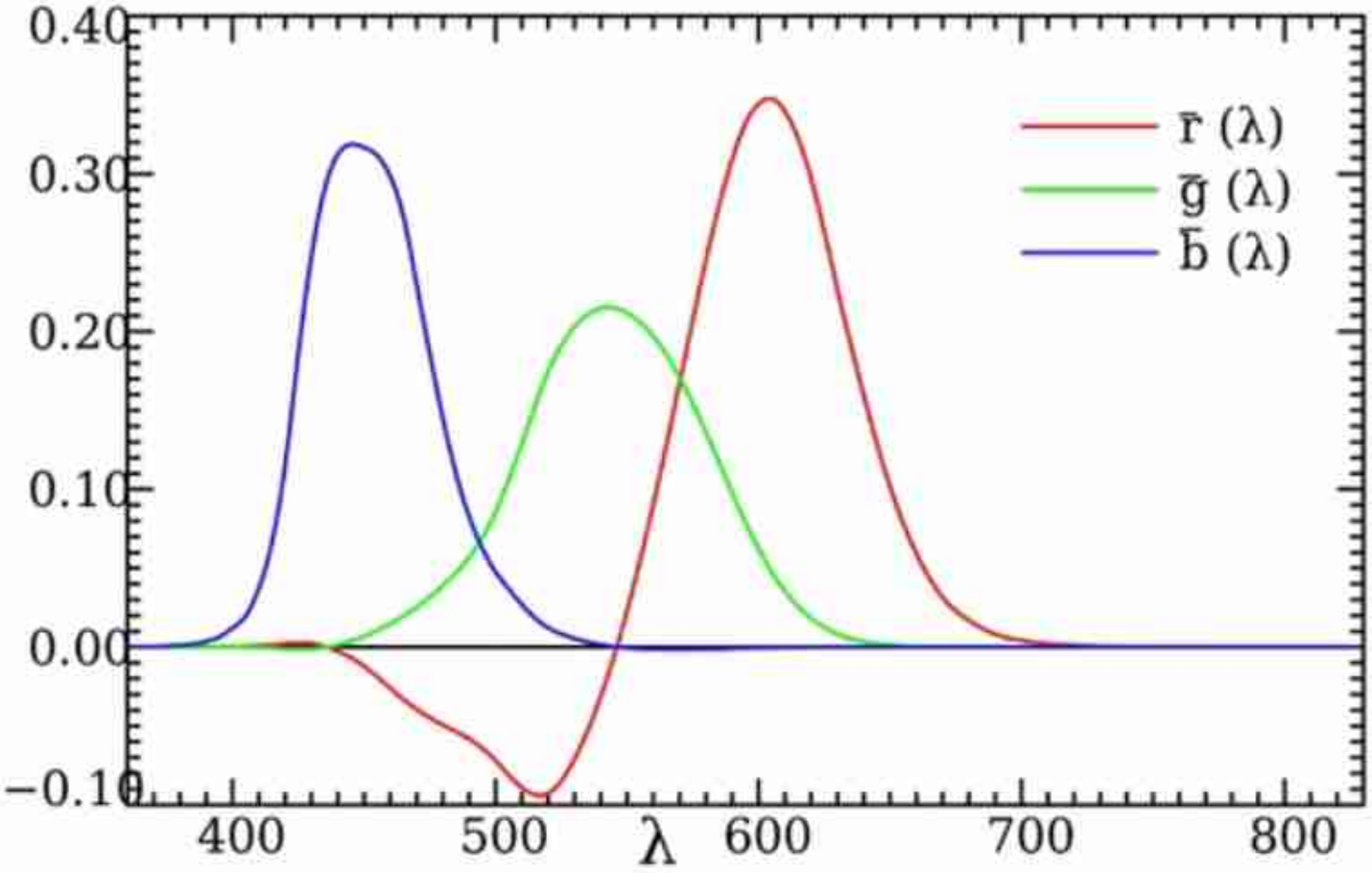


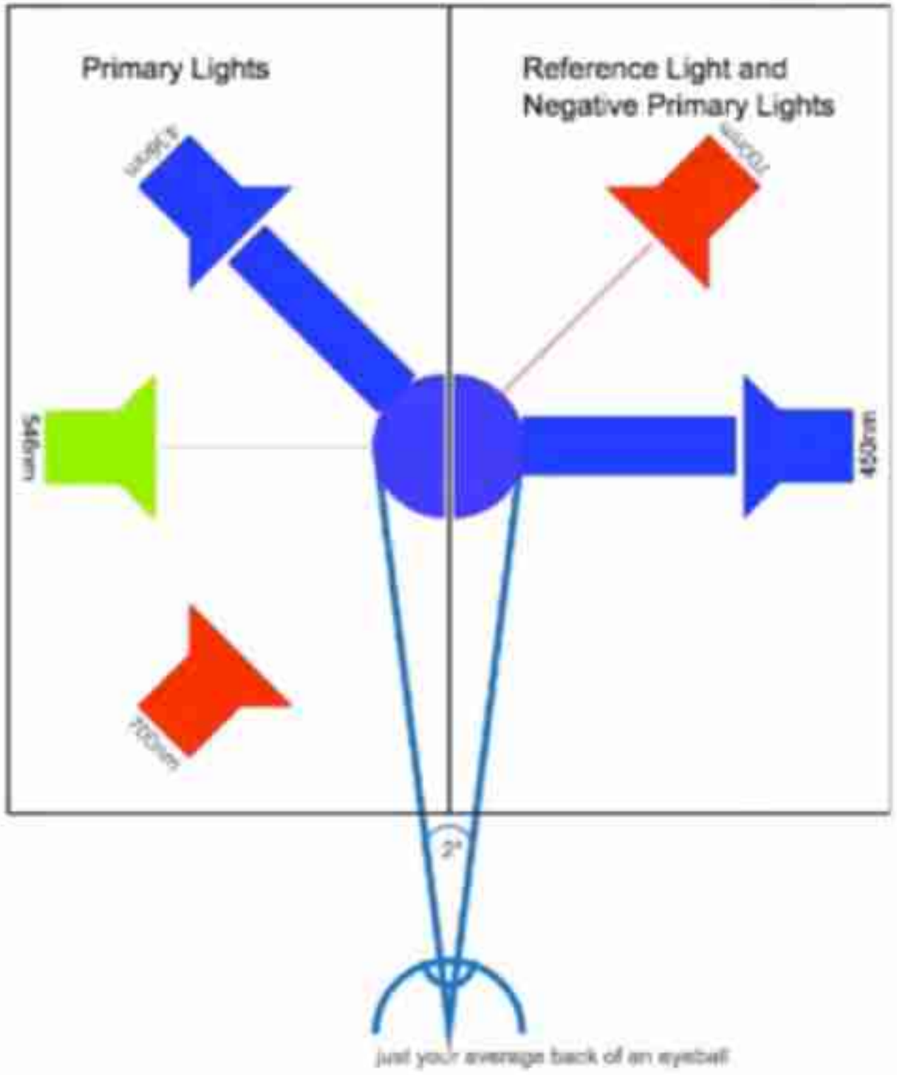


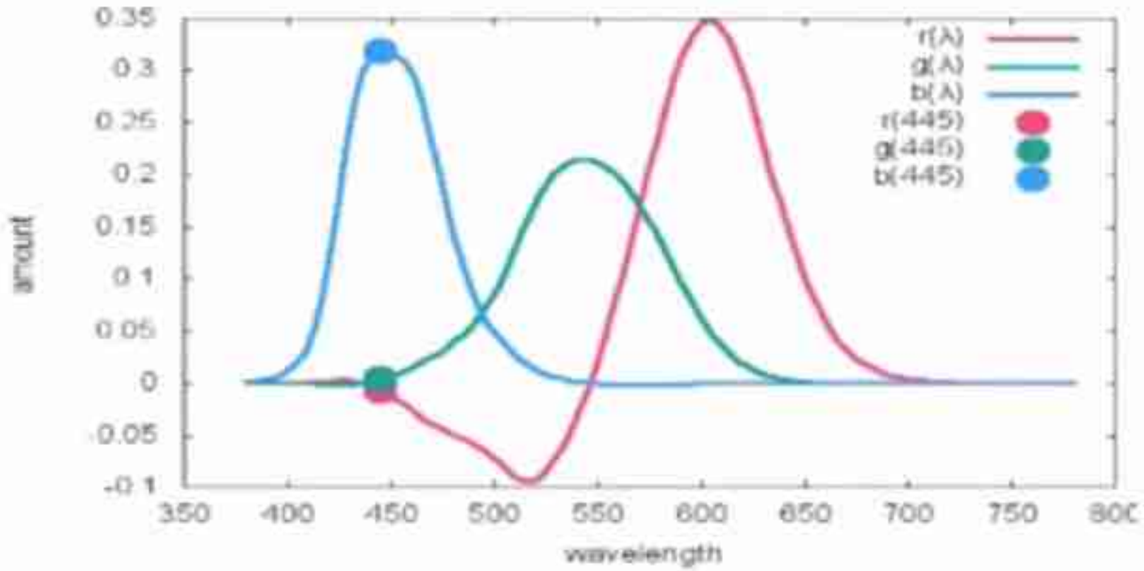


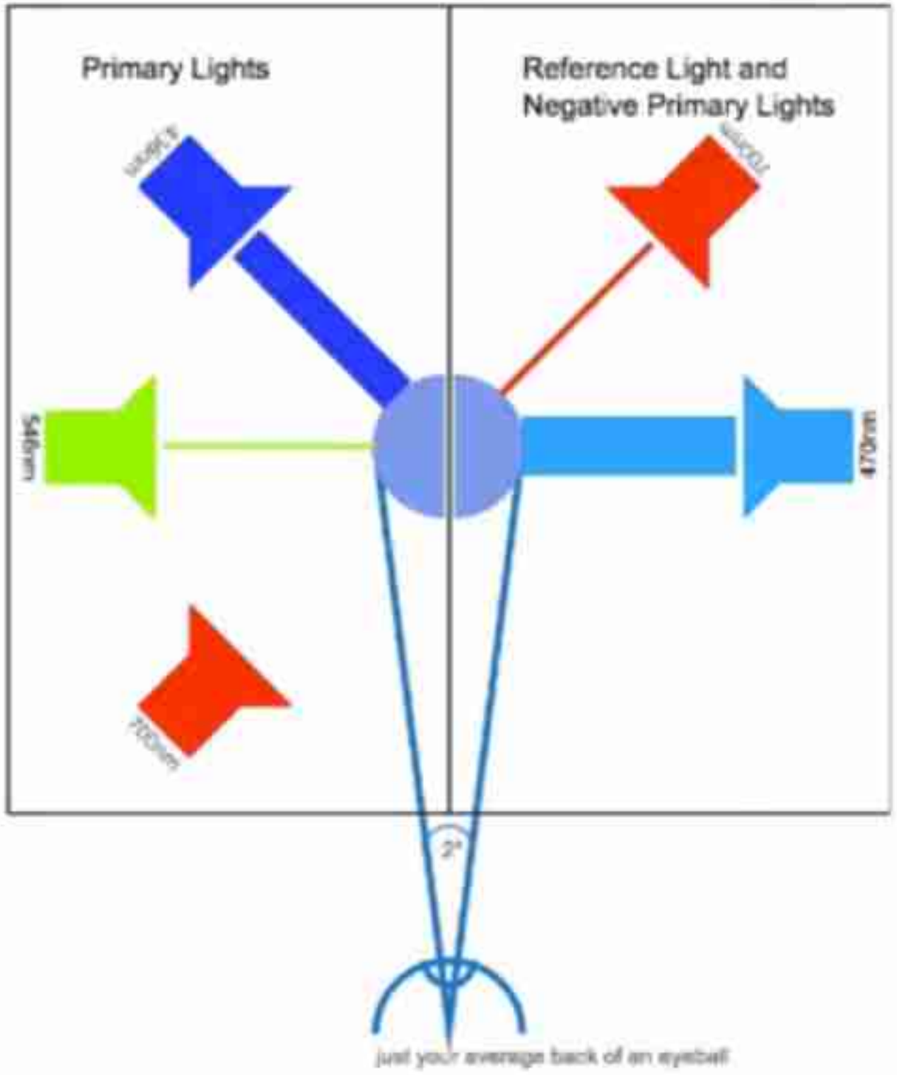
Source: W. Freeman

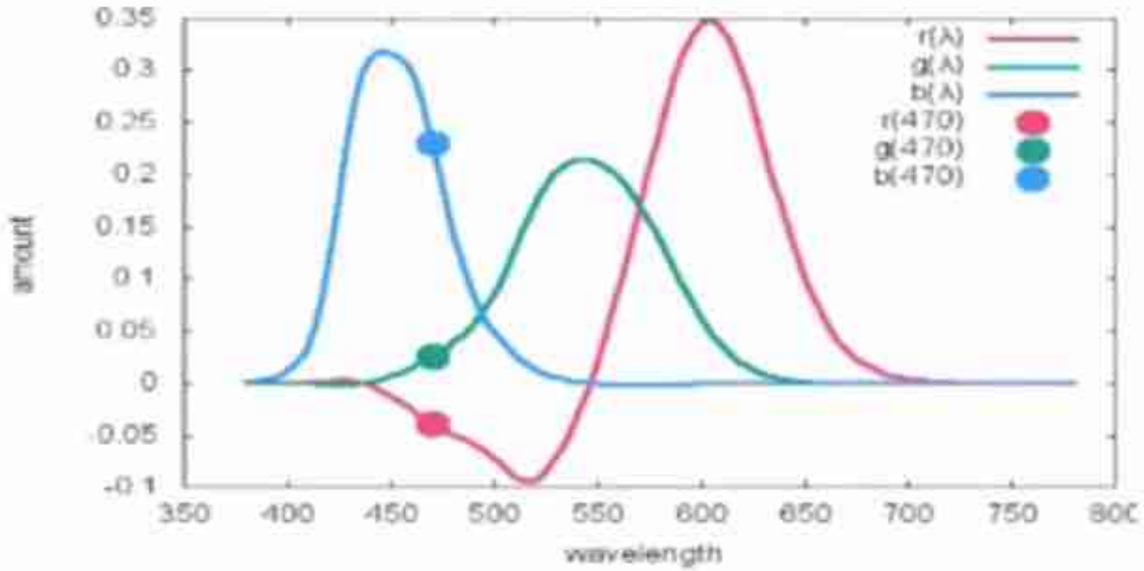

Source: W. Freeman


0.40
0.30
0.20
0.10
0.00
−0.10
400
500
600
700
800
λ
r̄ (λ)
ḡ (λ)
b̄ (λ)

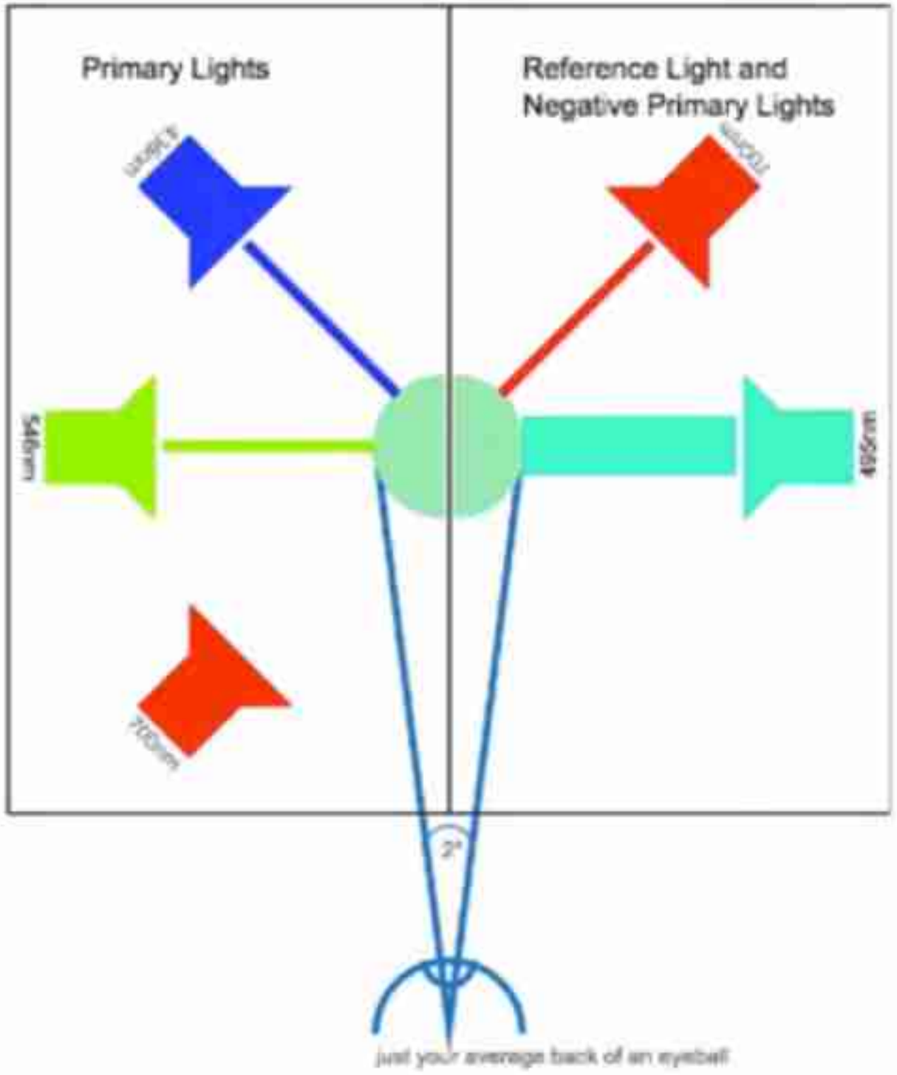

Primary Lights
Reference Light and
Negative Primary Lights
450nm
2°
just your average back of an eyeball

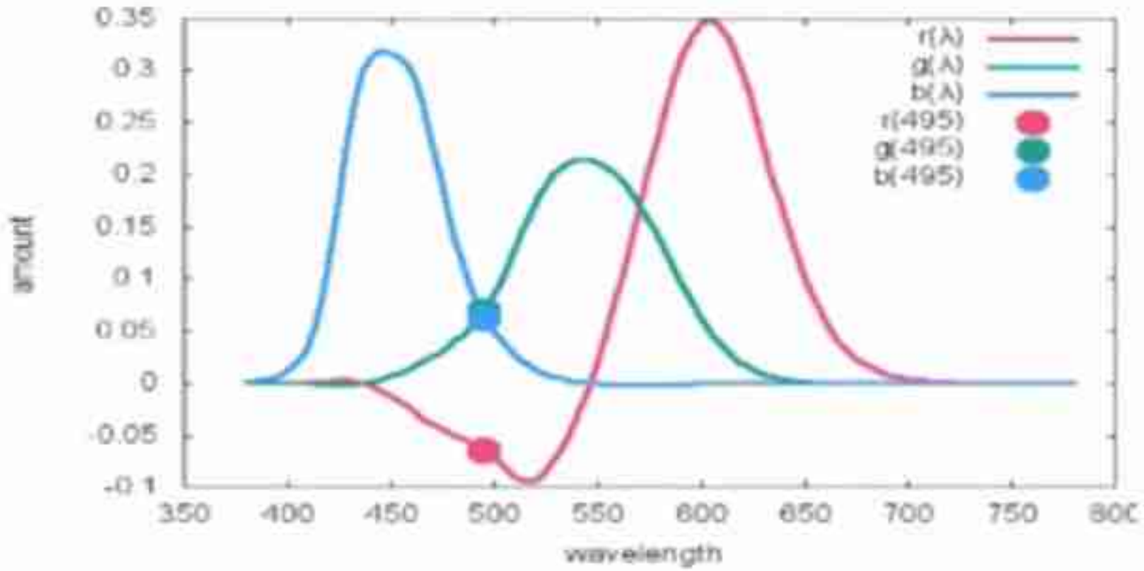

r(λ)
g(λ)
b(λ)
r(445)
g(445)
b(445)
amount
wavelength
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
350
400
450
500
550
600
650
700
750
800



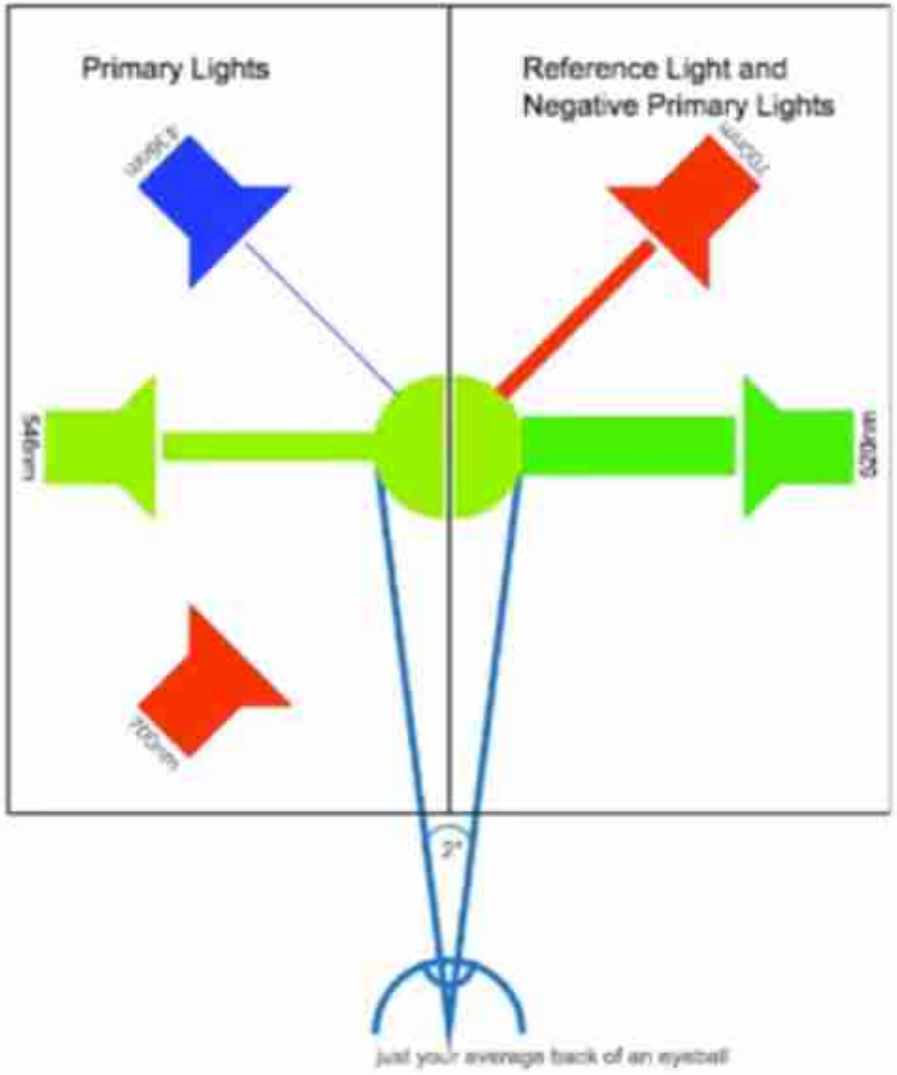



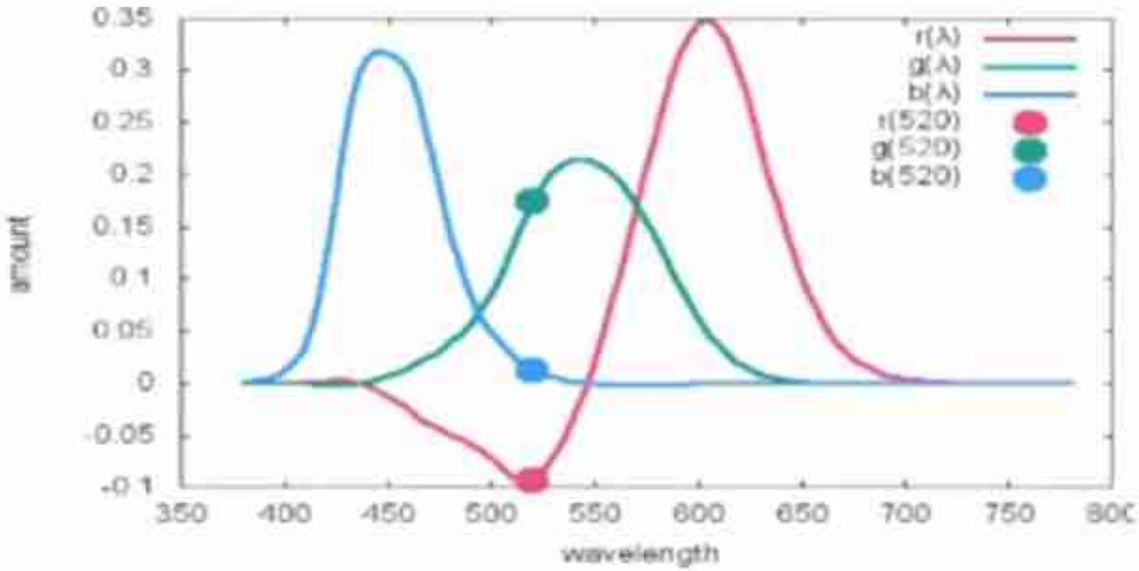

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



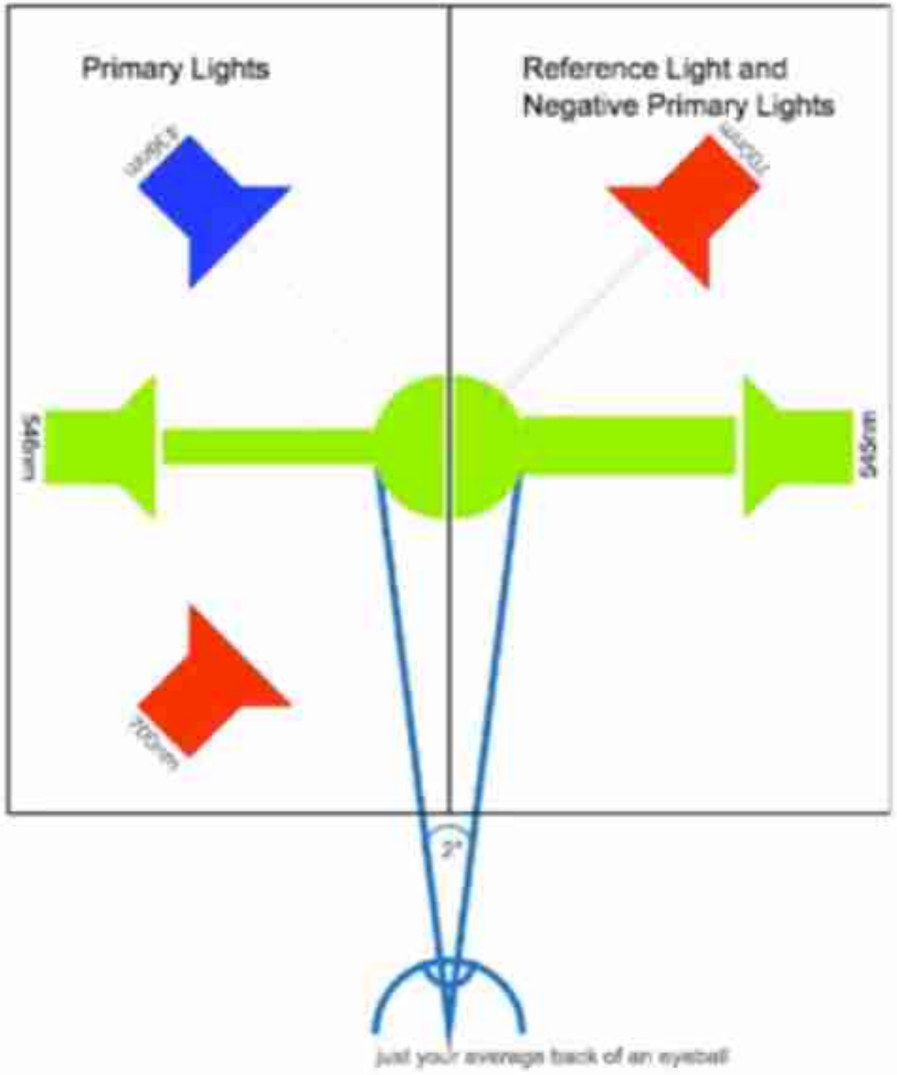



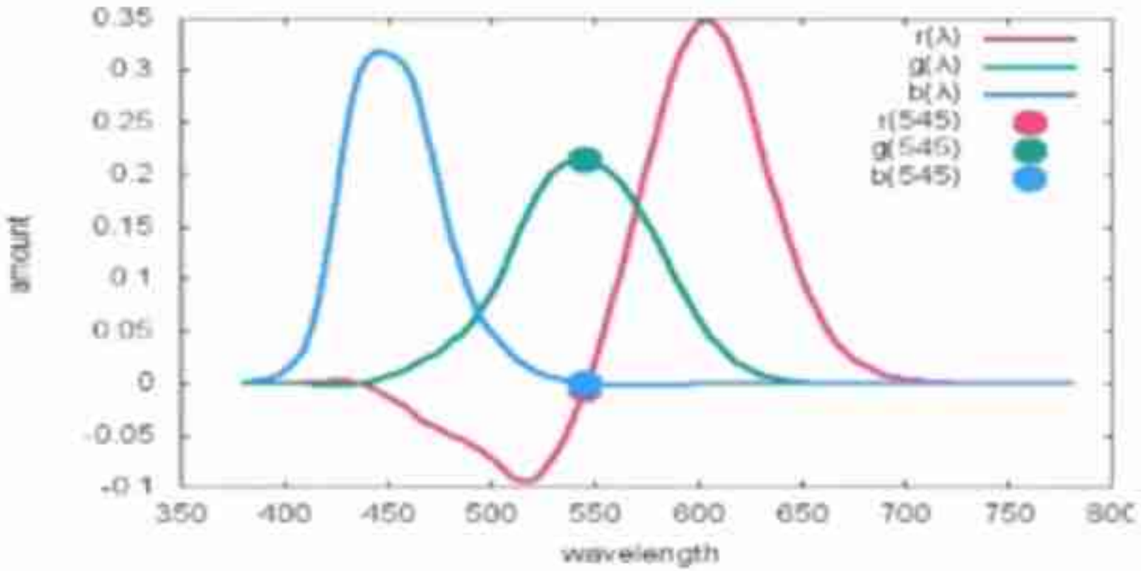

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



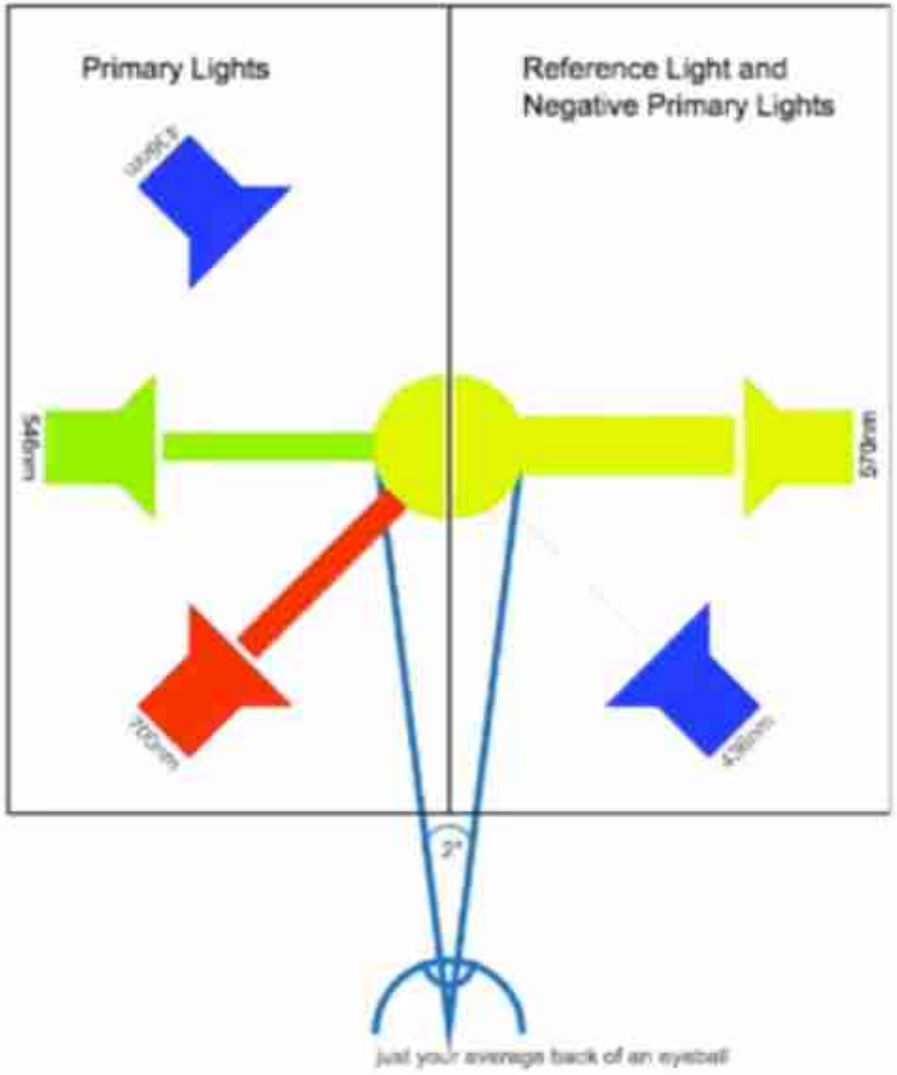



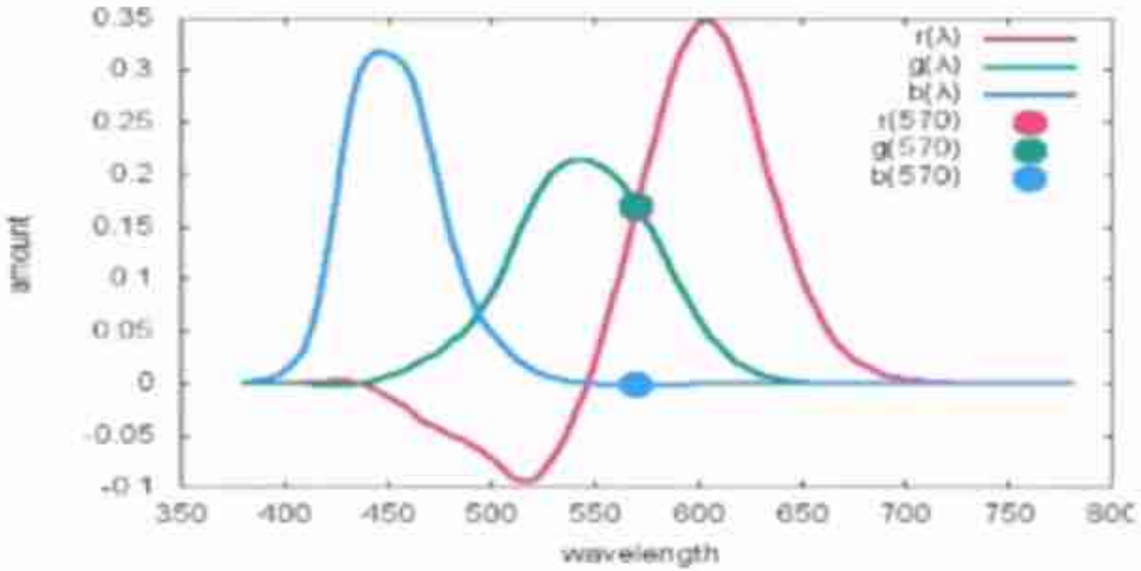

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



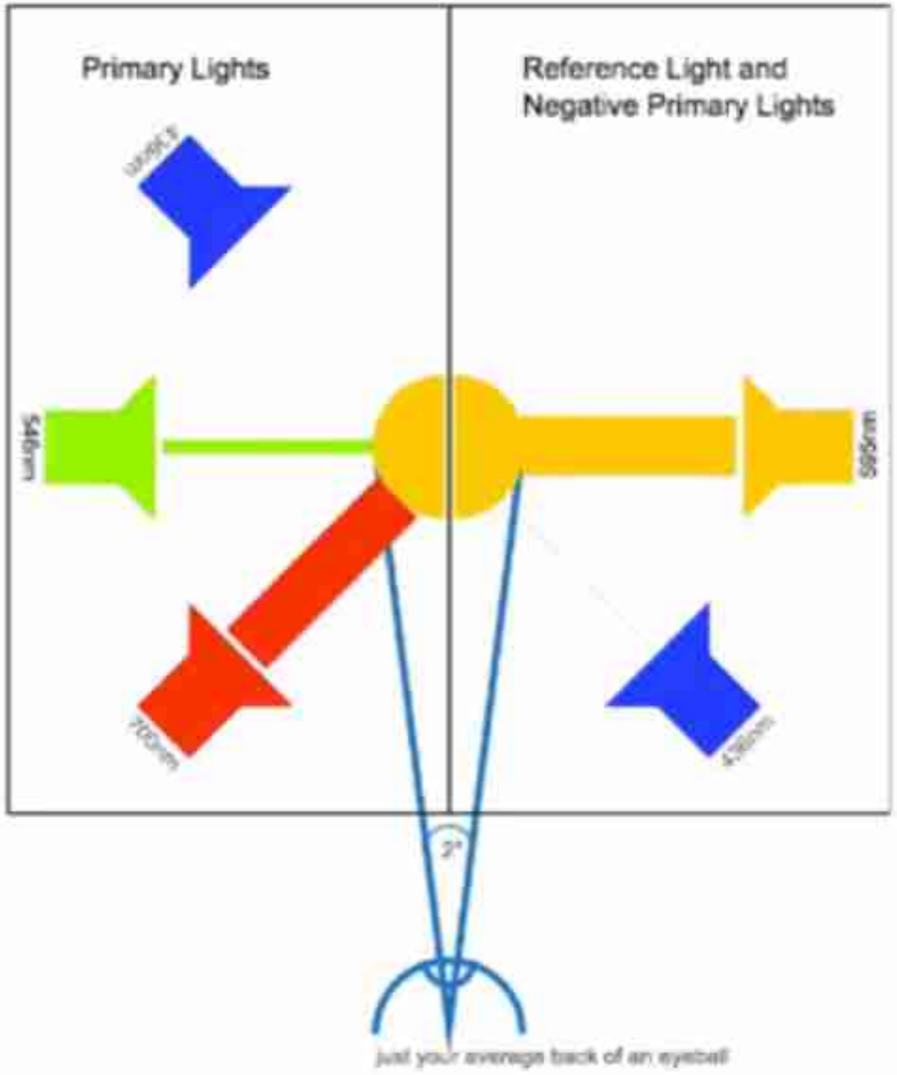



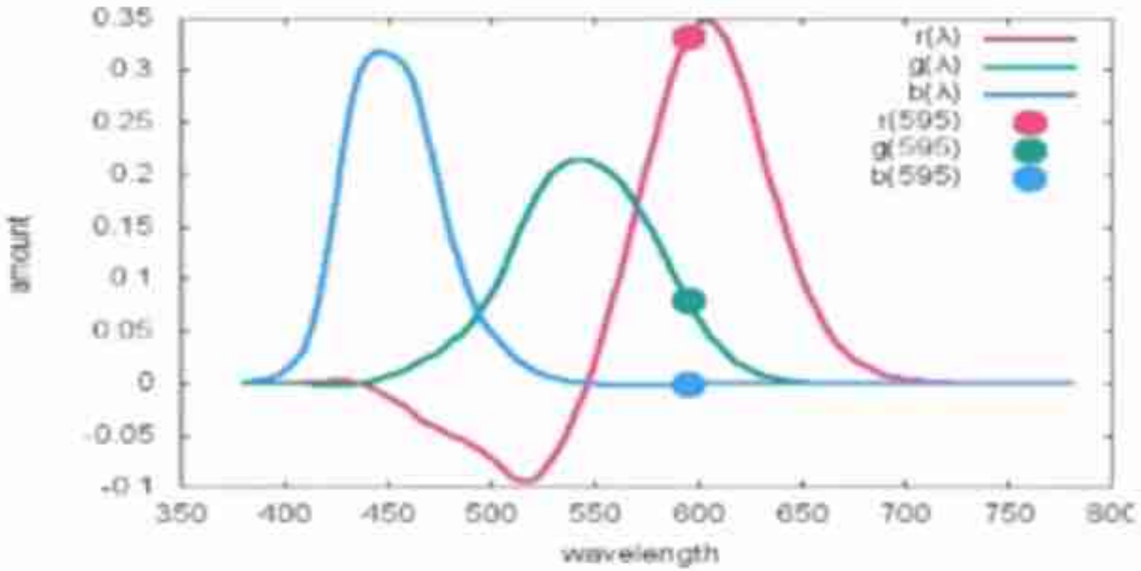

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



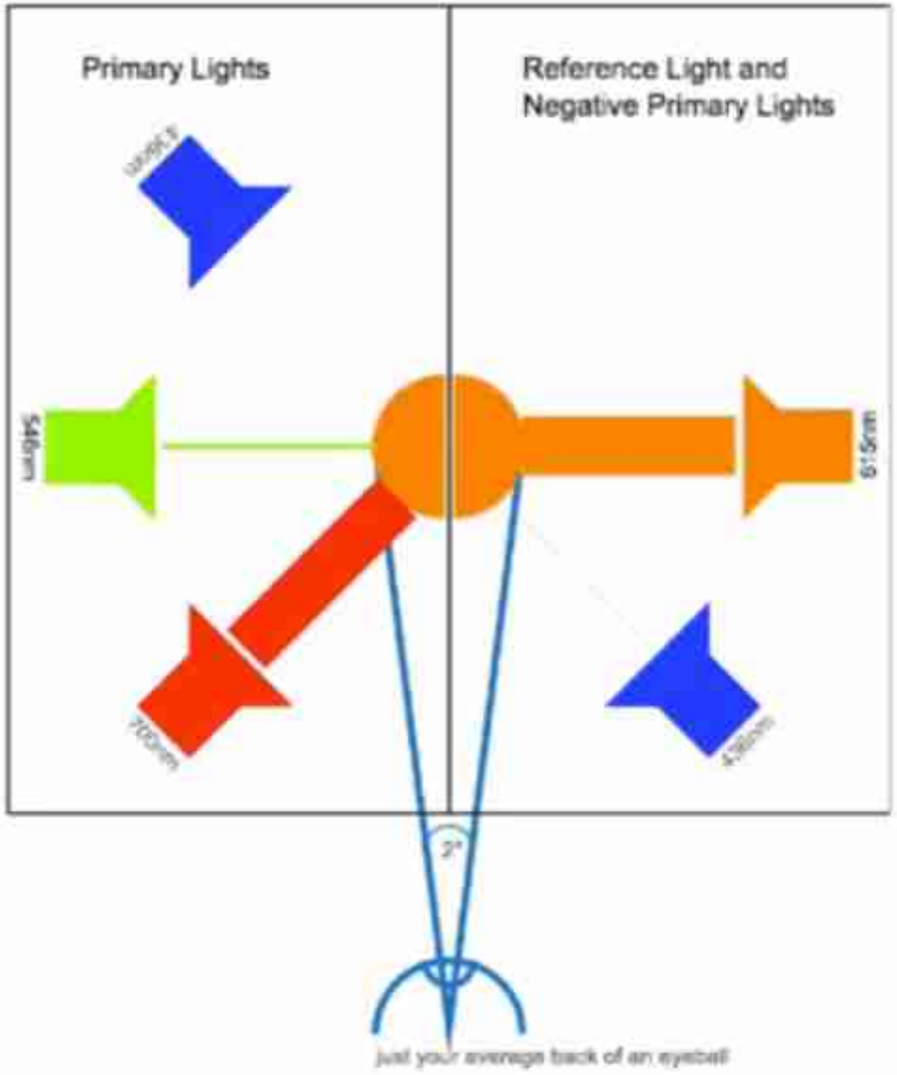



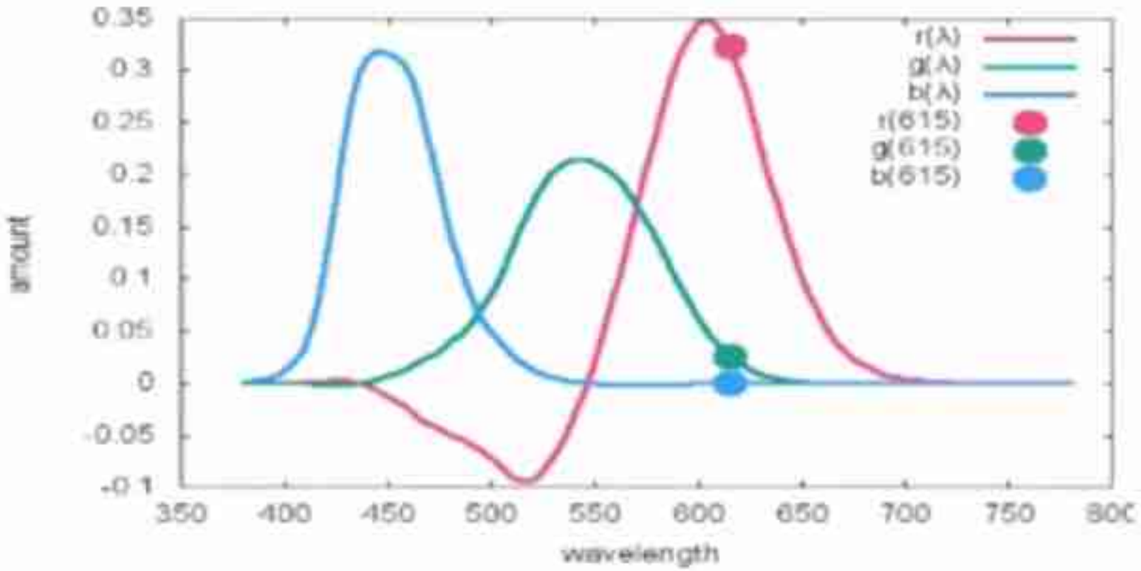

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



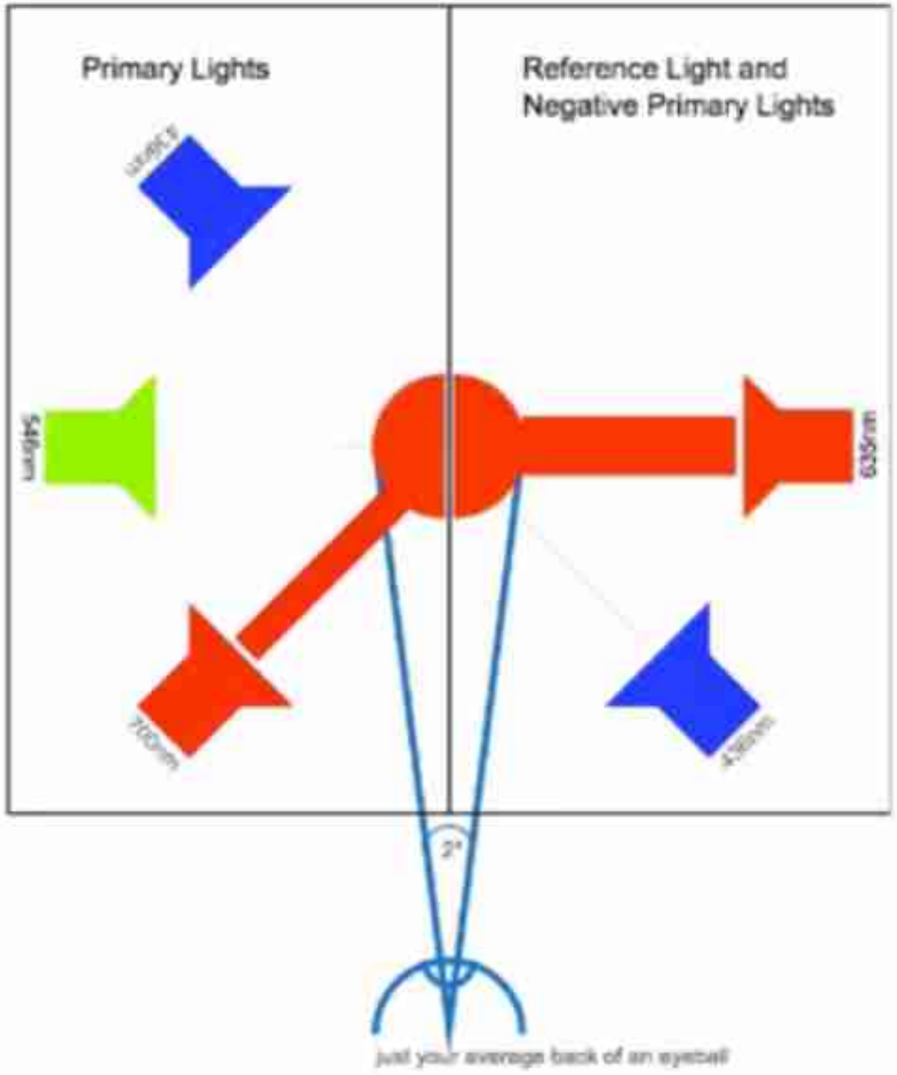



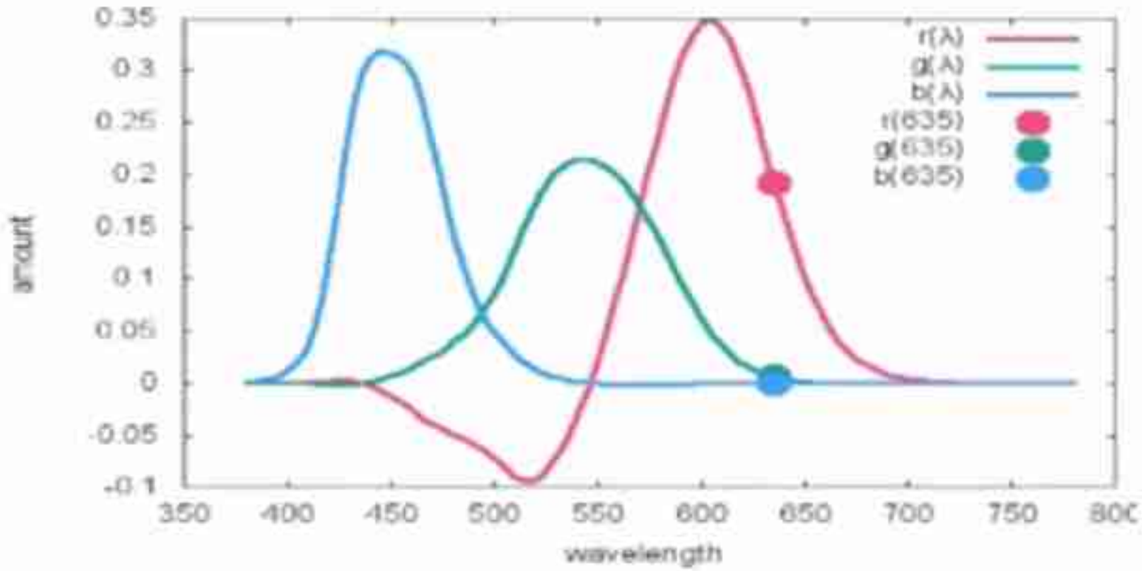

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



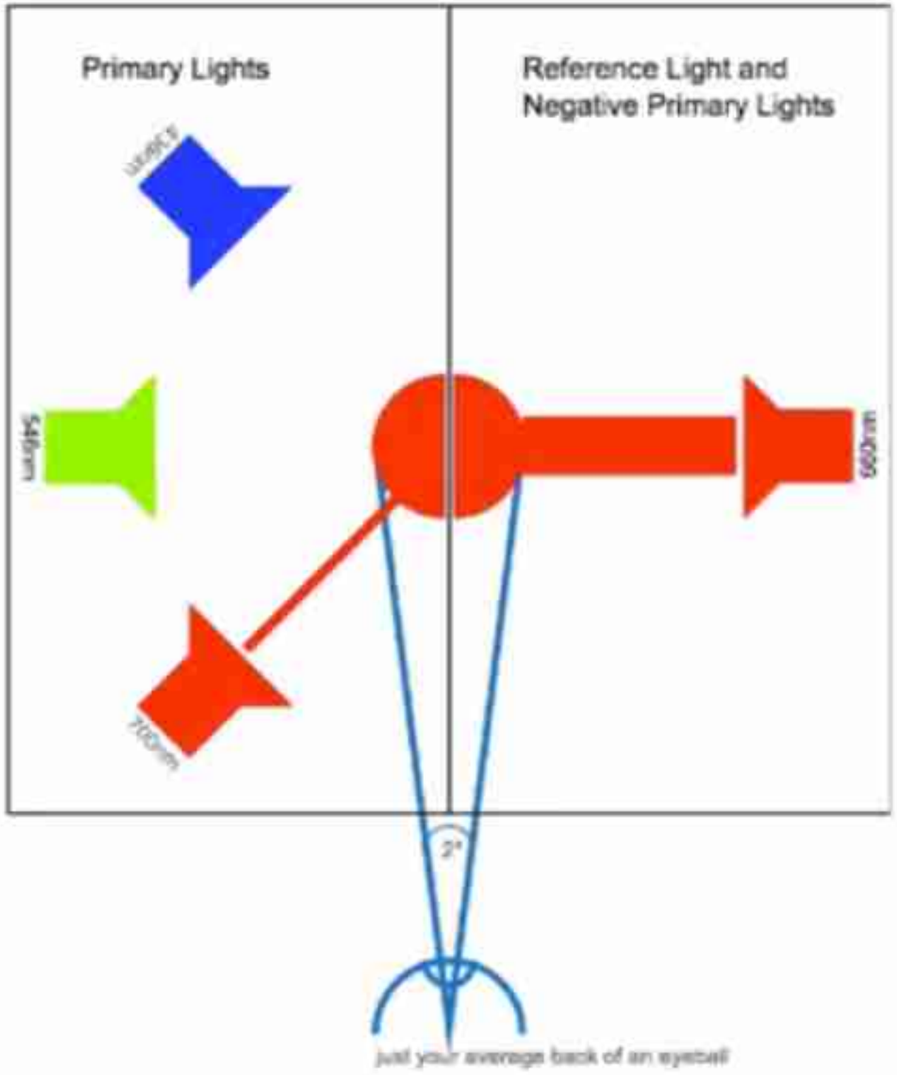



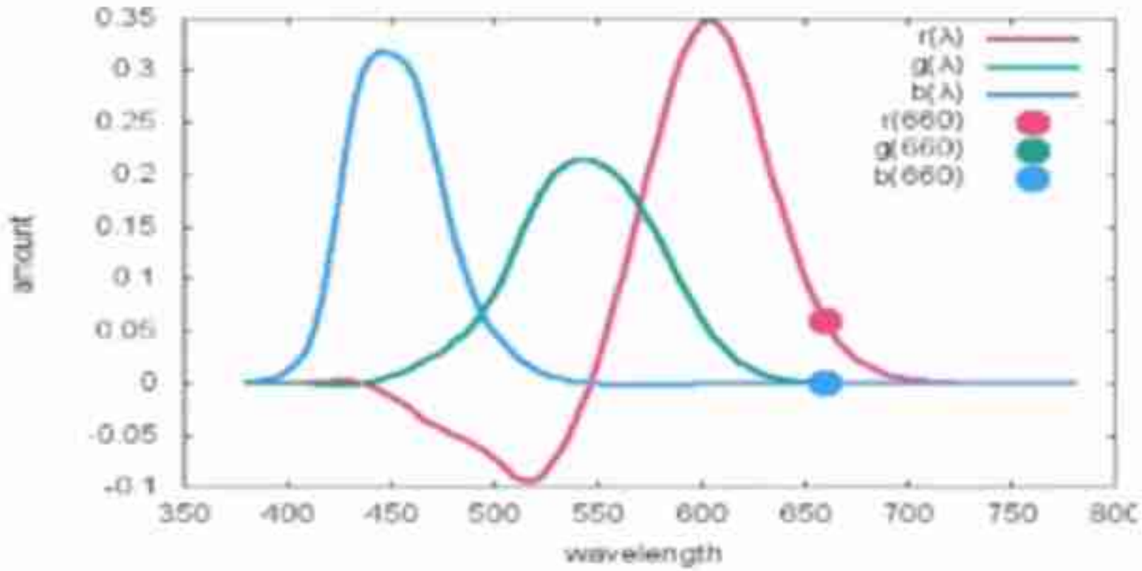

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



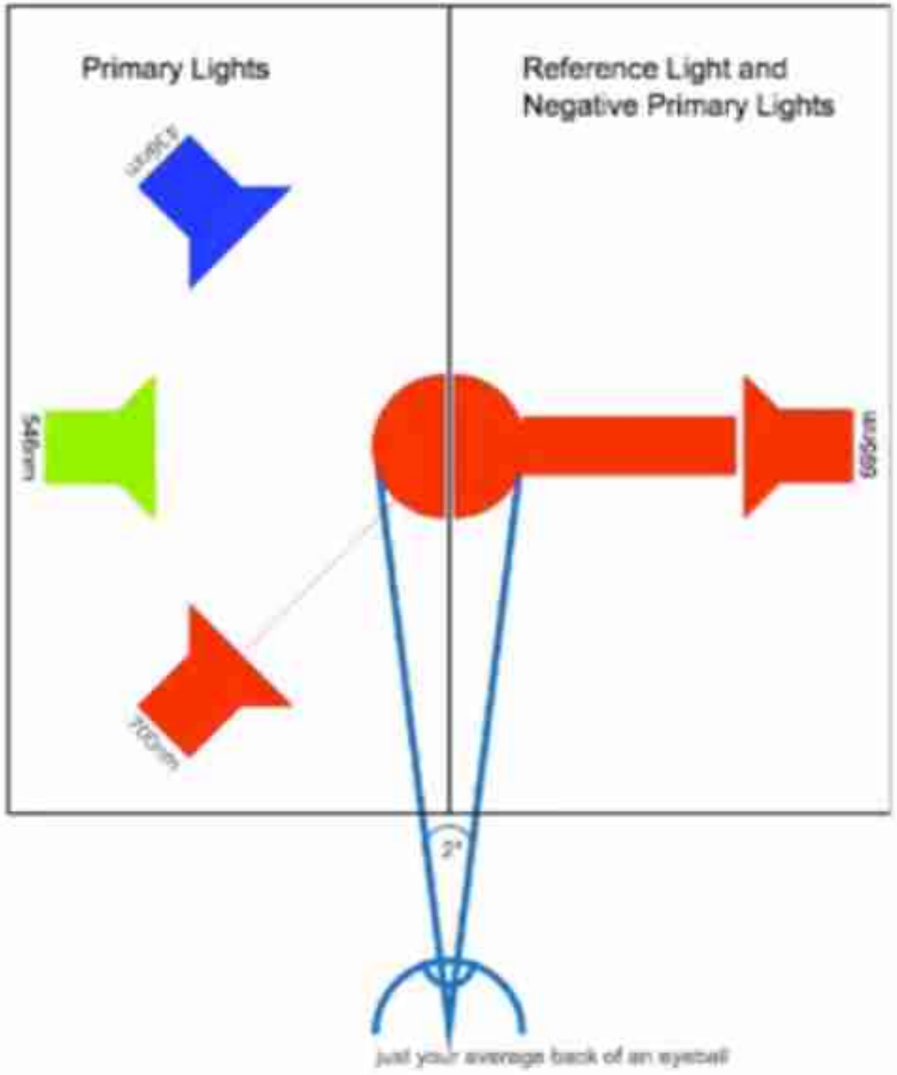



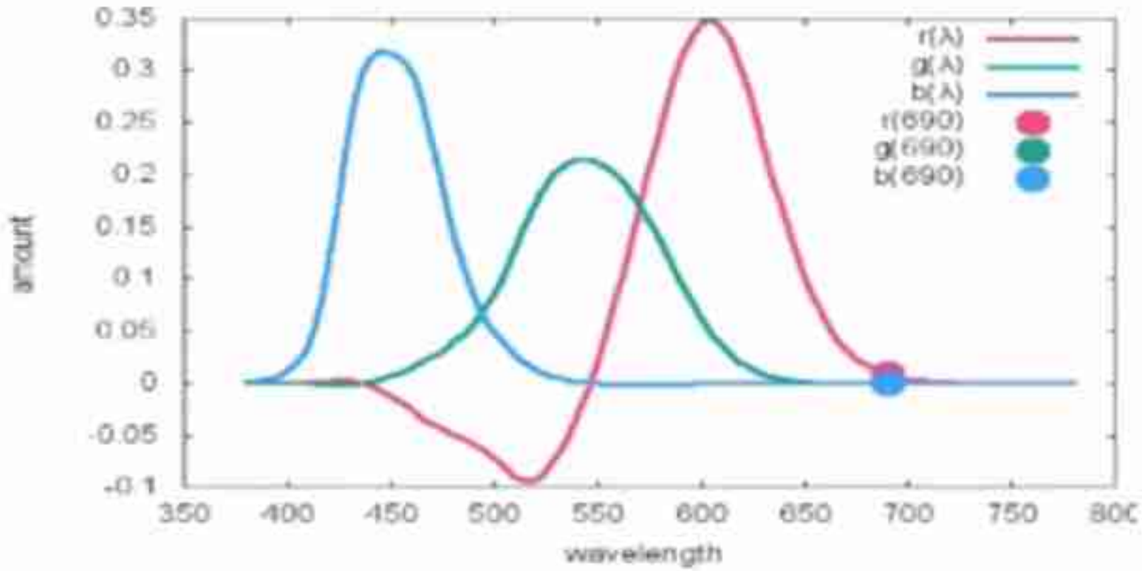

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



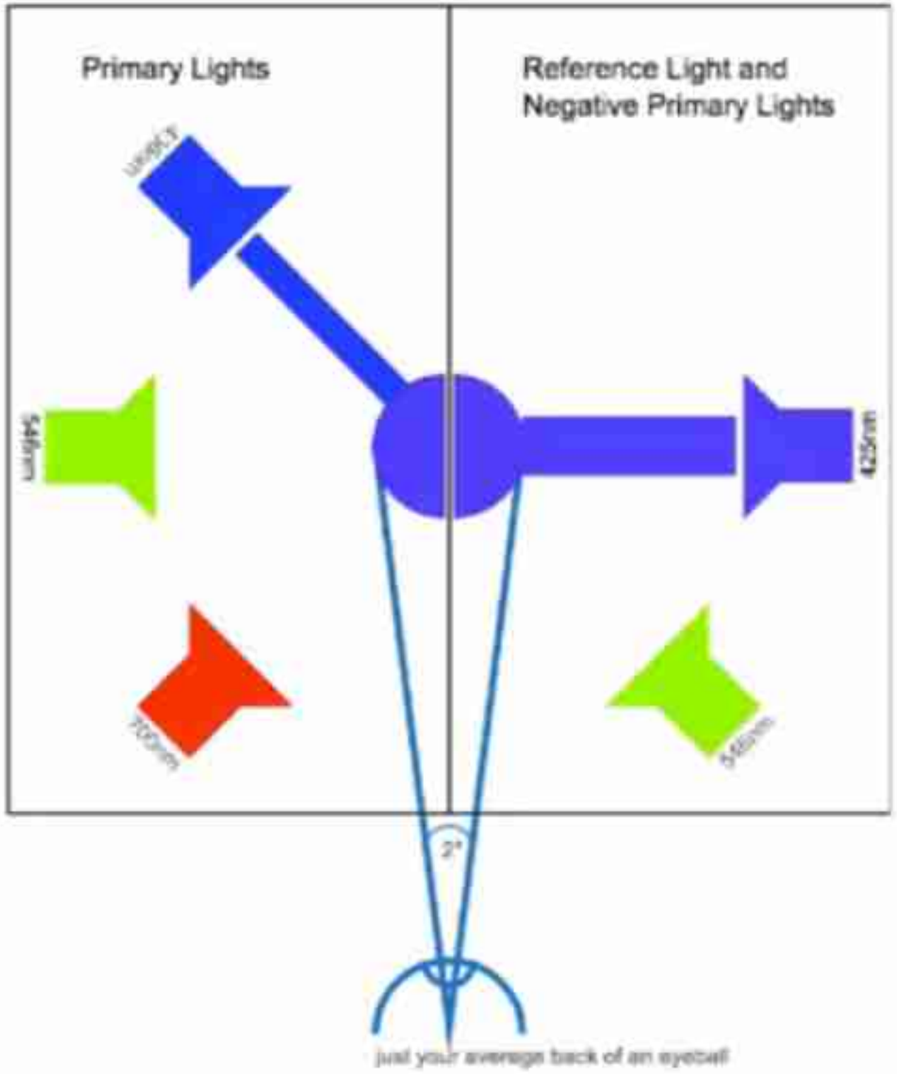



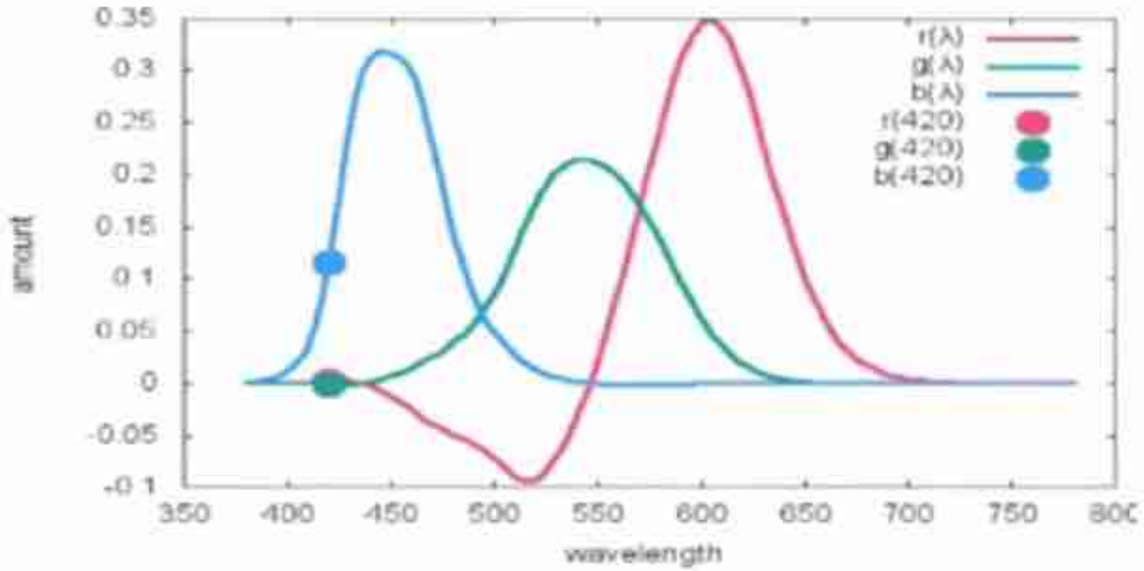

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry



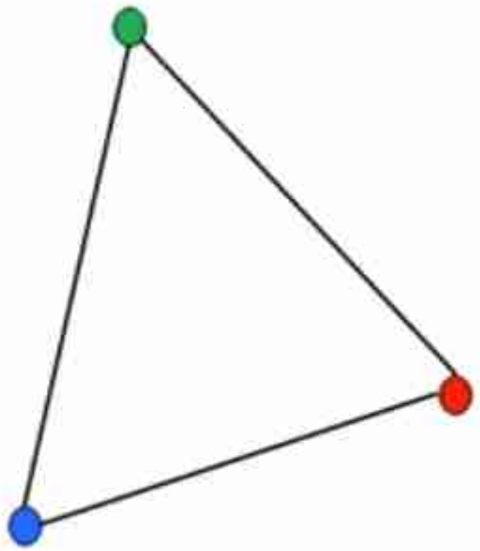



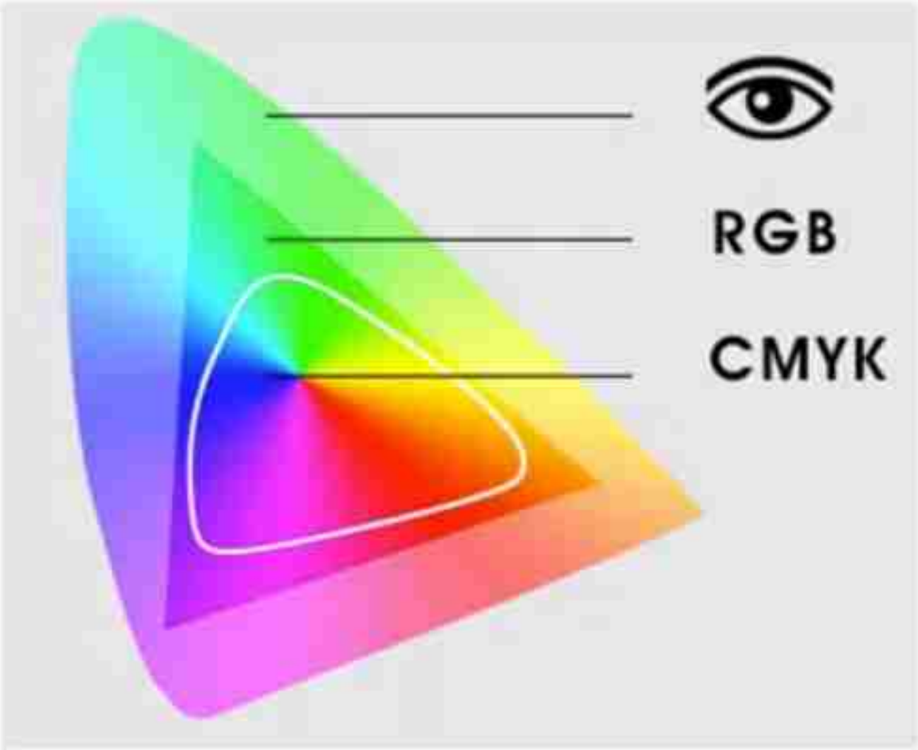

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry

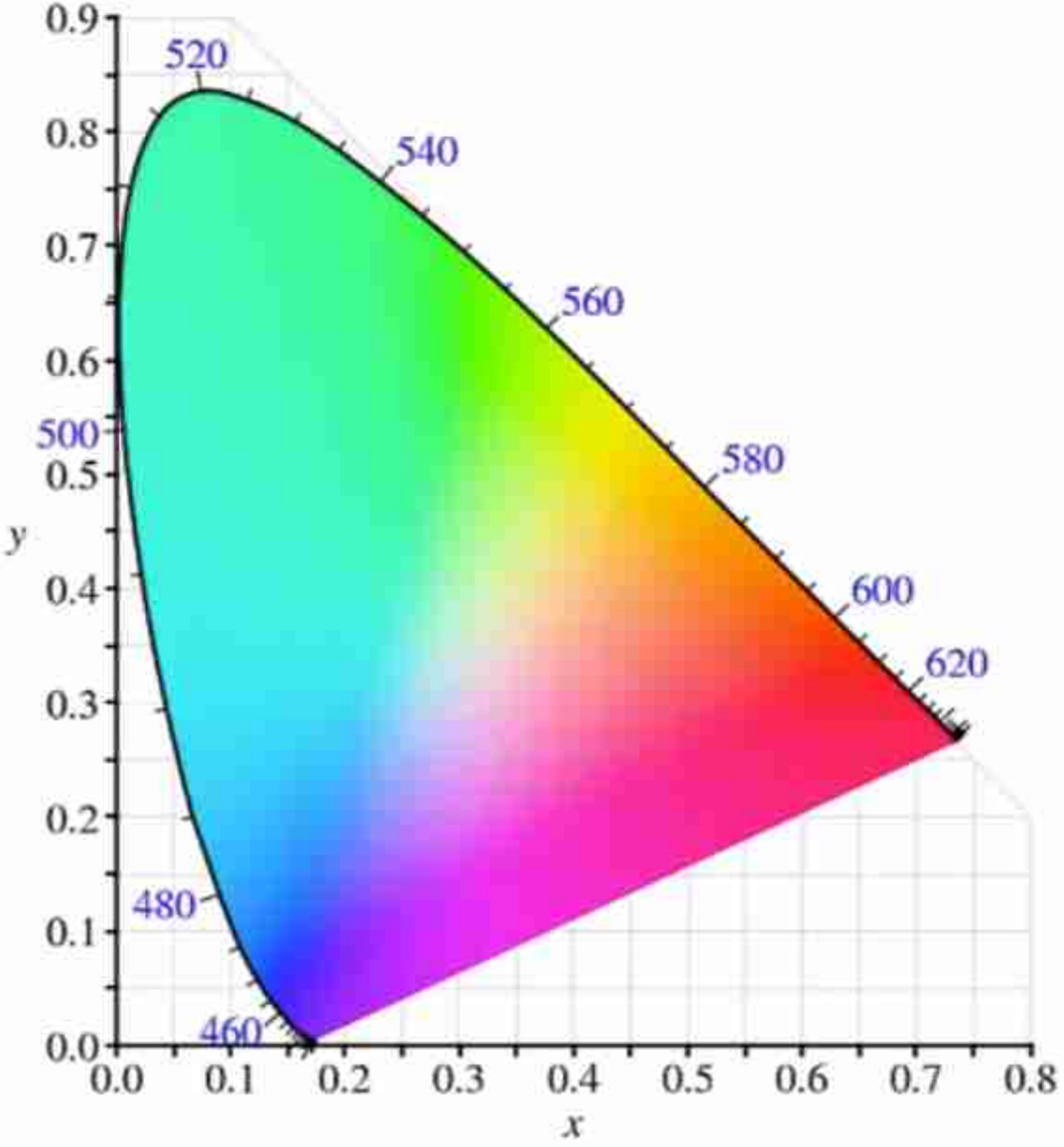


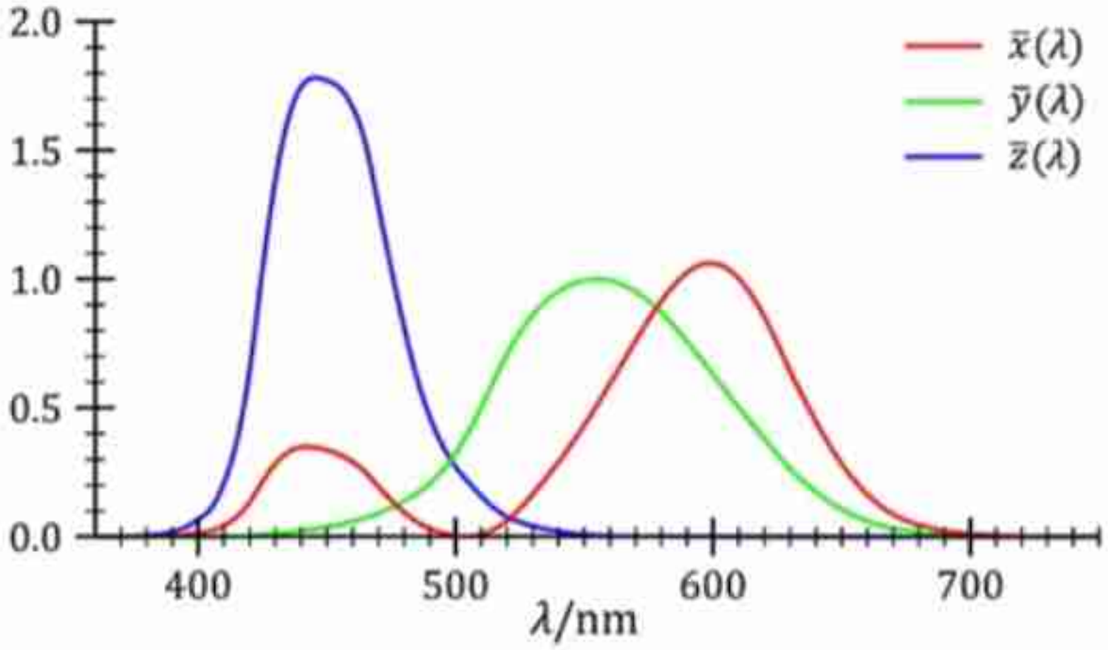

Credit: Chandler Abraham, A Beginner's Guide to (CIE) Colorimetry

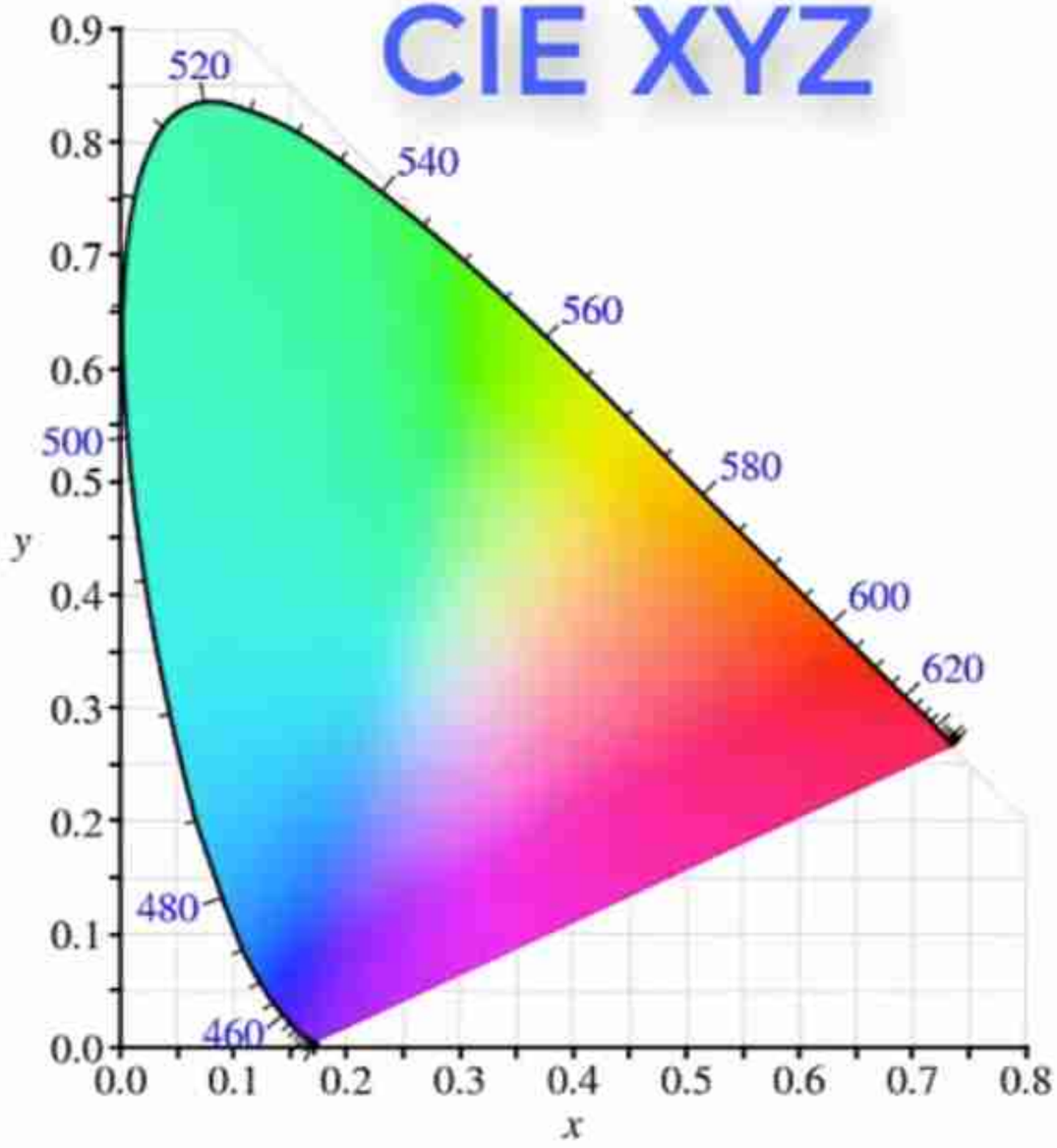



RGB
CMYK

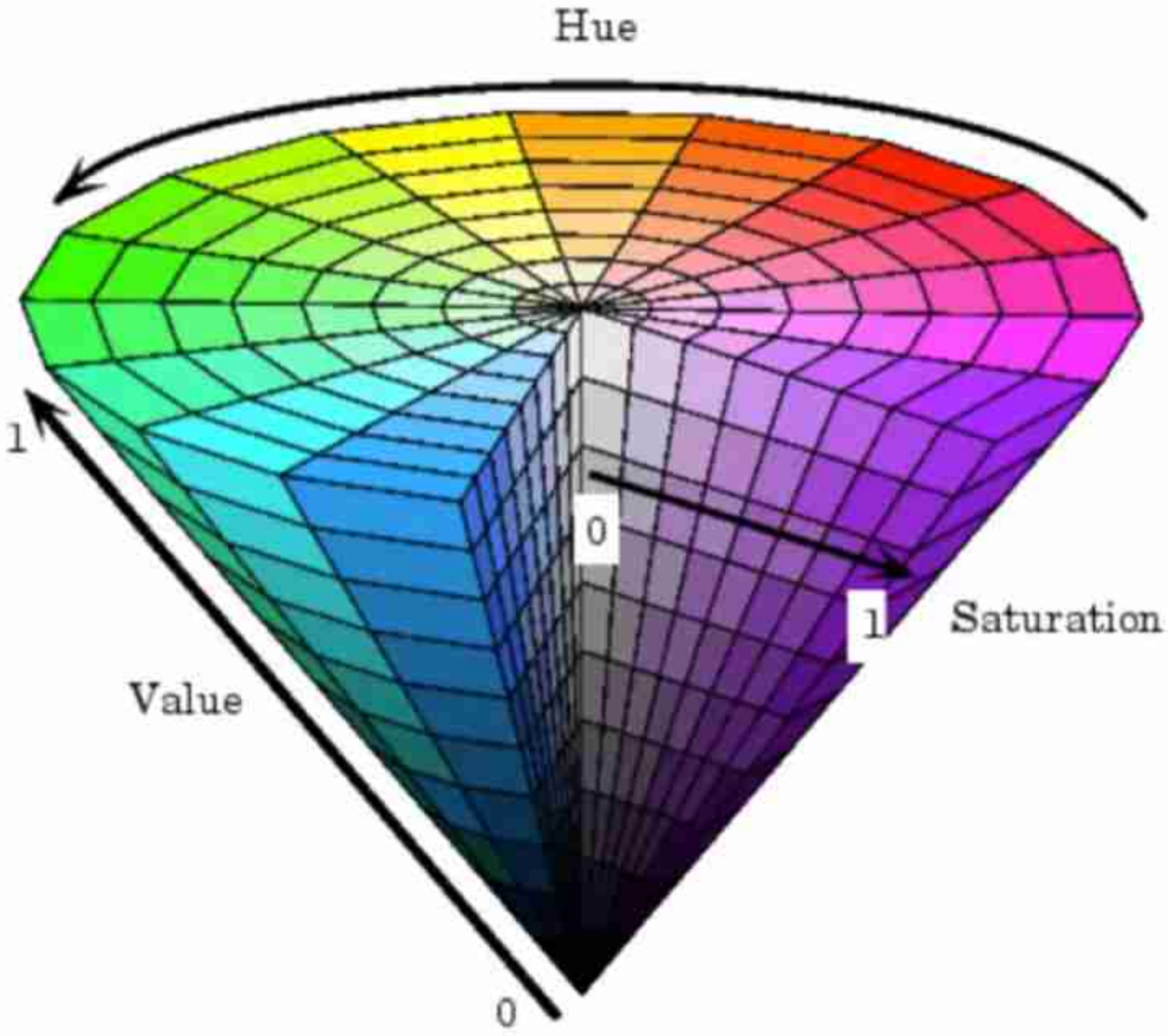

520
540
560
580
600
620
500
480
460
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
y
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
x

# CIE XYZ

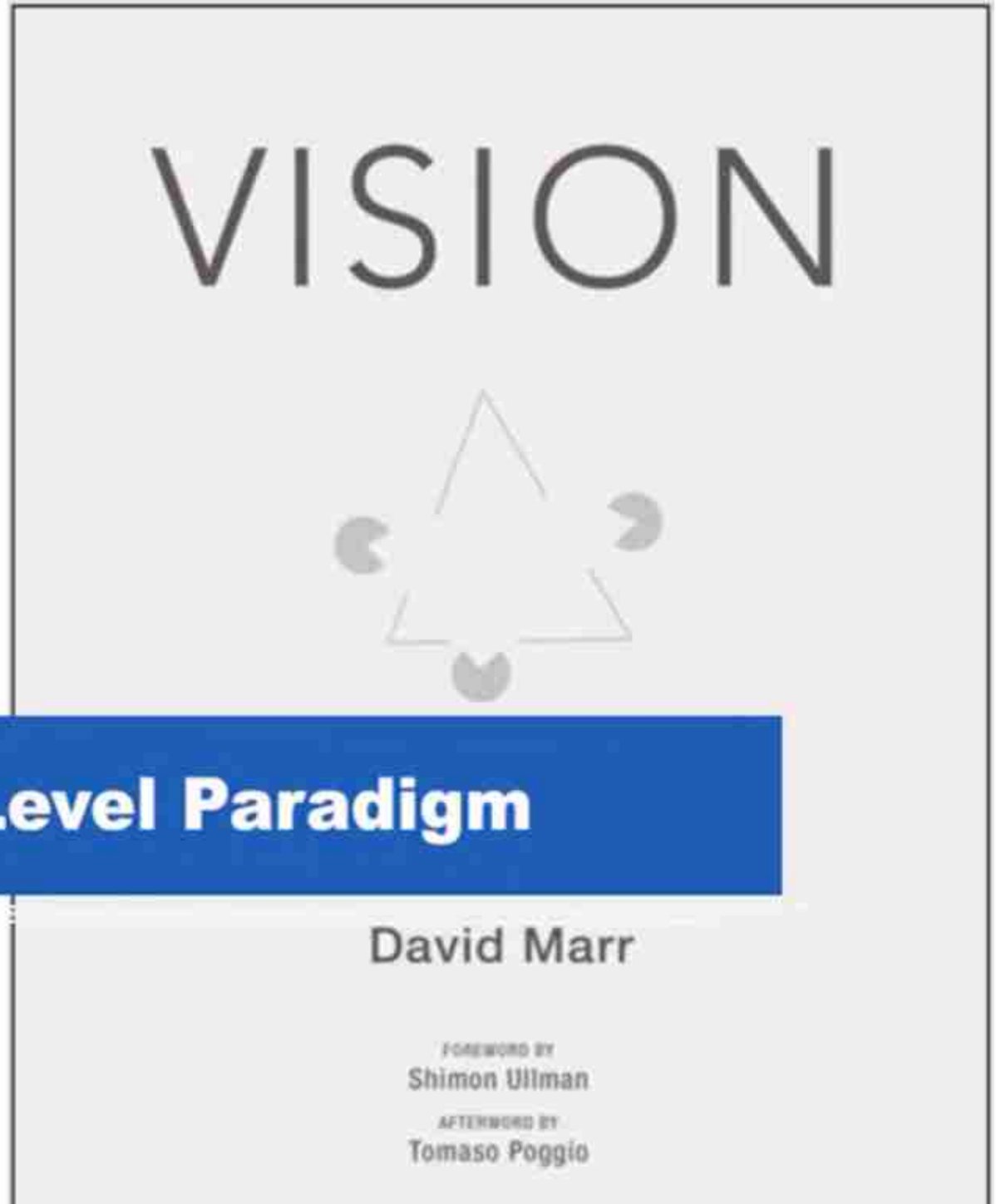

2.0
1.5
1.0
0.5
0.0
400
500
600
700
λ/nm
x̄(λ)
ȳ(λ)
z̄(λ)


0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
y
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
x
520
540
560
580
600
620
500
480
460

# HSV



# Three-Level Paradigm



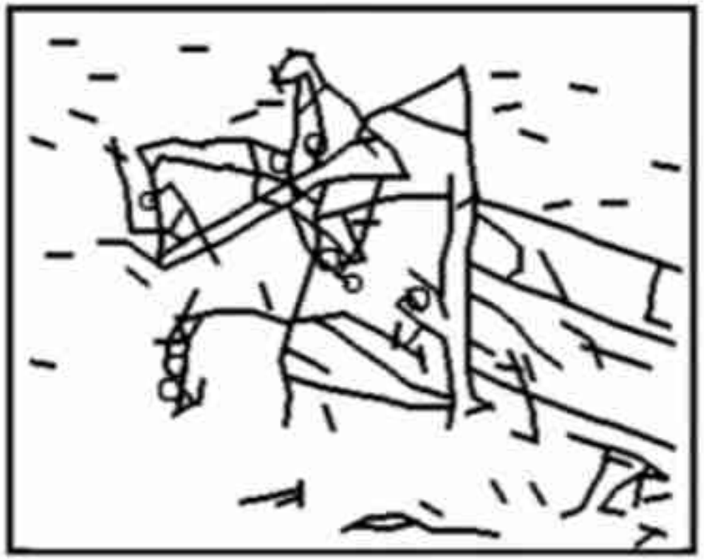
# Three Levels - Paradigm

computational theory

representation and algorithm

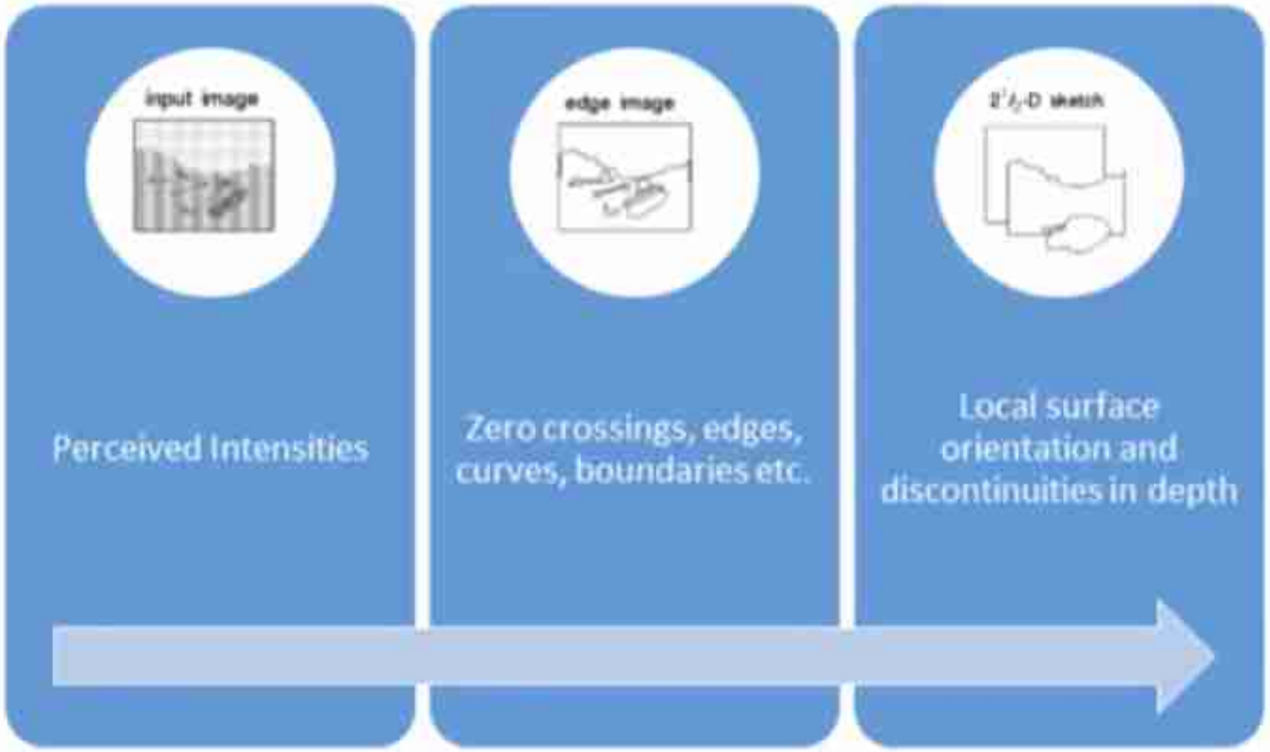
implementation

# Primal Sketch

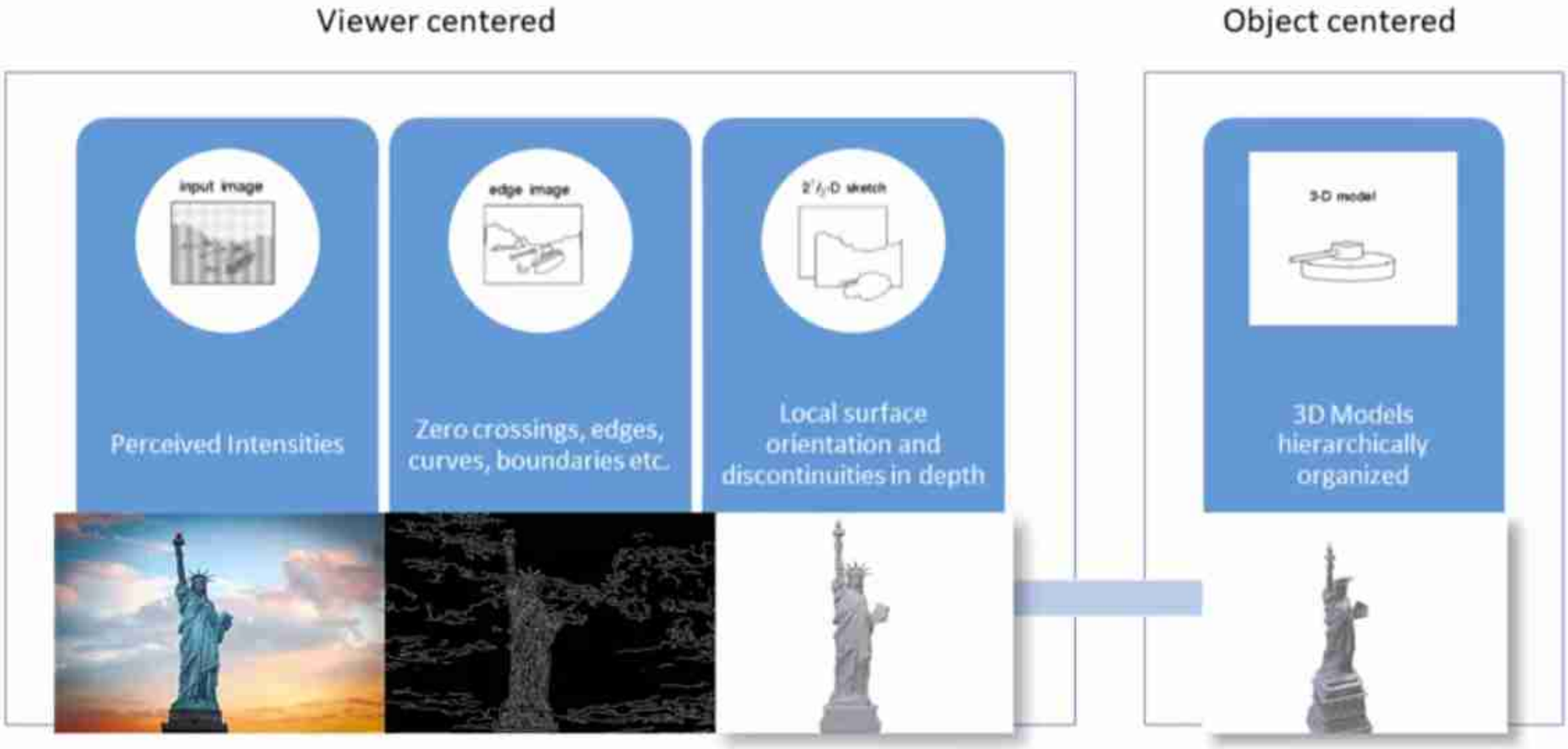


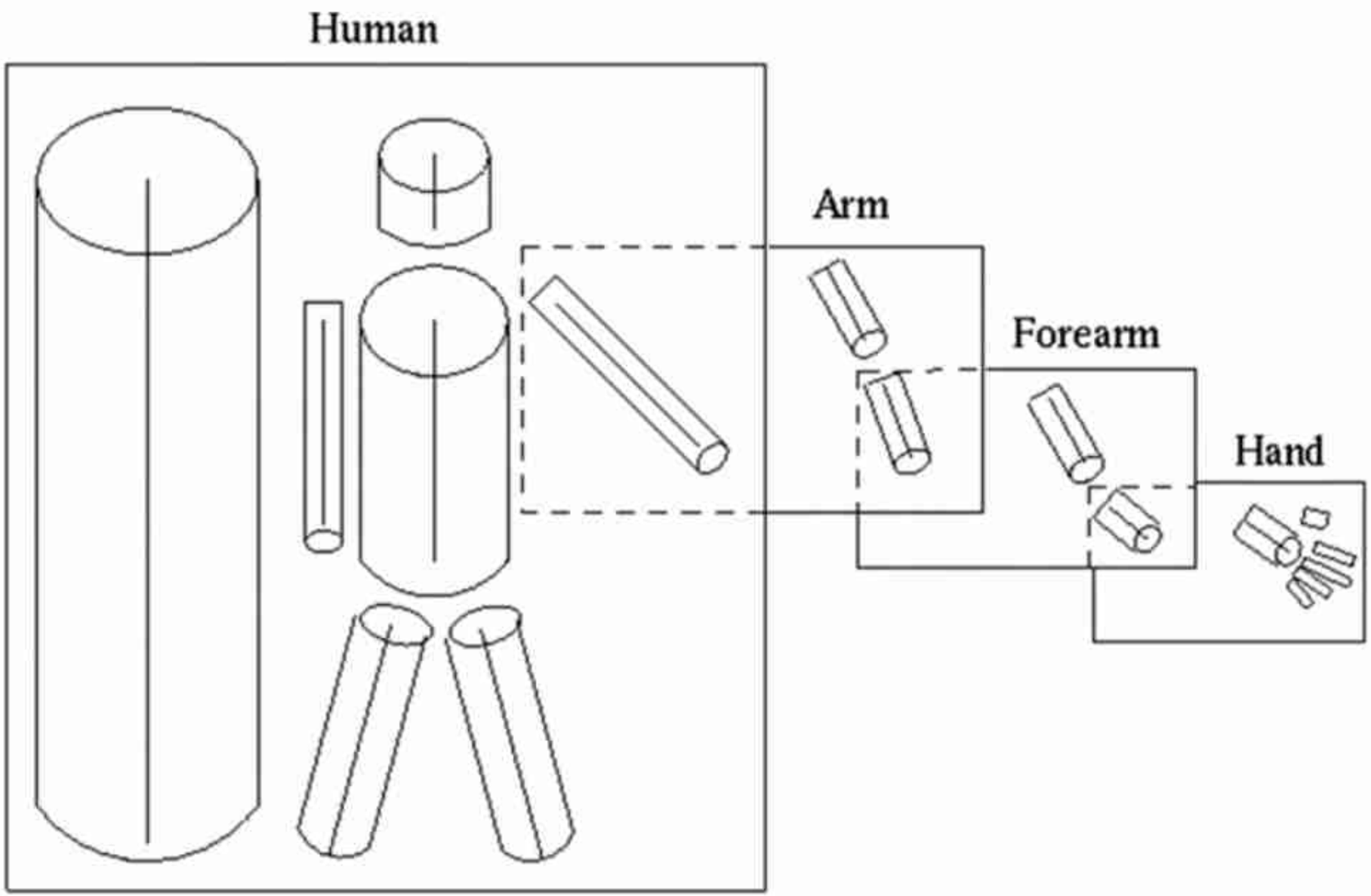
Source: C. Guo, S.C. Zhu, Y.N. Wu, A Mathematical Theory of Primal Sketch and Sketchability

# From Primal Sketch to 3D Models

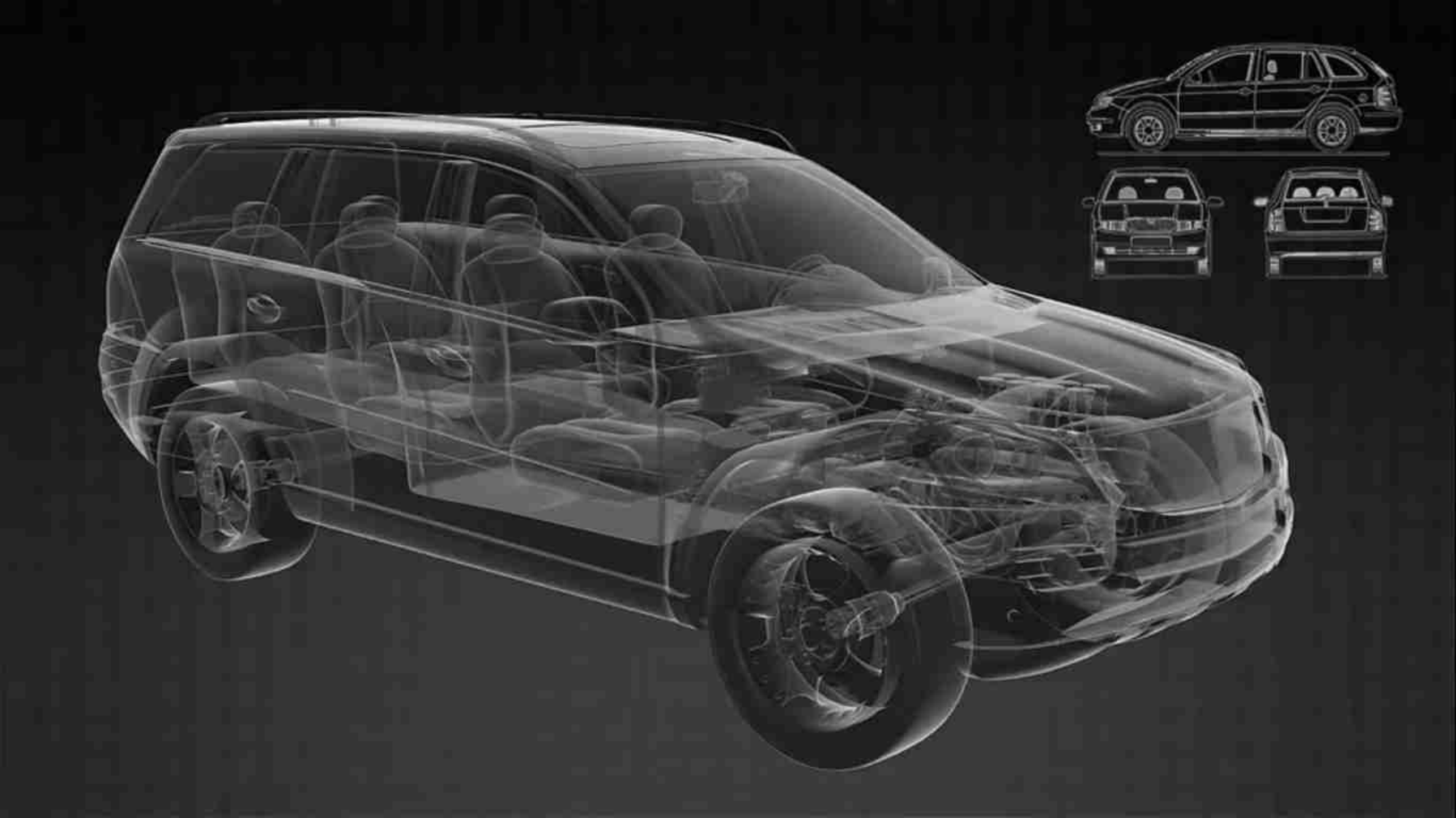


# From Primal Sketch to 3D Models






Source: Vision by David Marr

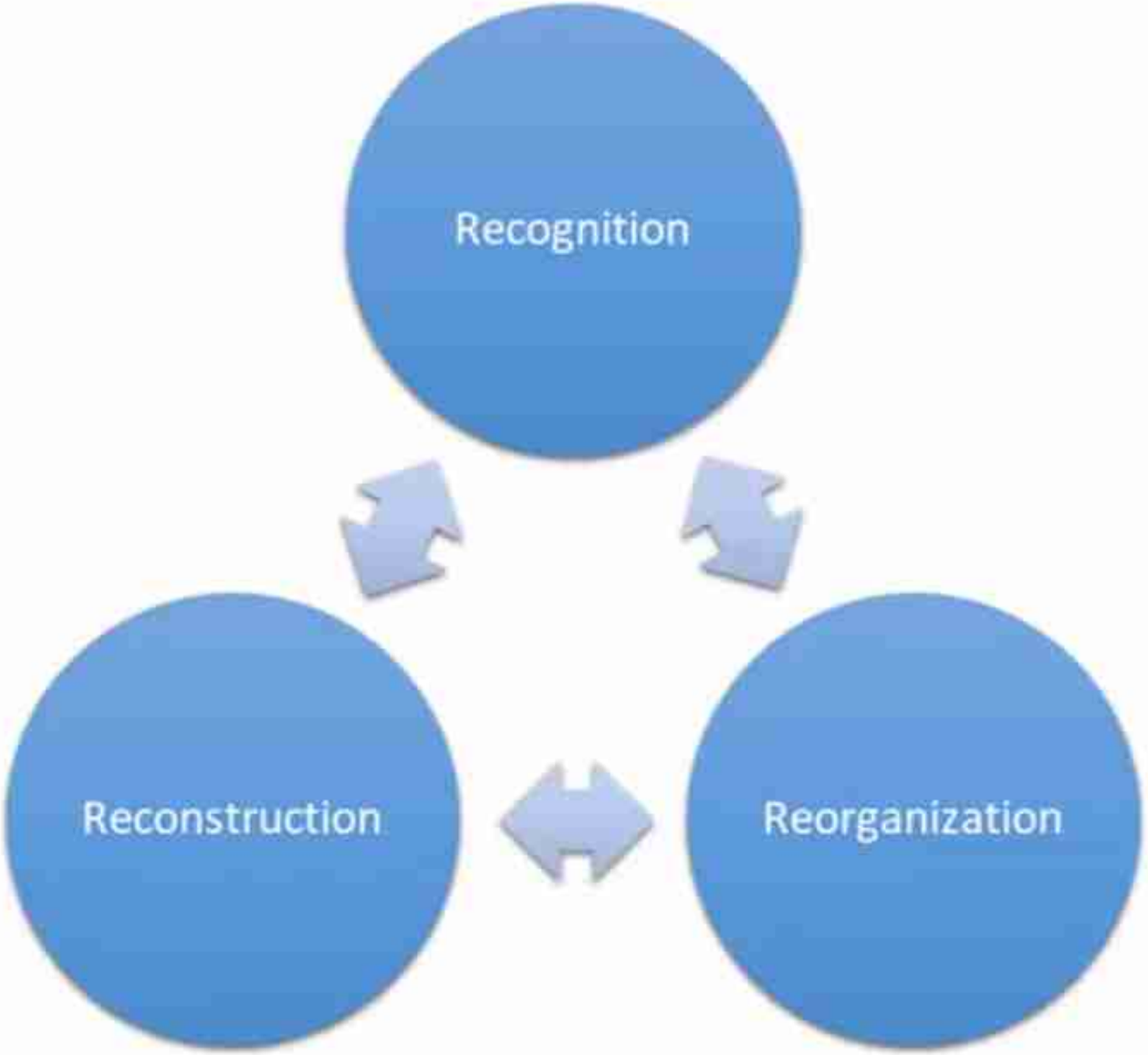


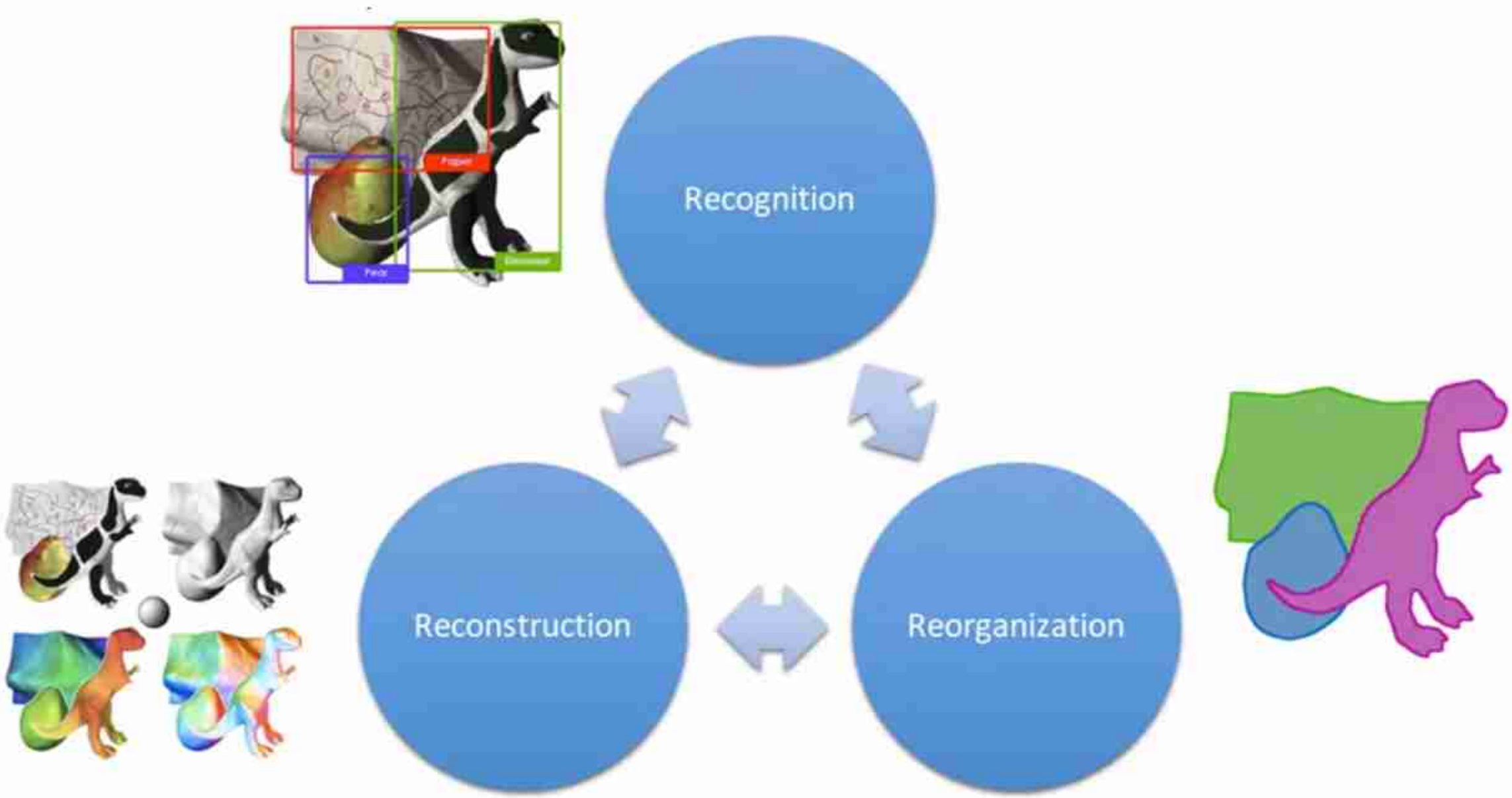
Audi

GDP
BANK
demand
teamwork
tax
NO STOCK

# Three R's - Paradigm

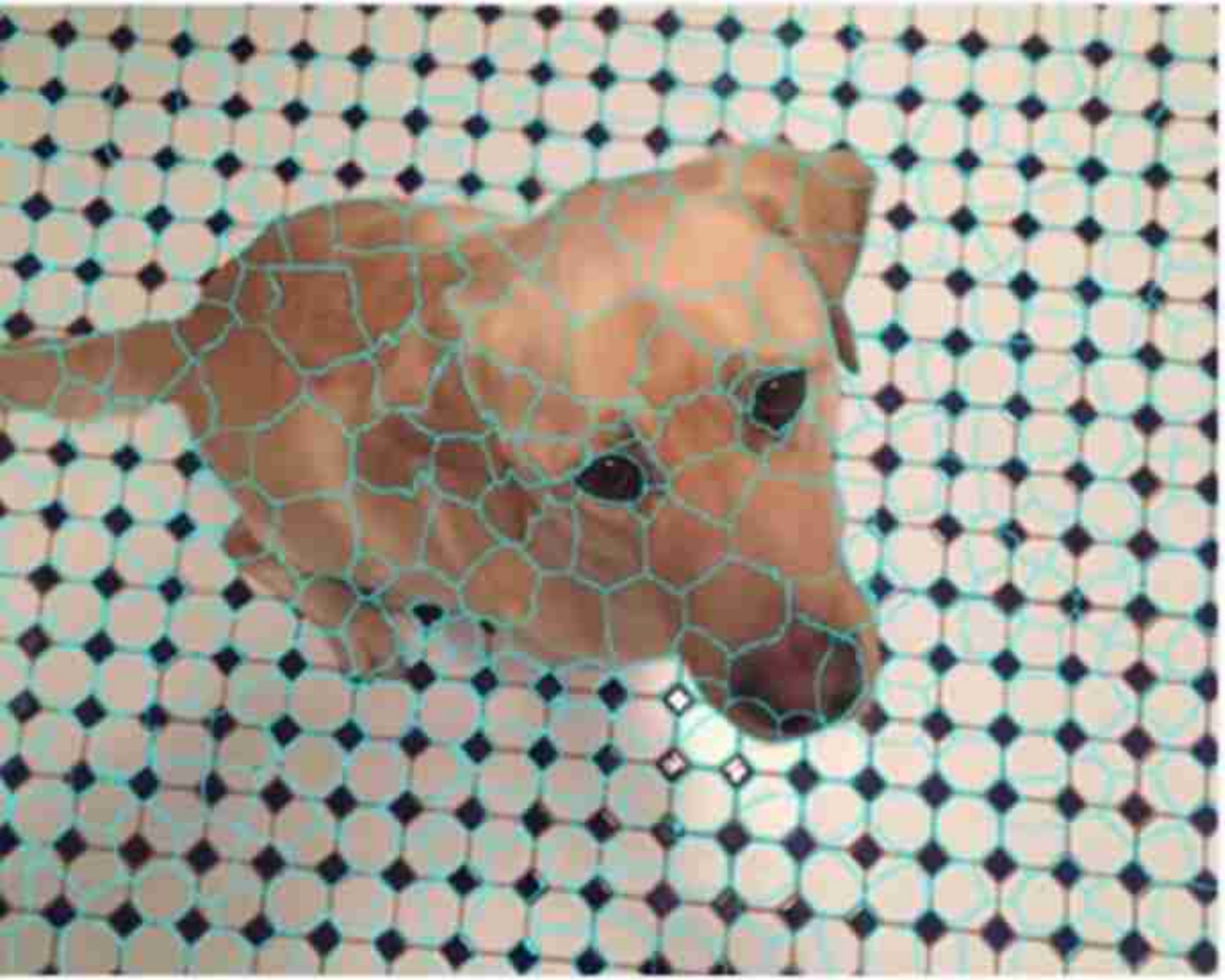
recognition
reconstruction
reorganization


Recognition
Reorganization
Reconstruction




Credit: Dr. Jitendra Malik

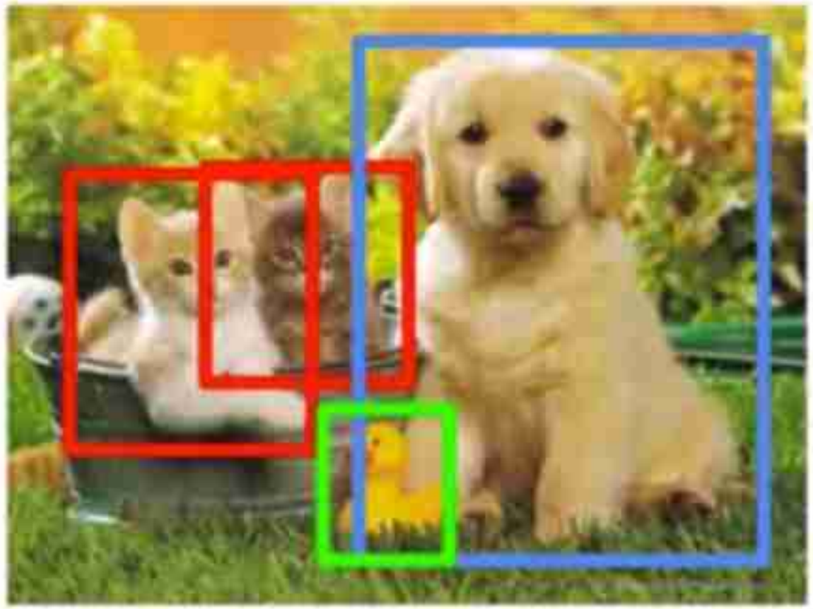
# Reorganization --> Recognition



## Classification + Localization

CAT

## Object Detection



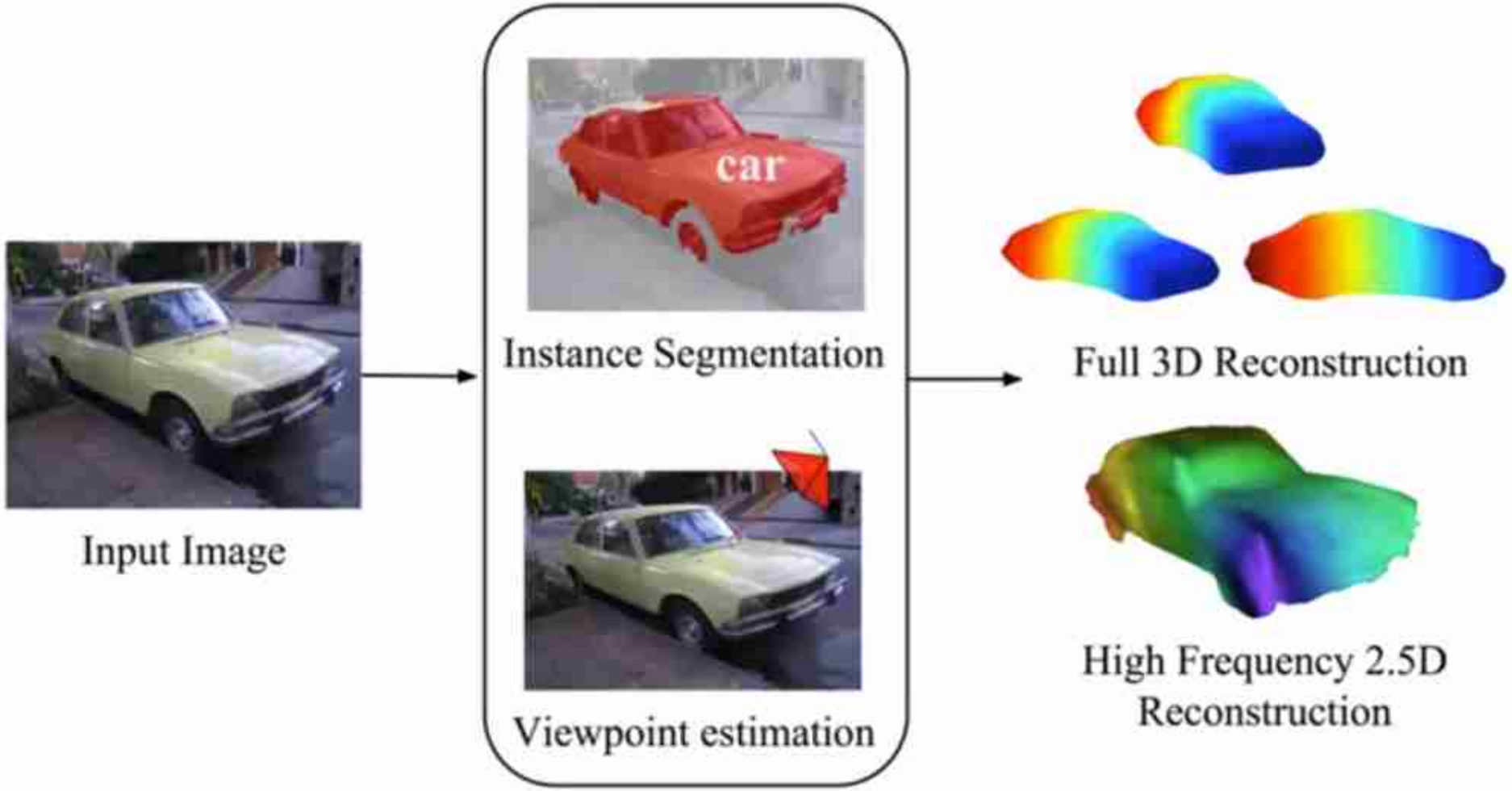
CAT, DOG, DUCK

## Instance Segmentation

CAT, DOG, DUCK

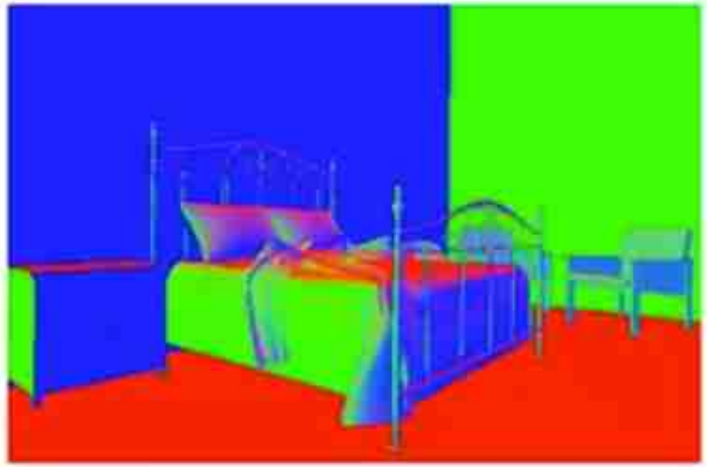
e object

Multiple objects

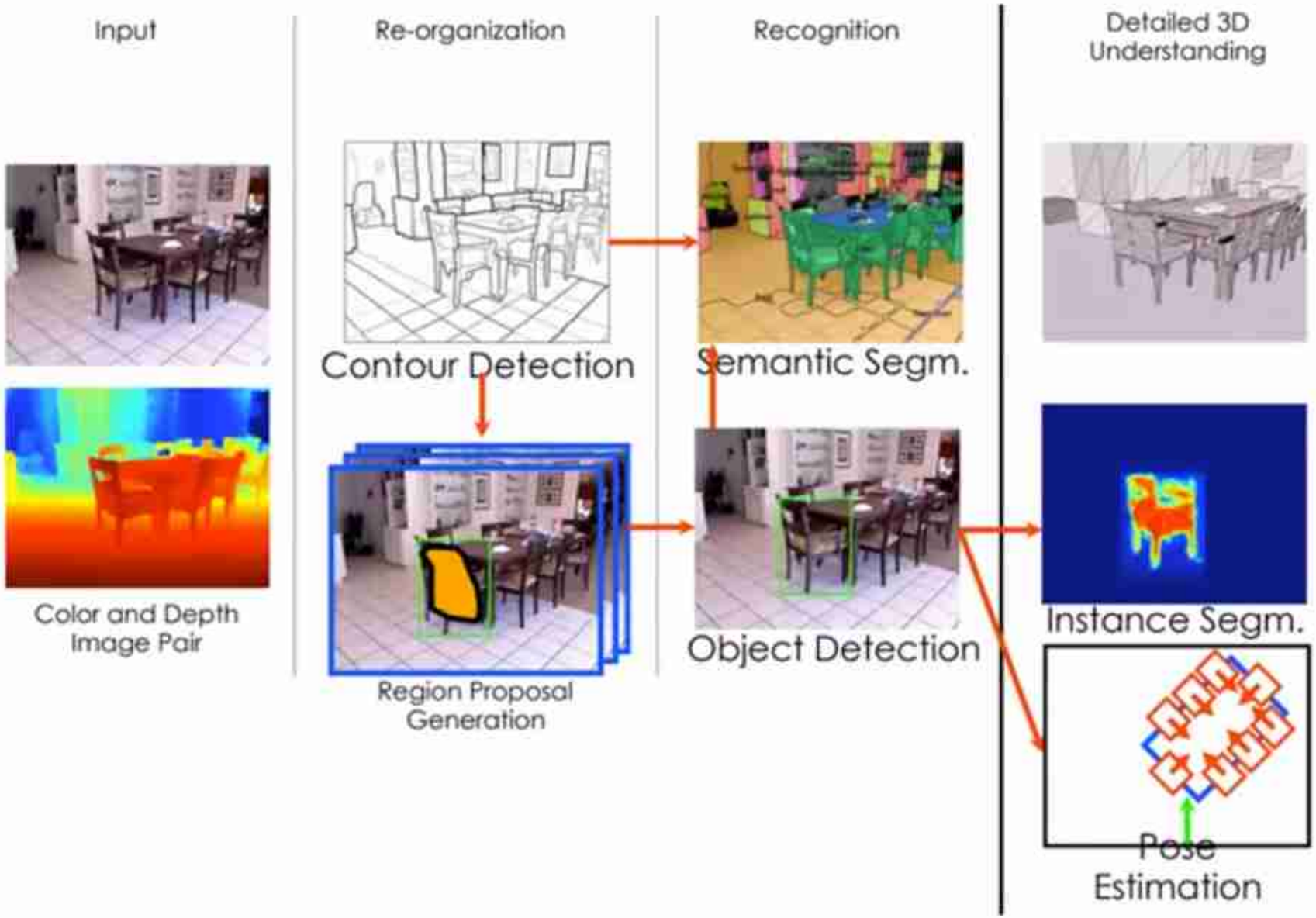



Source: Category Specific Object Reconstruction, Kar, Tulsiani, Carreira & Malik

Color and Depth Image Pair



3D Image Understanding

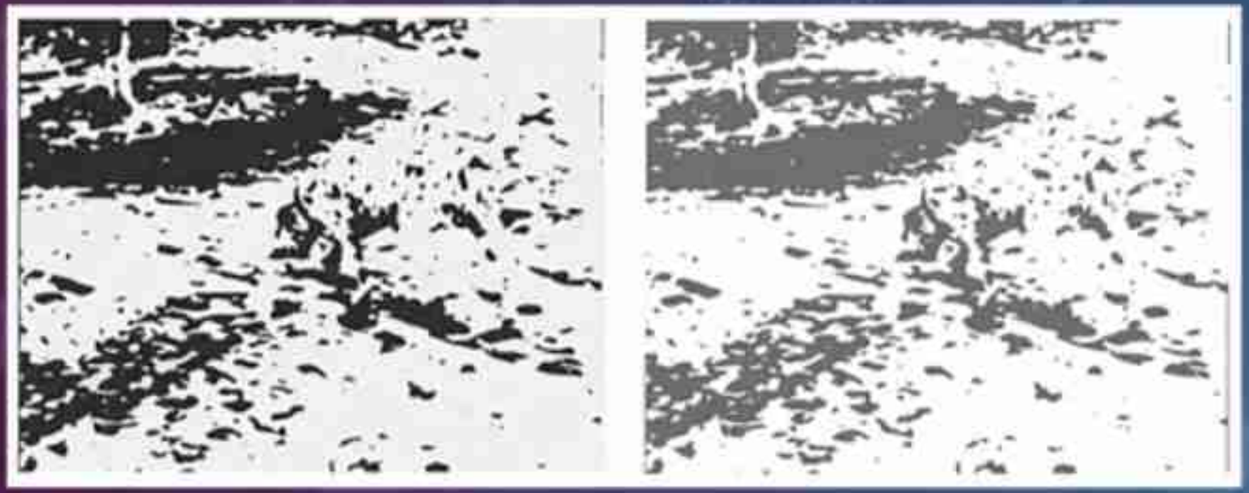

Input
Re-organization
Recognition
Detailed 3D Understanding
Color and Depth Image Pair
Contour Detection
Region Proposal Generation
Semantic Segm.
Object Detection
Instance Segm.
Pose Estimation

# Low Level Vision

image processing techniques, feature detection and matching and early segmentation

# Mid Level Vision

come together attributing meaning



# High Level Vision

algorithms which make sense of the visual content and make computer vision live up to the capabilities of human vision

# Low-Level or Early Vision

Considers local properties of an image

"There are several edges in this image!"

Share
Low Level Vision
0:02 / 8:36

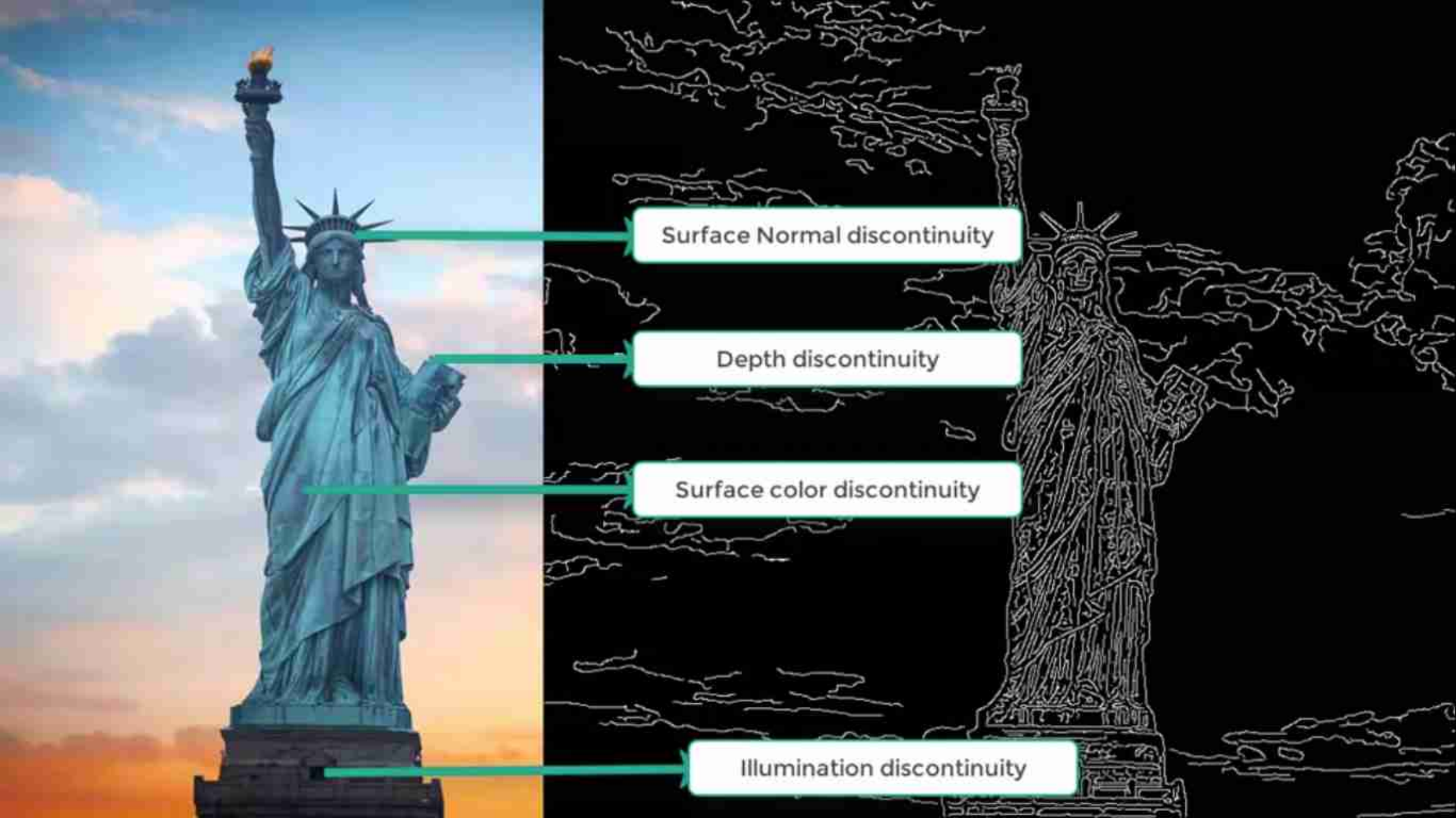
# Low Level Image Processing

edges

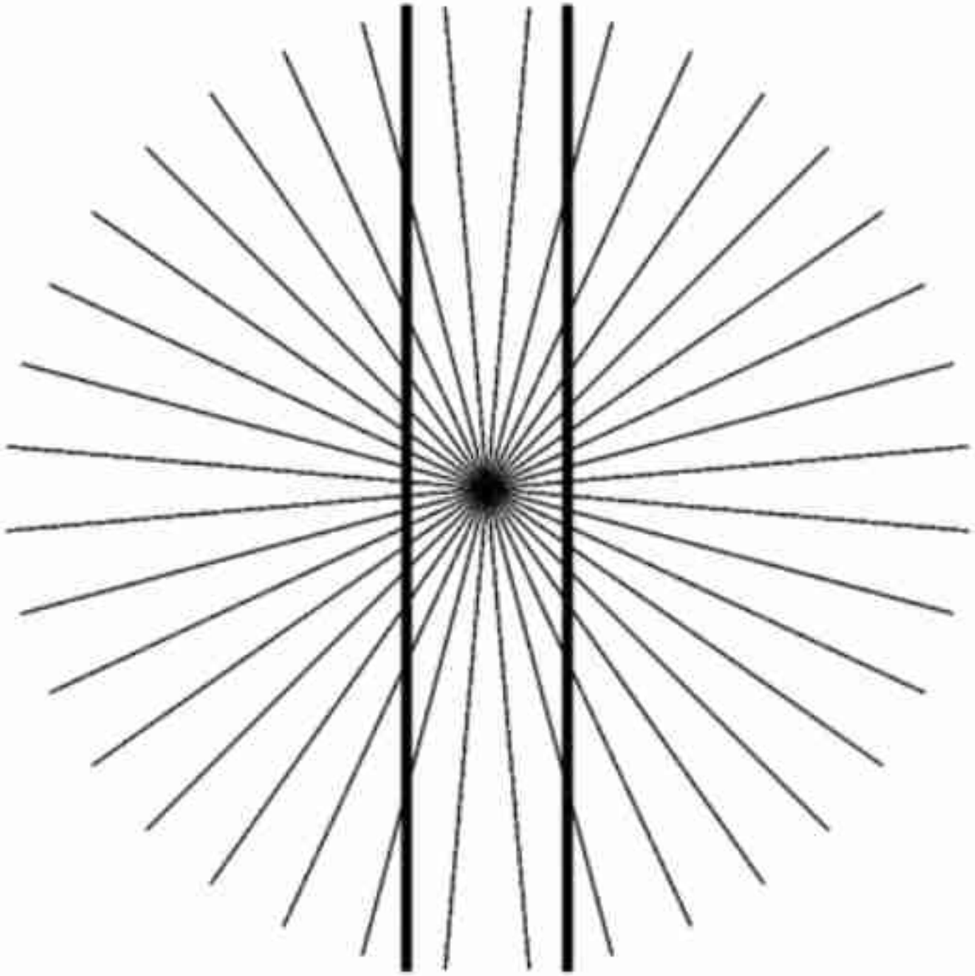
corners

performing filtering

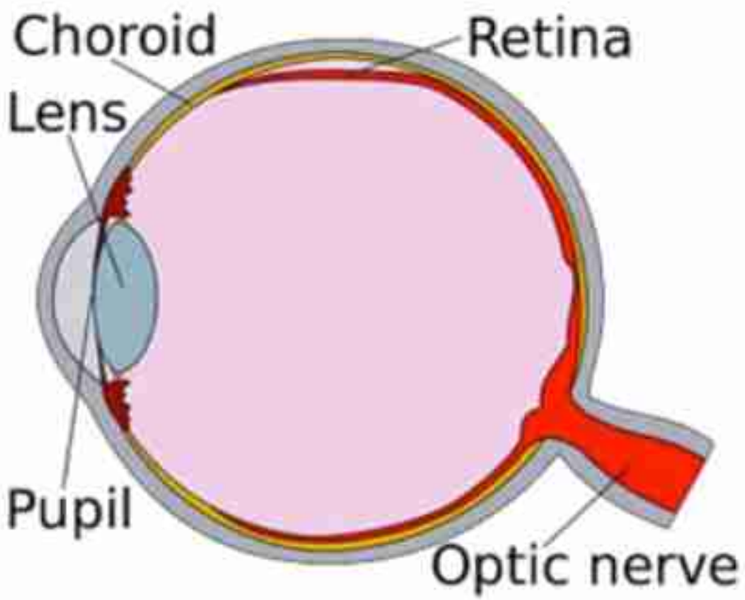
morphology

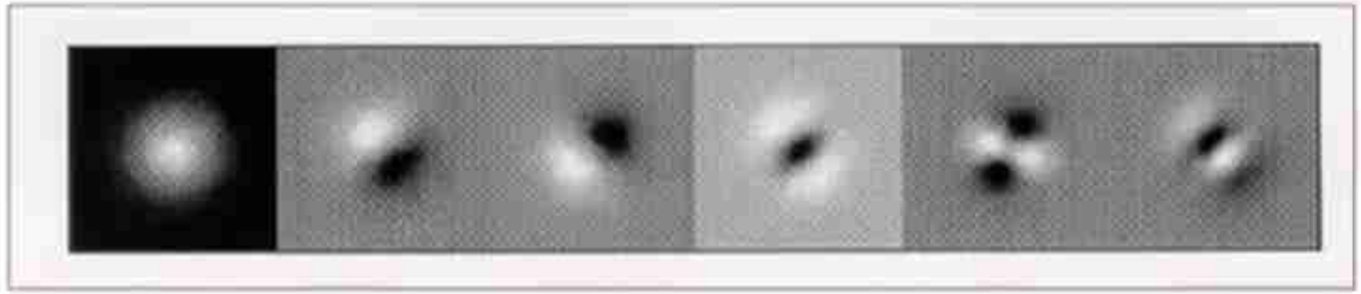

Surface Normal discontinuity
Depth discontinuity
Surface color discontinuity
Illumination discontinuity

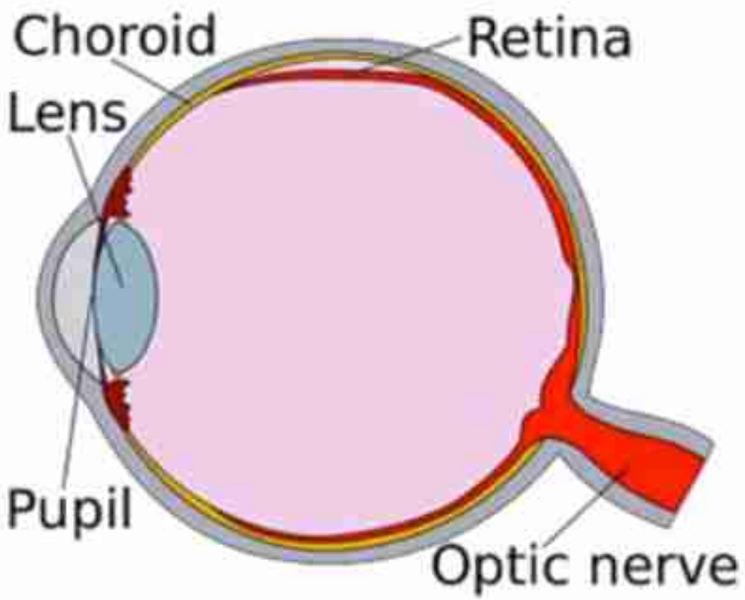
Share

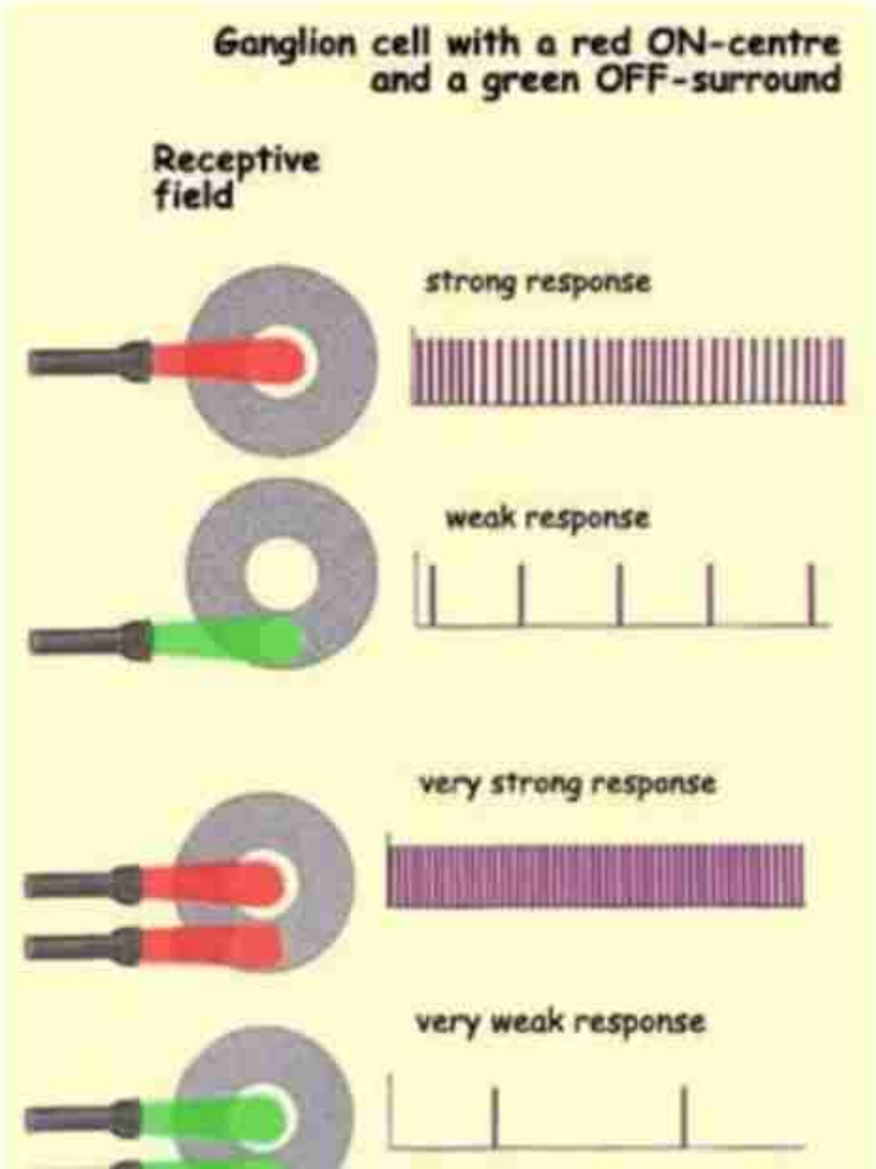


0:53 / 8:36

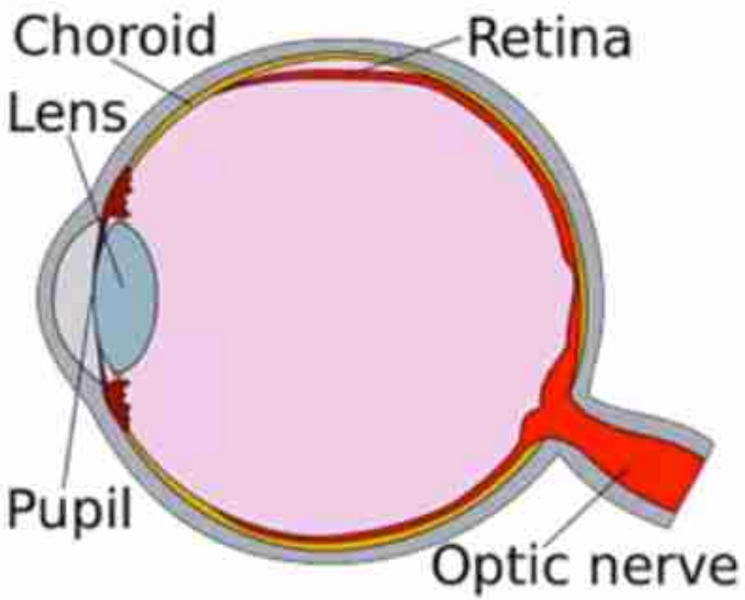

Choroid
Retina
Lens
Pupil
Optic nerve

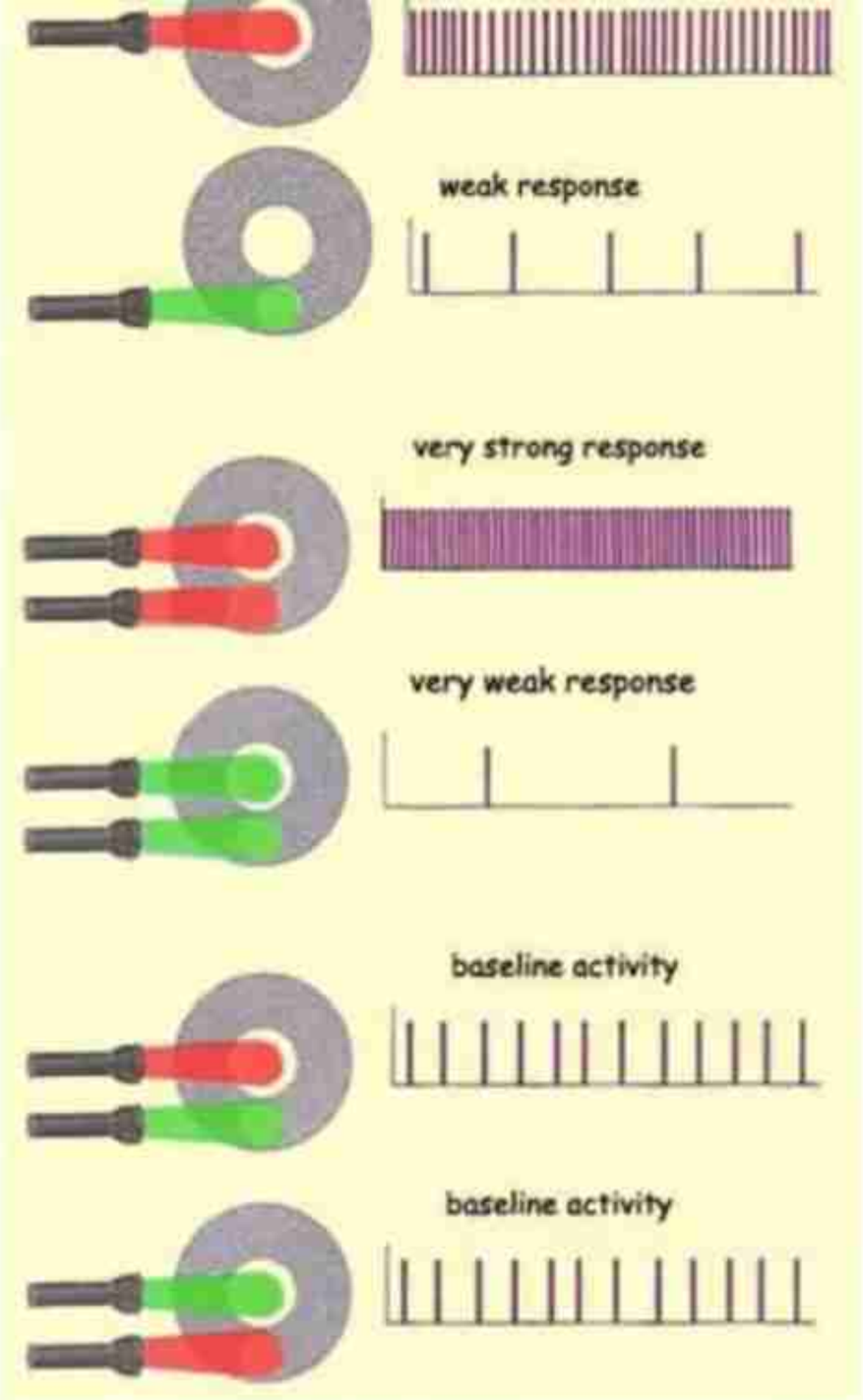


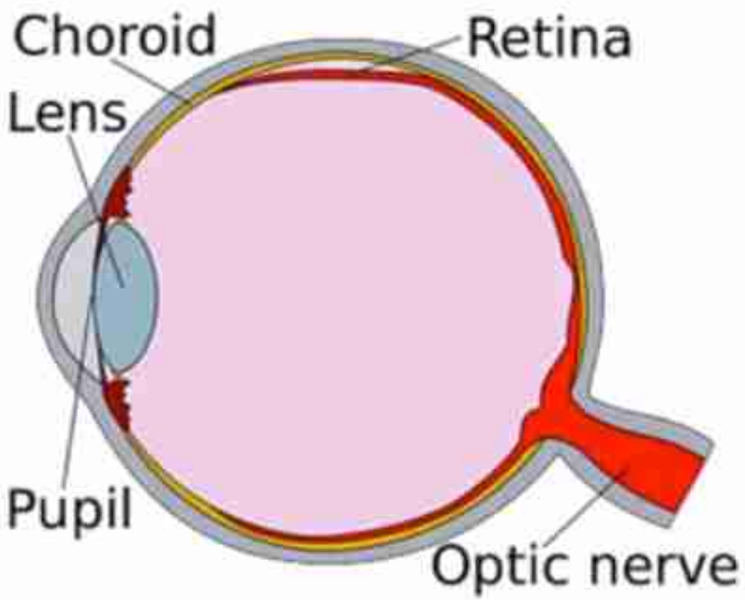


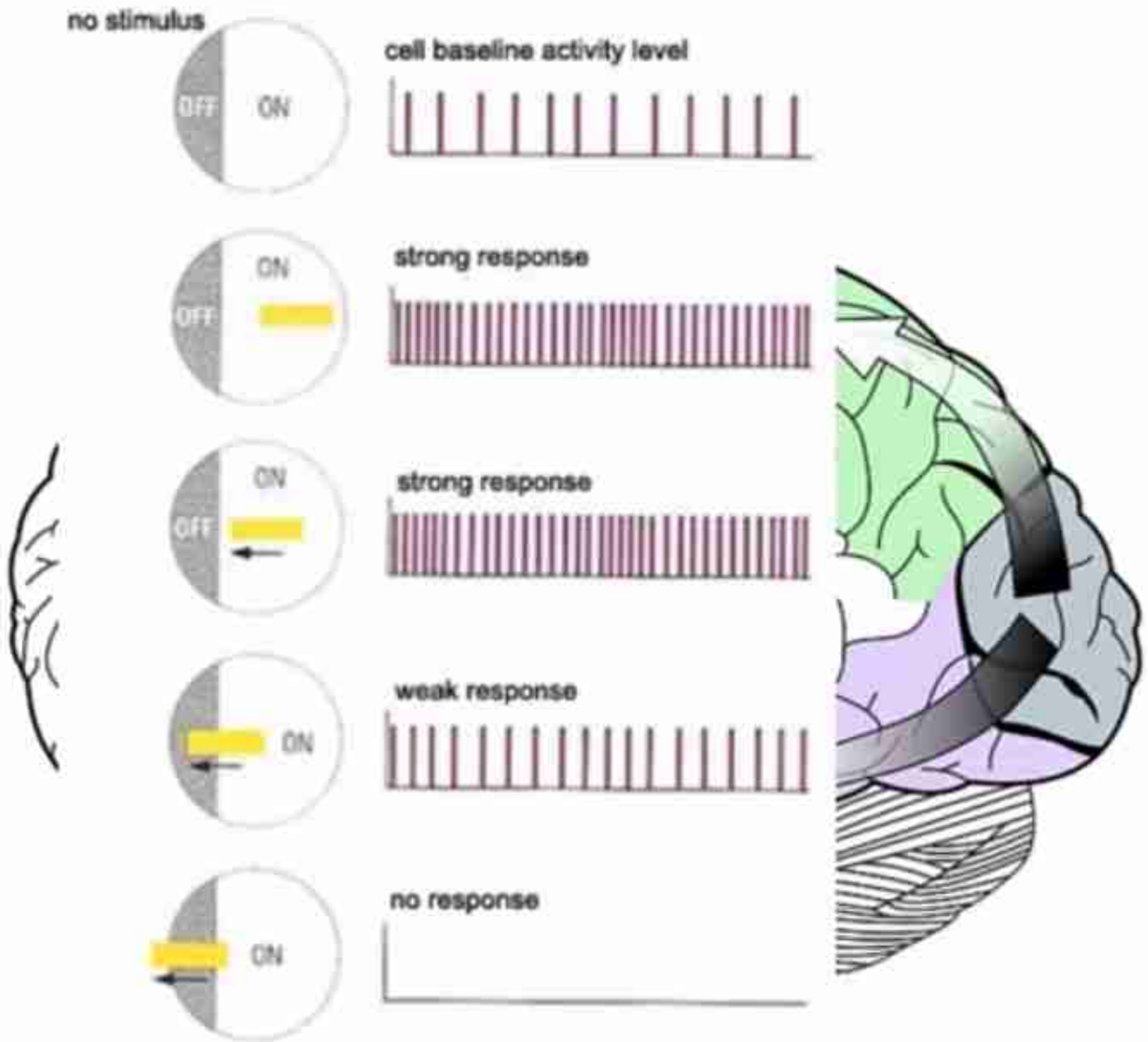



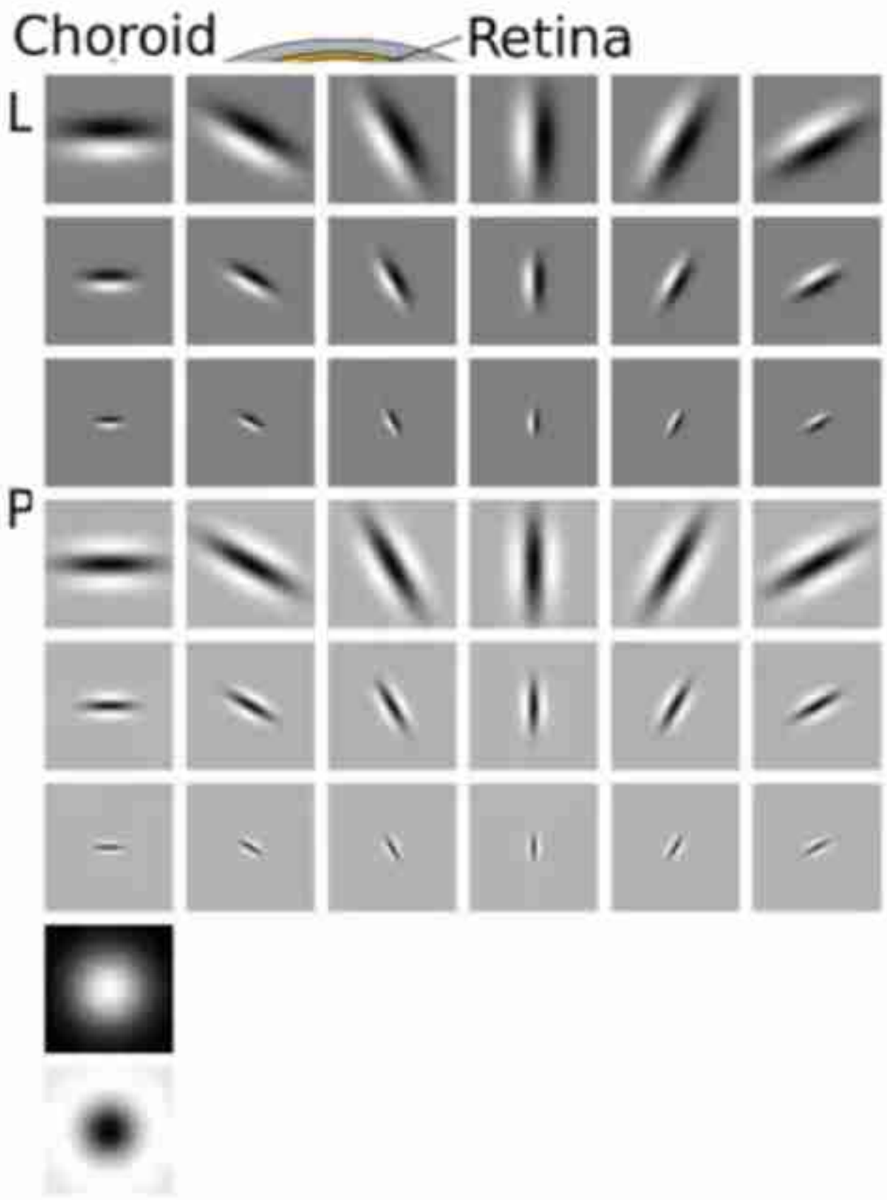

Source: The brain from Top to Bottom

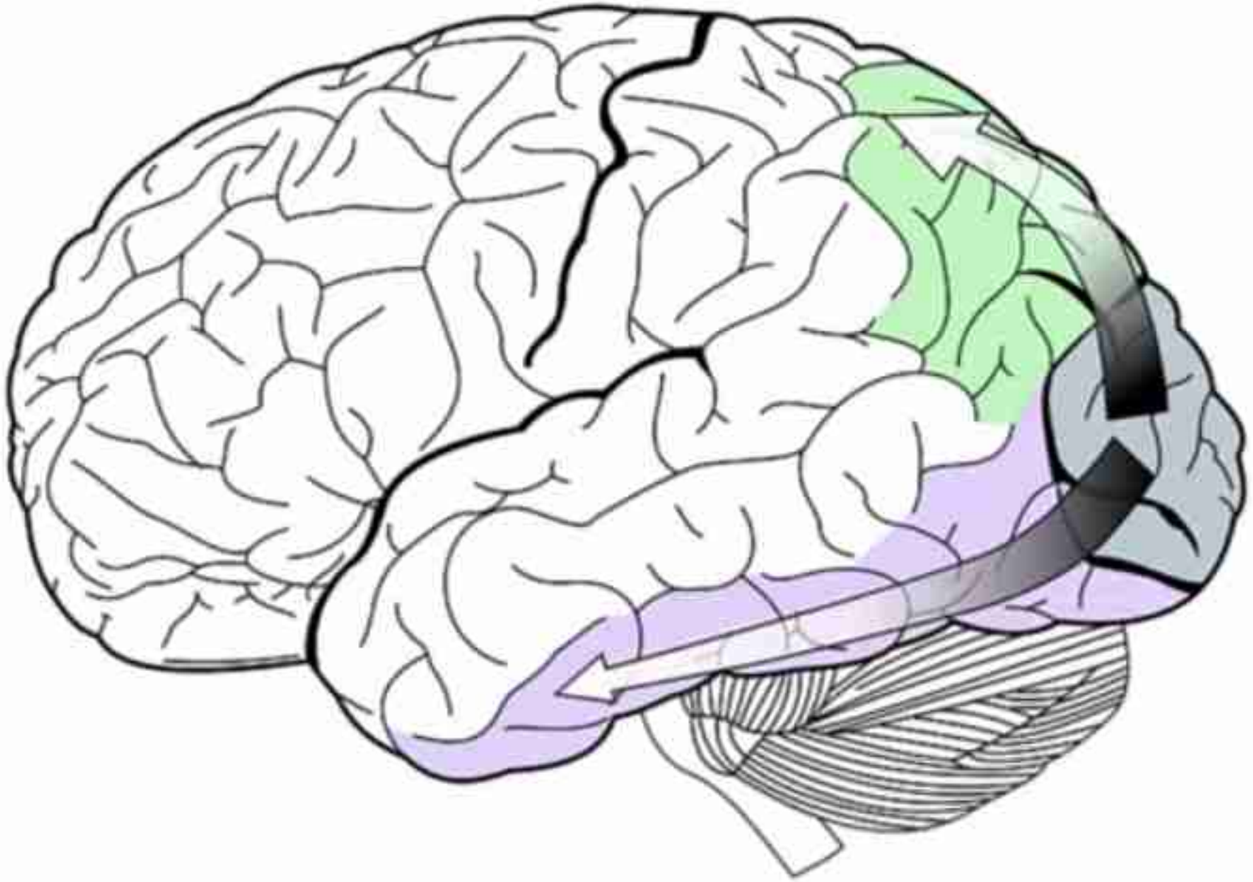

Choroid
Retina
Lens
Pupil
Optic nerve


weak response
very strong response
very weak response
baseline activity
baseline activity

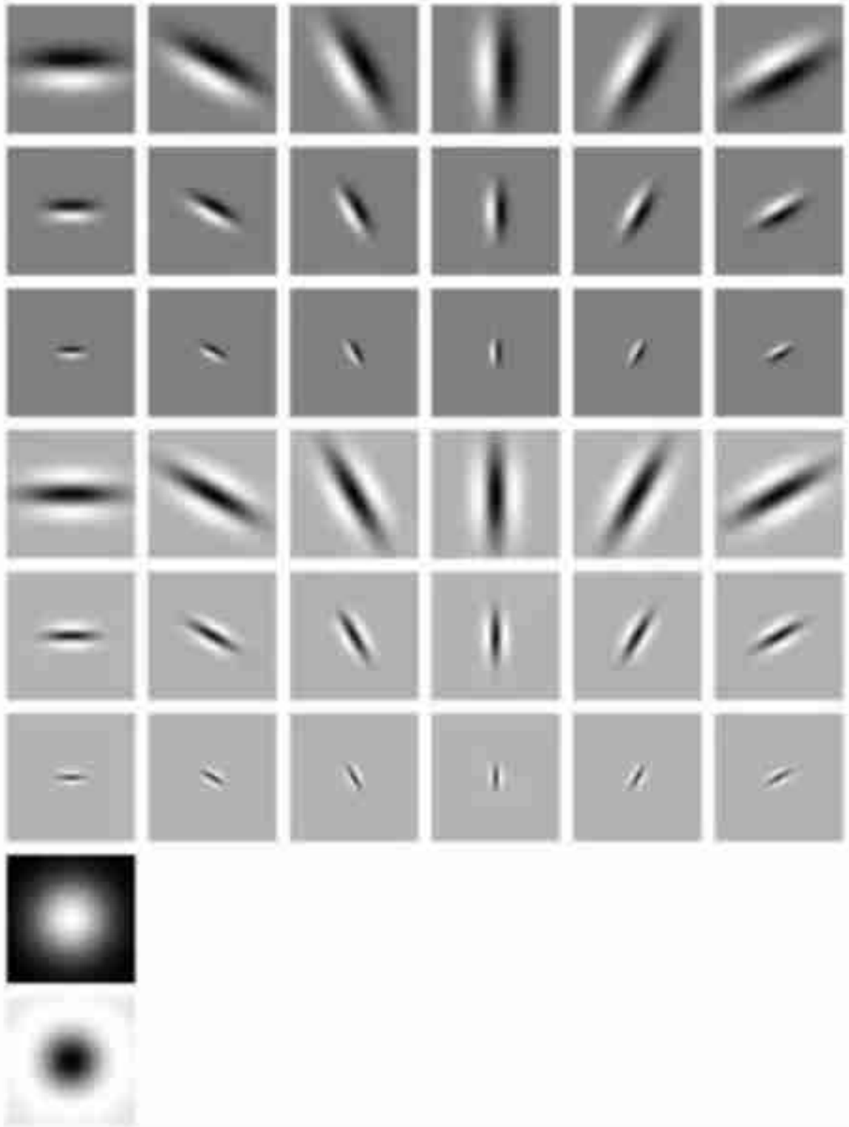

Choroid
Retina
Lens
Pupil
Optic nerve


no stimulus
OFF
ON
cell baseline activity level
ON
OFF
strong response
ON
OFF
strong response
weak response
ON
no response
ON


Choroid
Retina
L
P



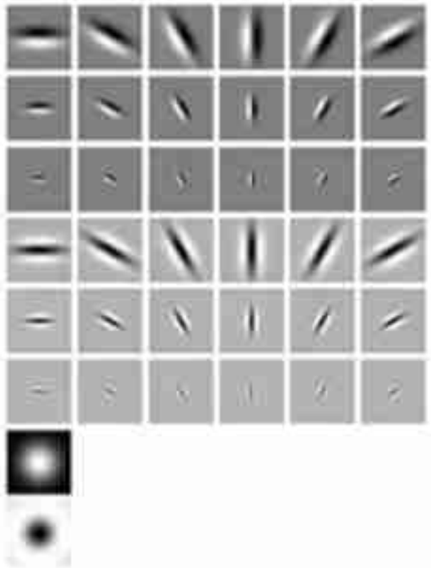
Share



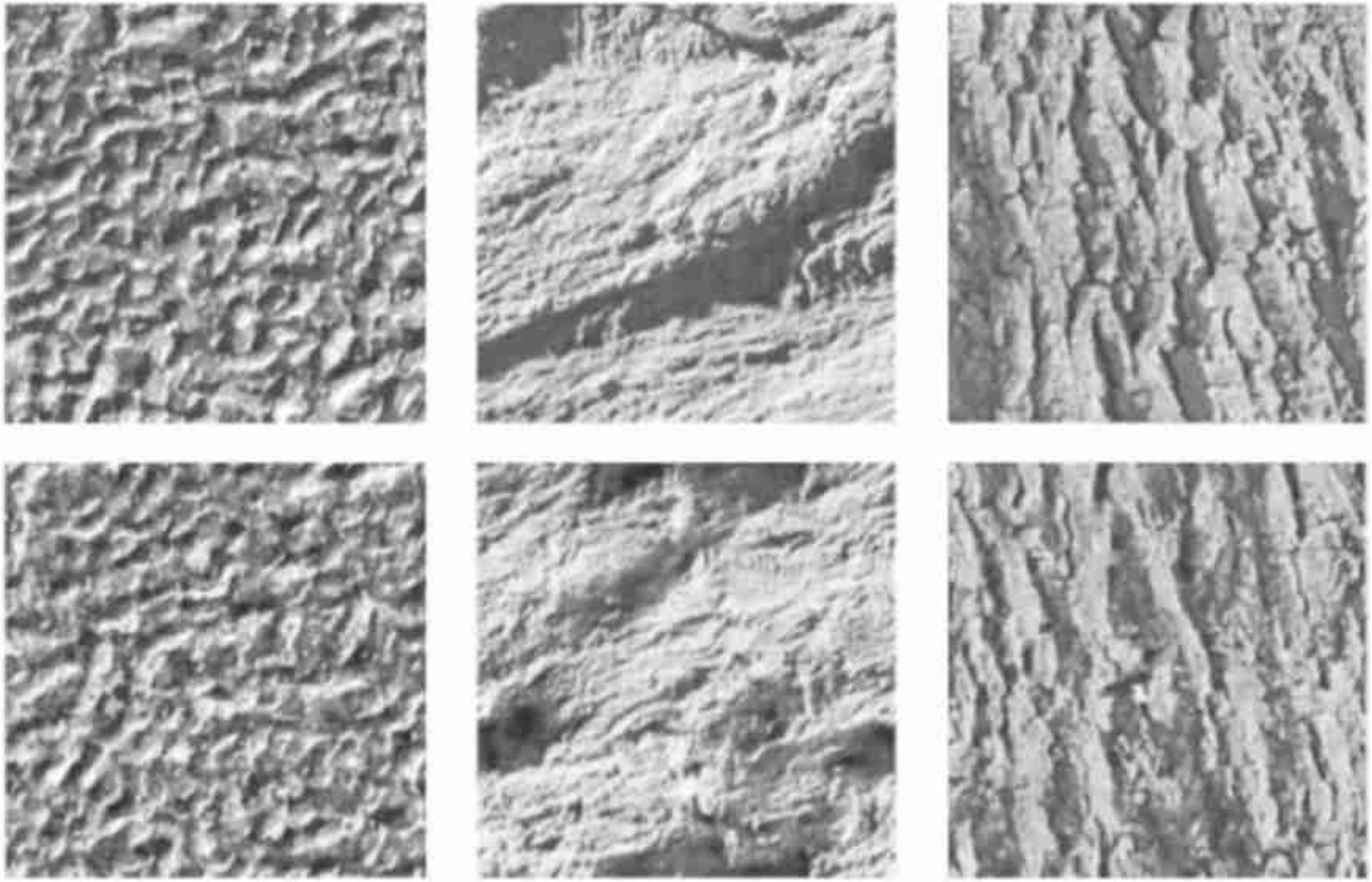
Filters

Convolution

1:43 / 8:36



Synthesized Texture



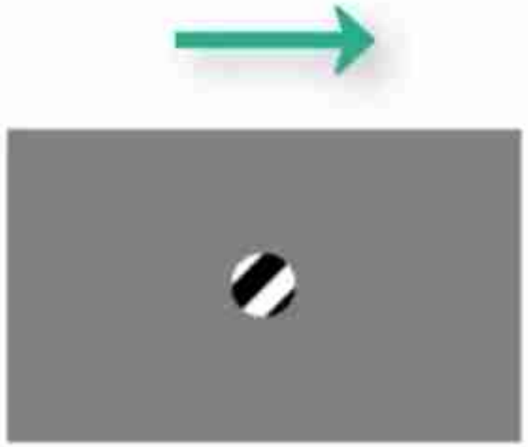
Synthesized

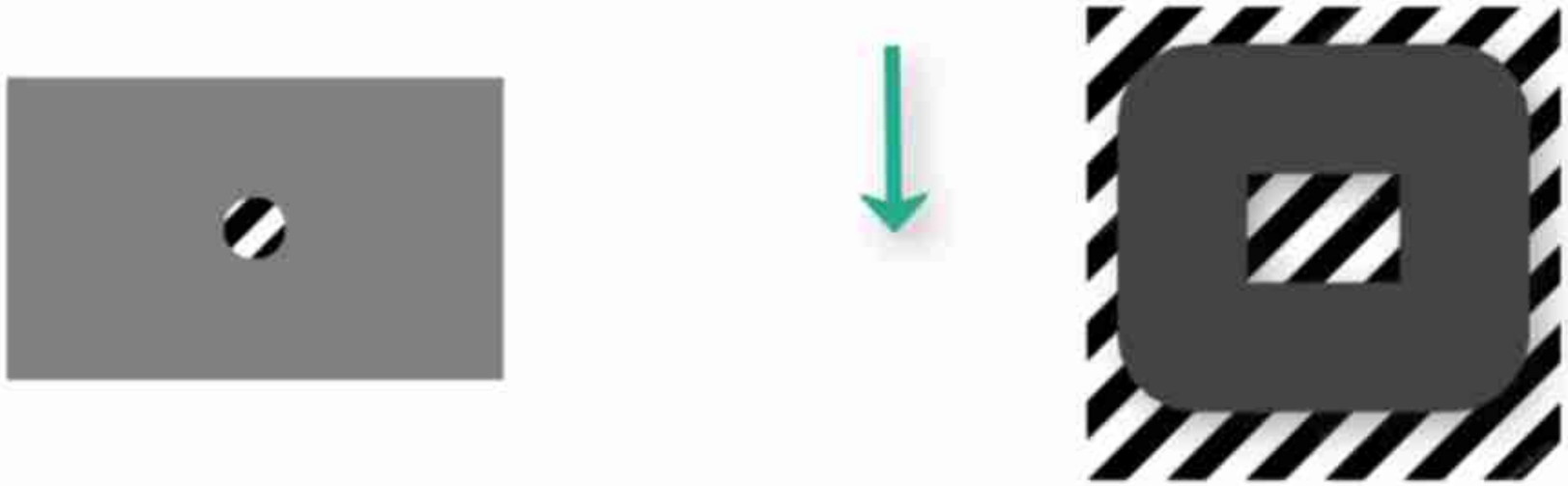
Synthesized

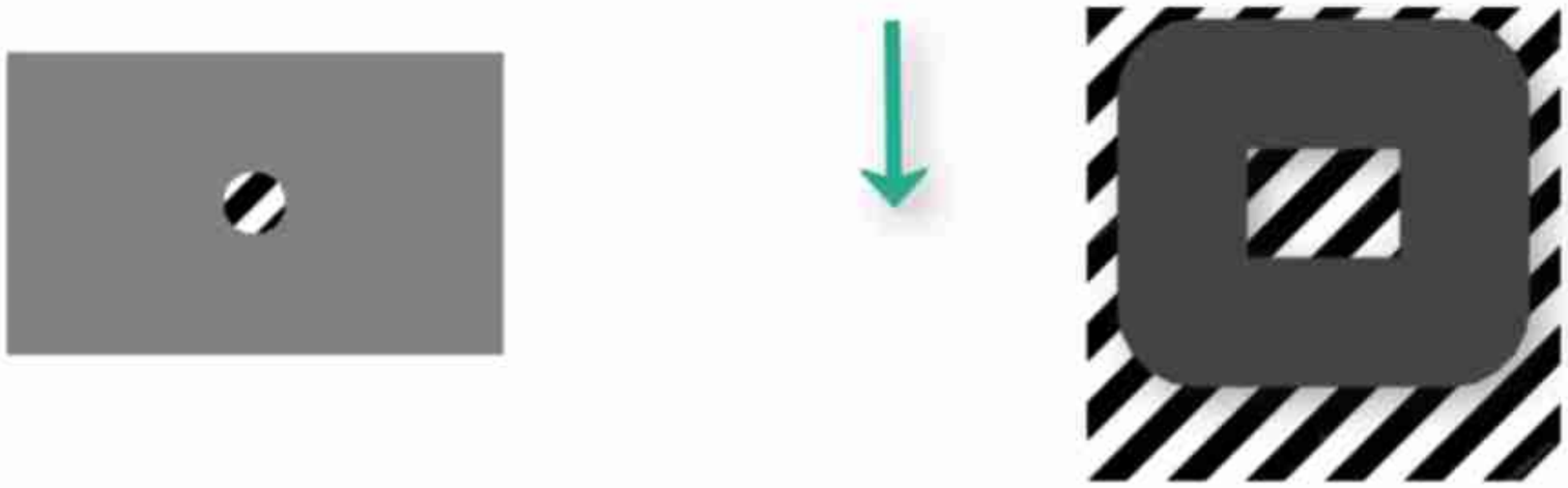
Credit: Tom Ouyang

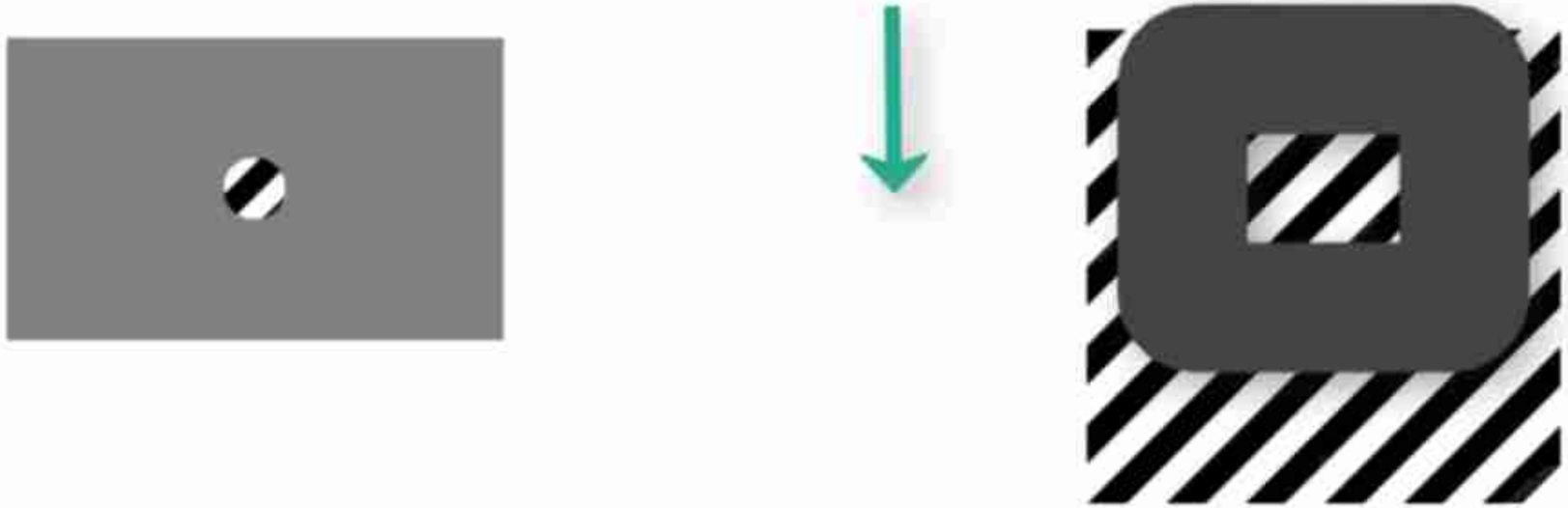
Share

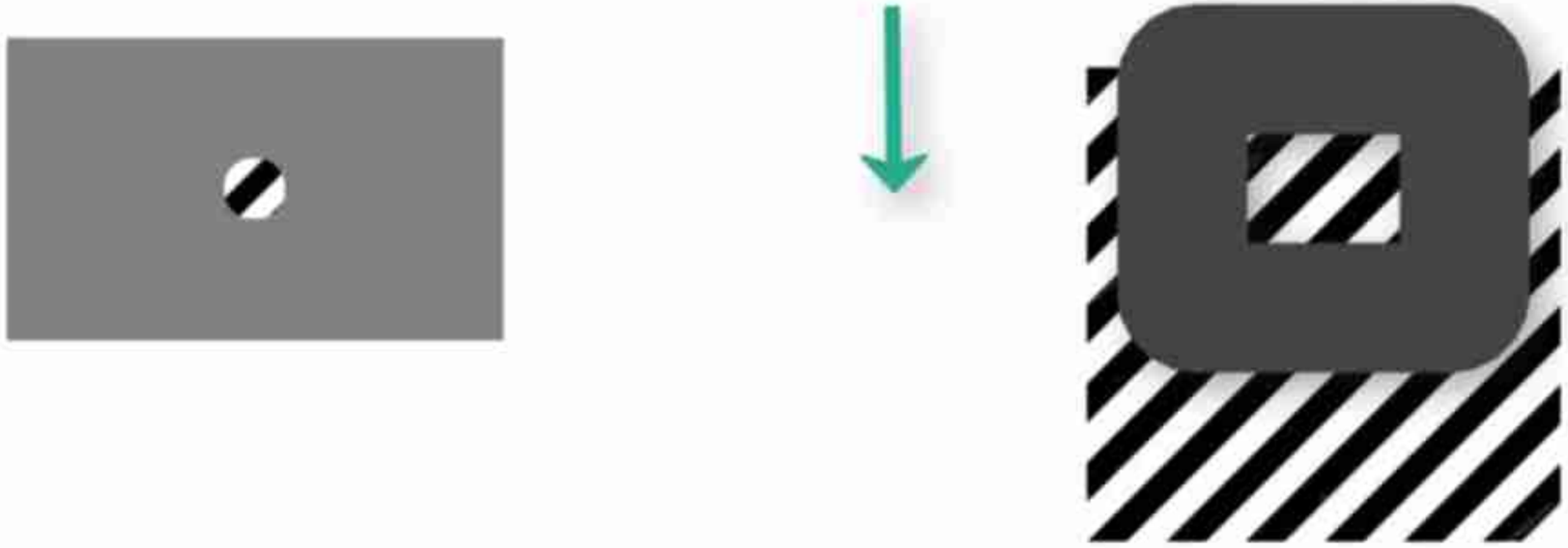
2:22 / 8:36

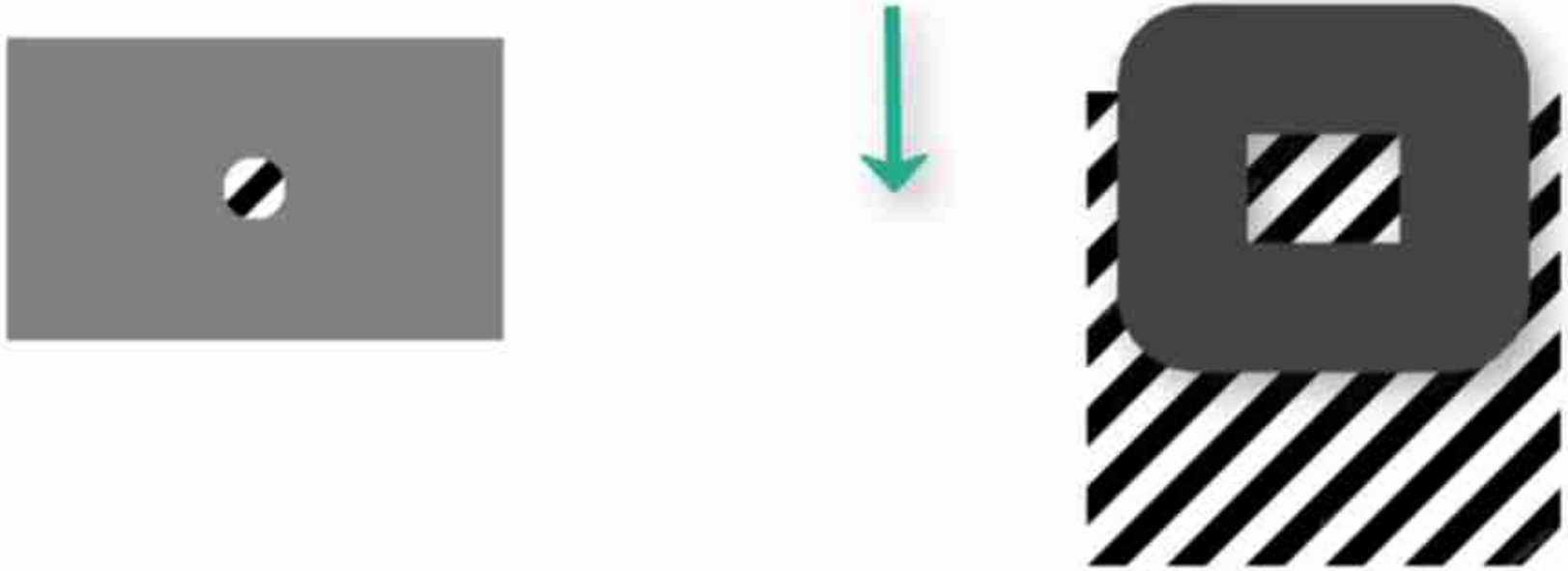


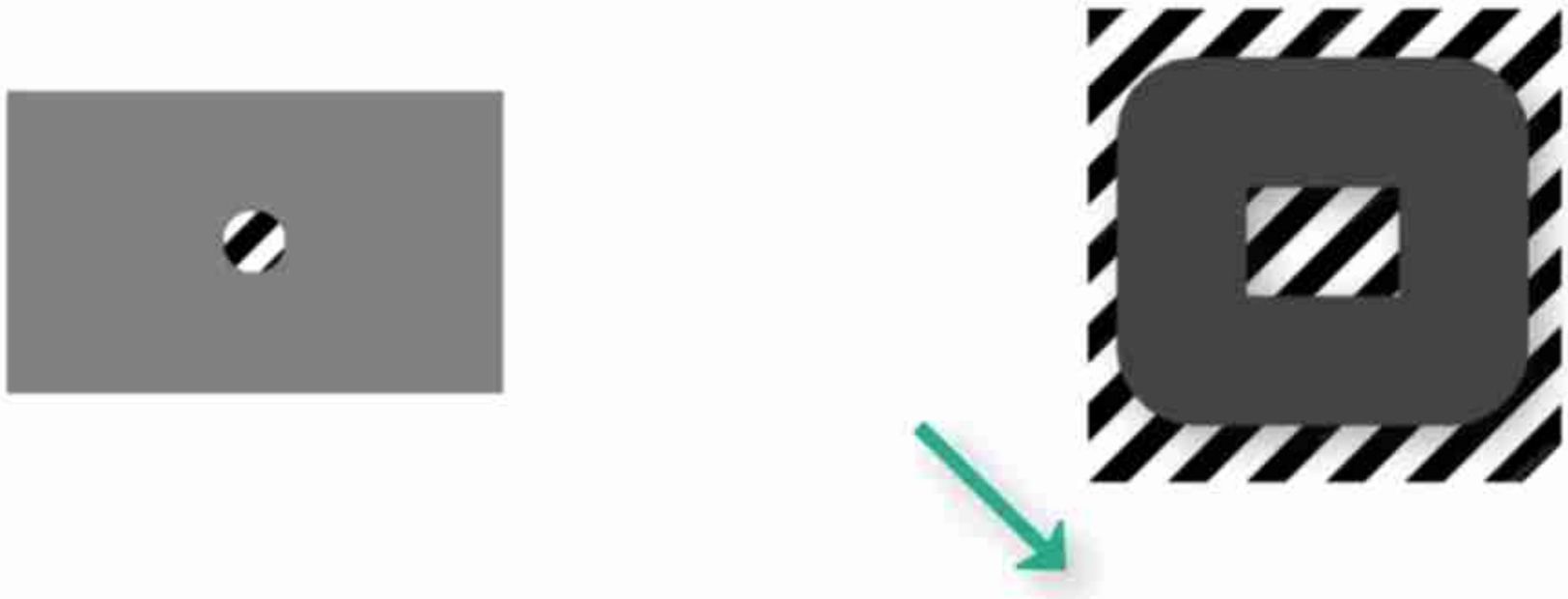


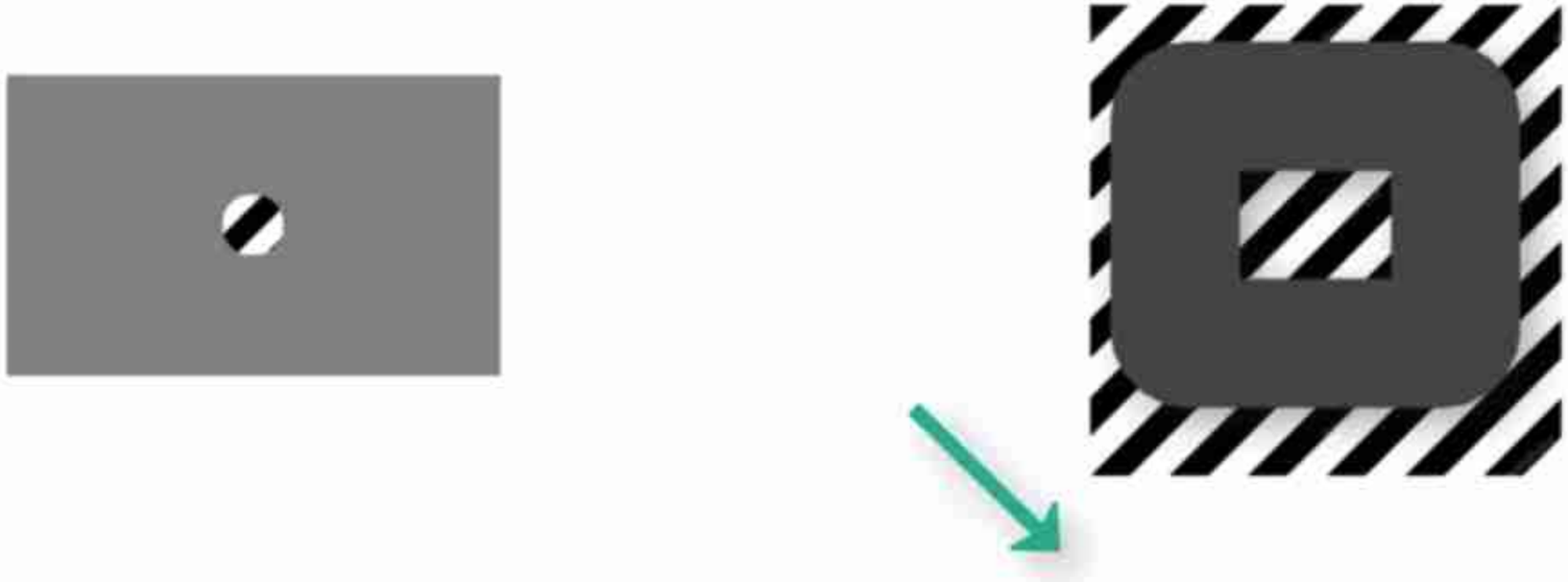


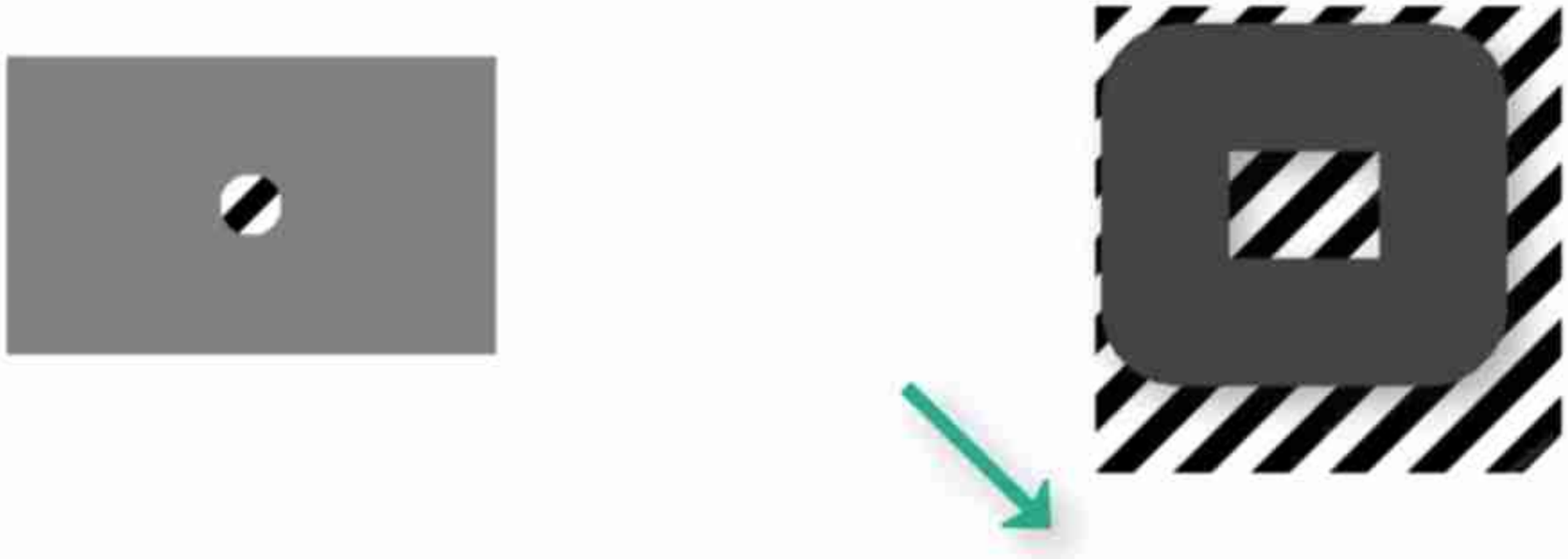


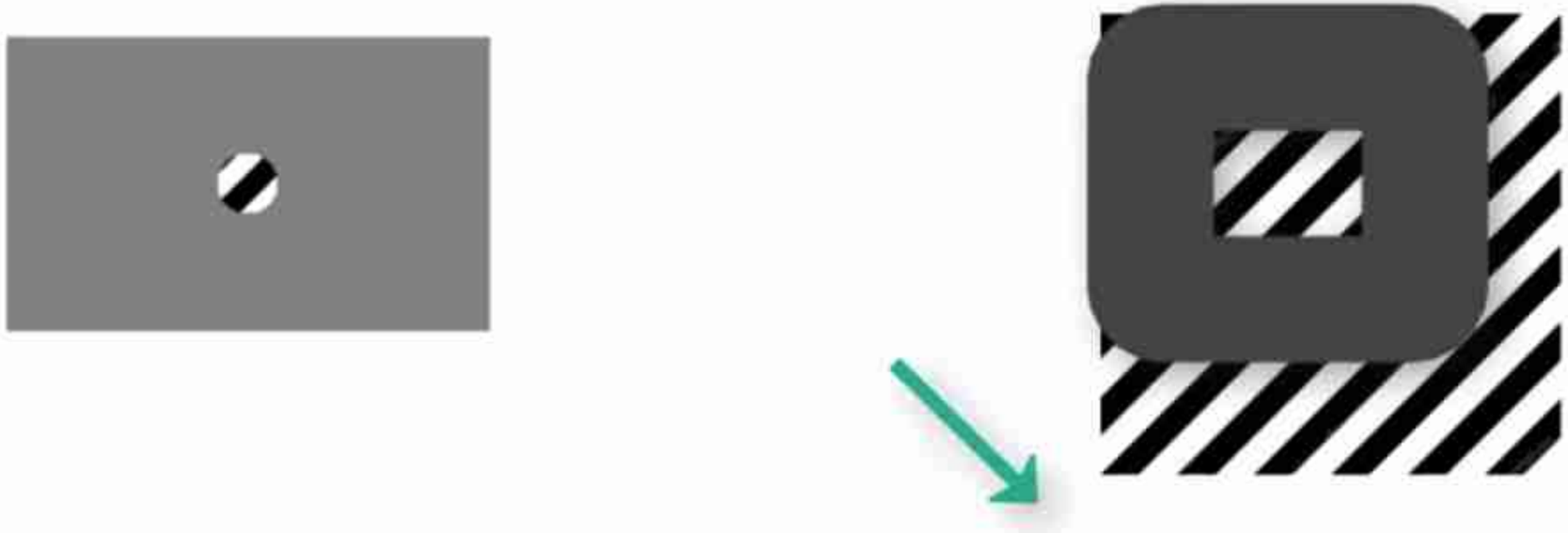


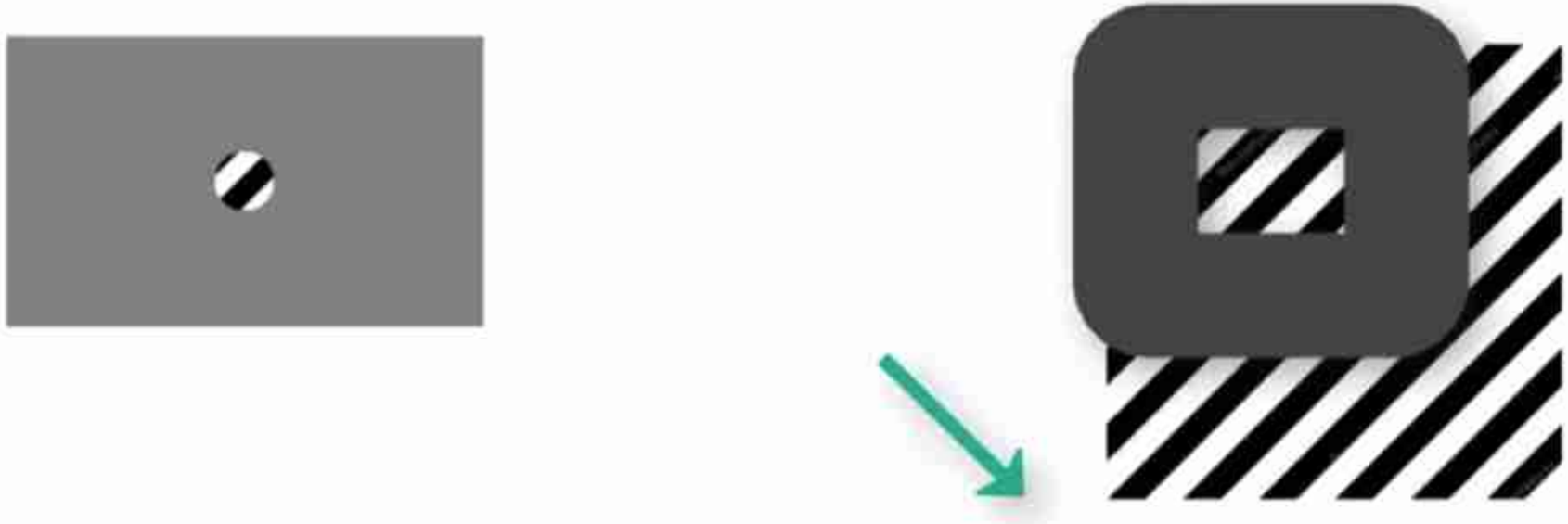




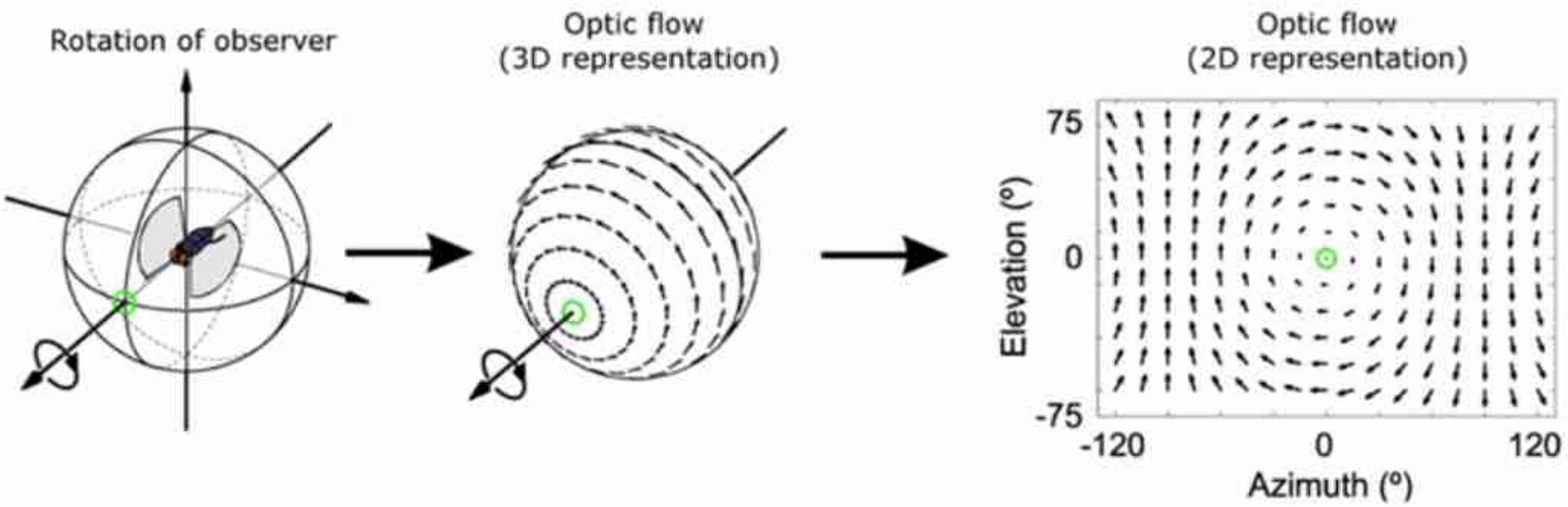


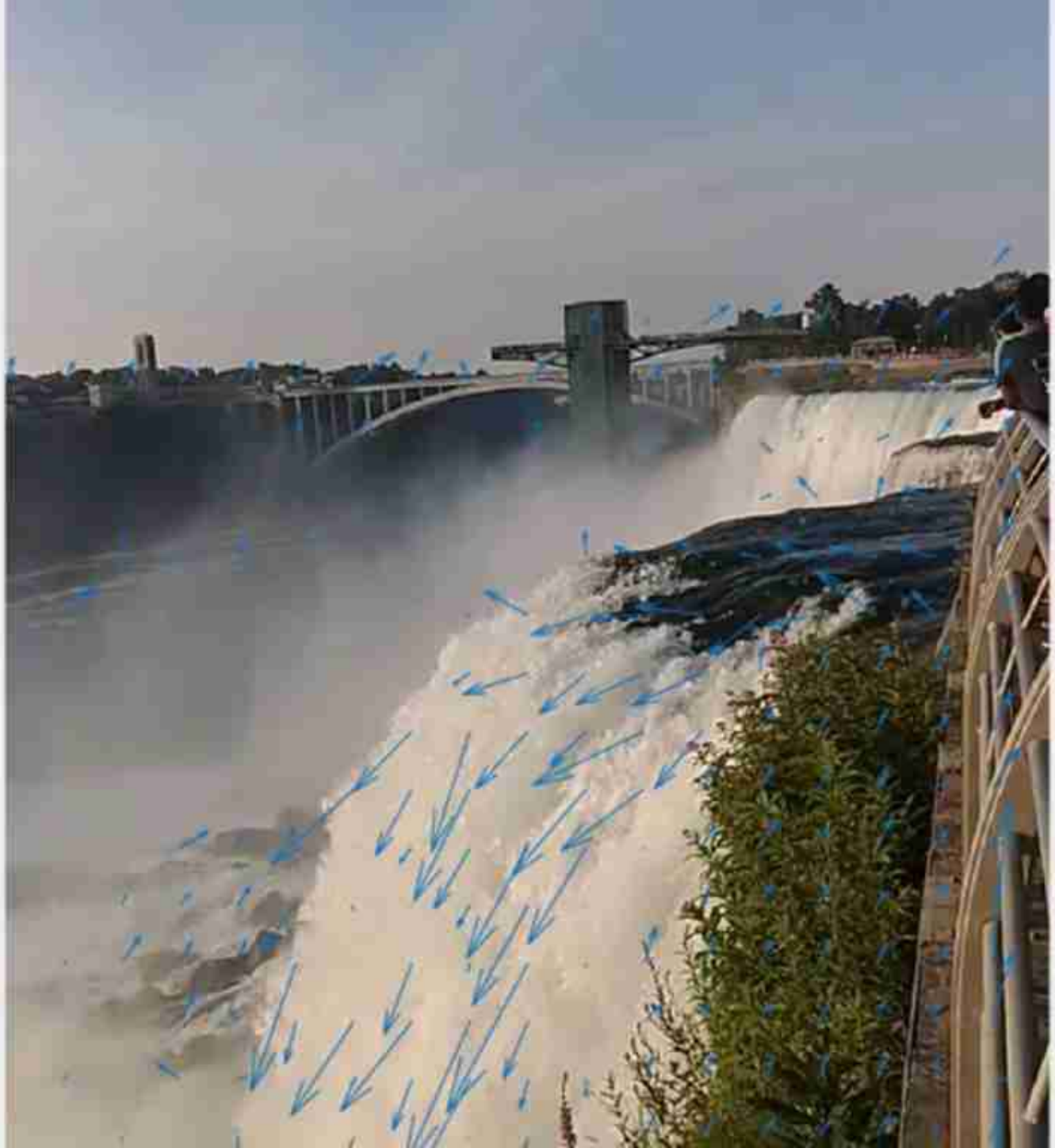


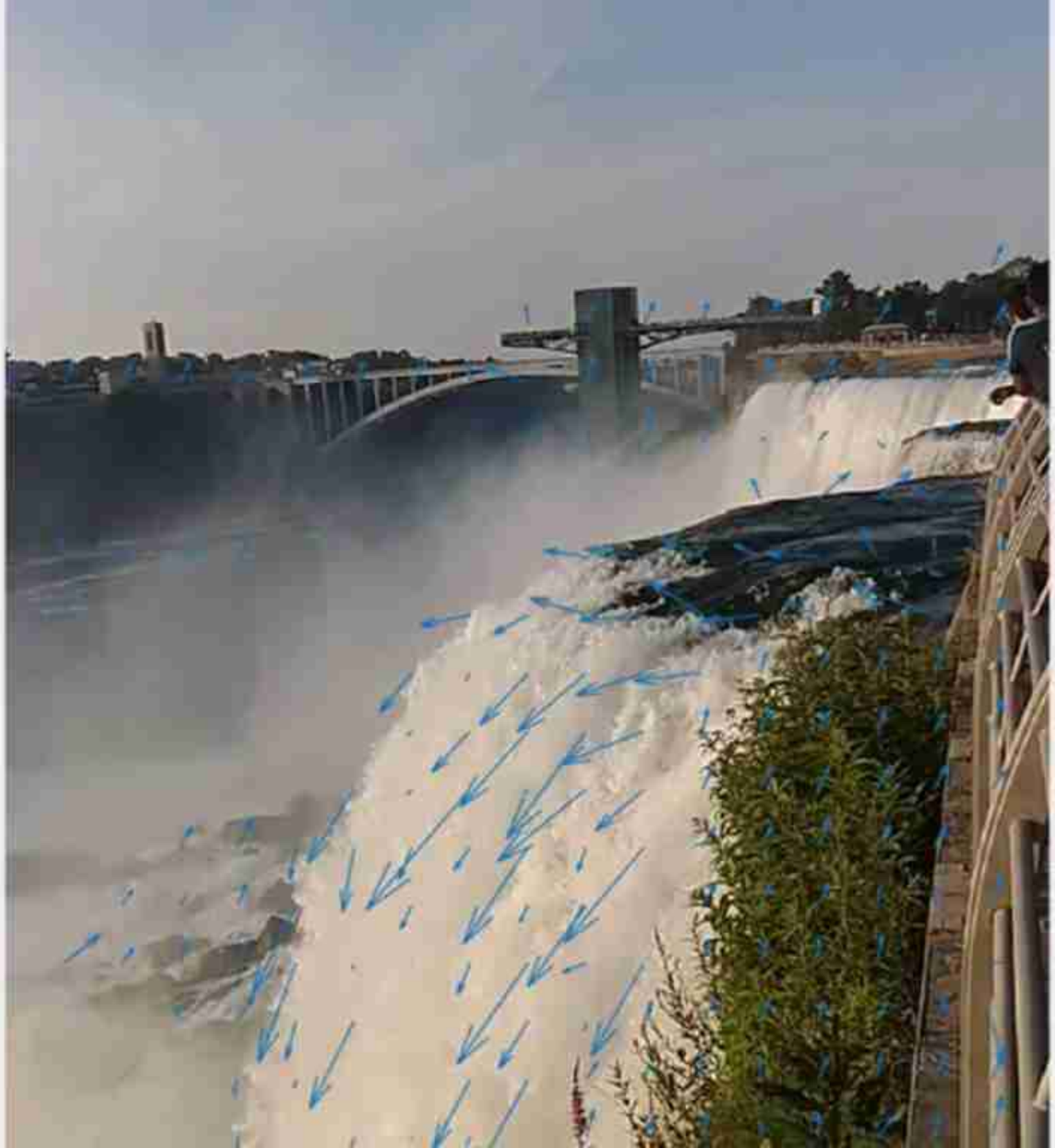


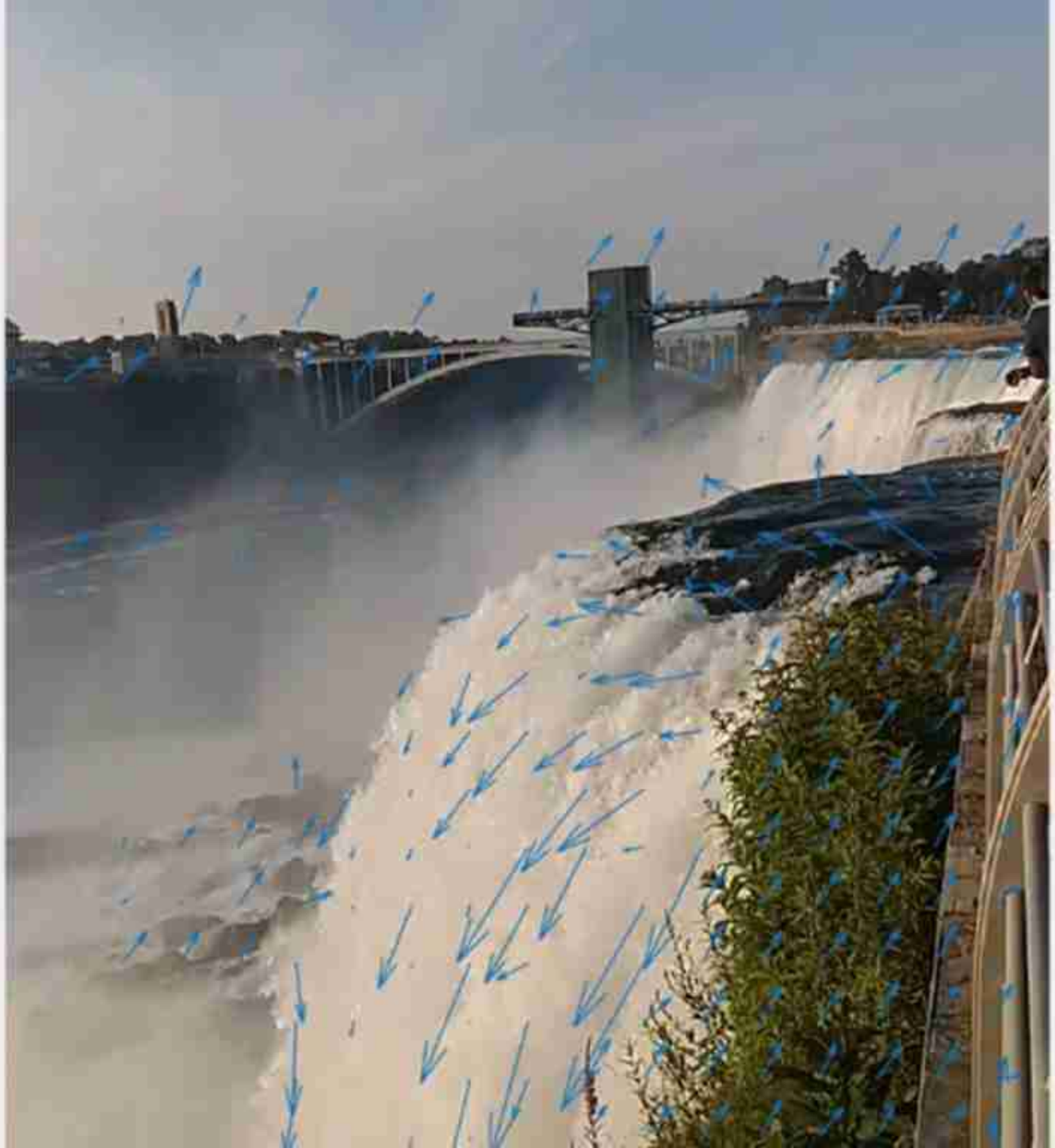


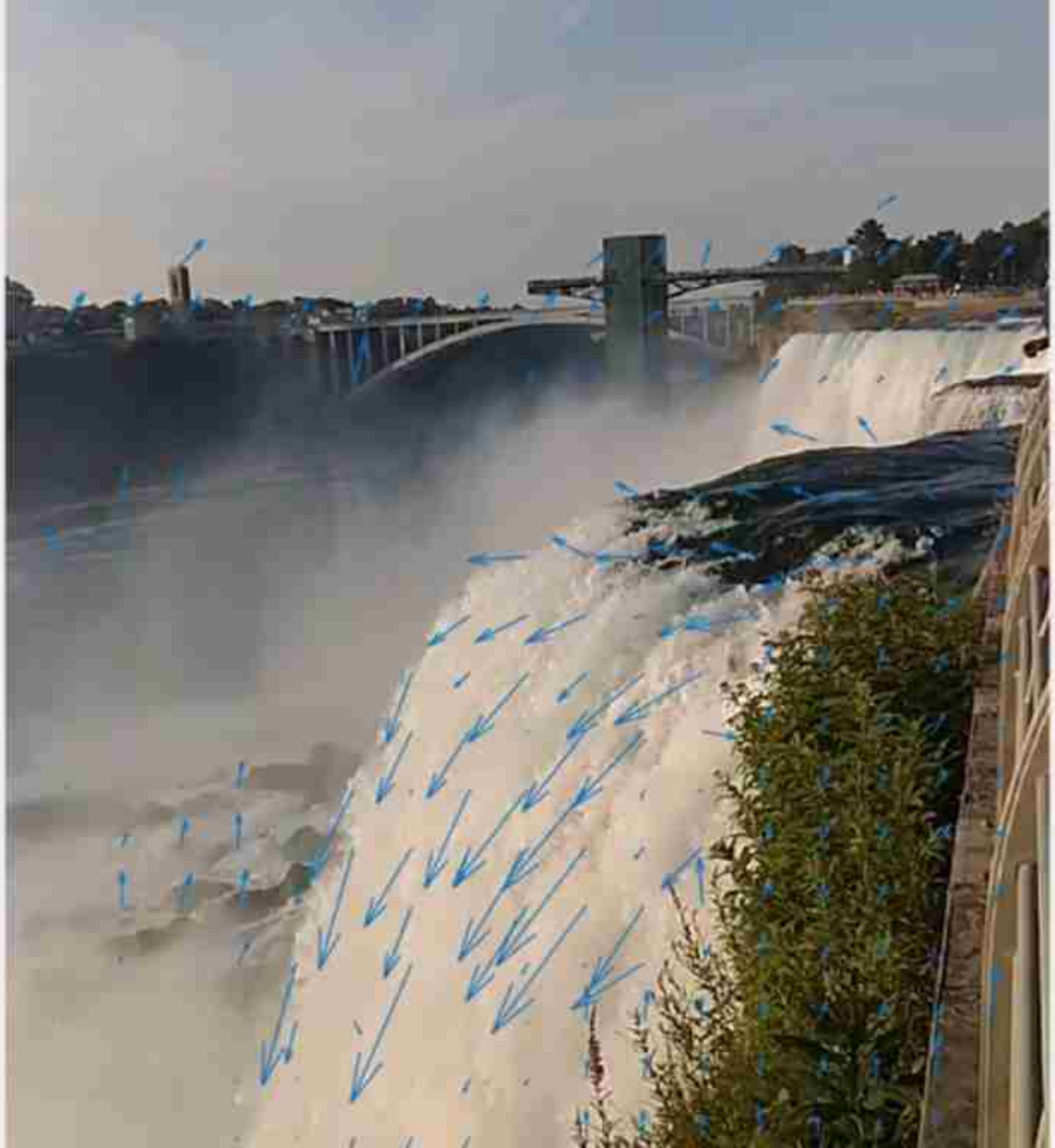
# Aperture Problem


Rotation of observer
Optic flow
(3D representation)
Optic flow
(2D representation)
75
0
-75
Elevation (°)
-120
0
120
Azimuth (°)

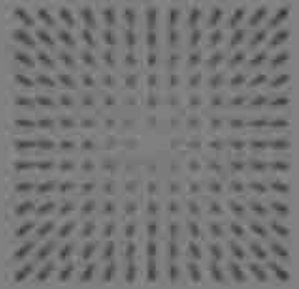






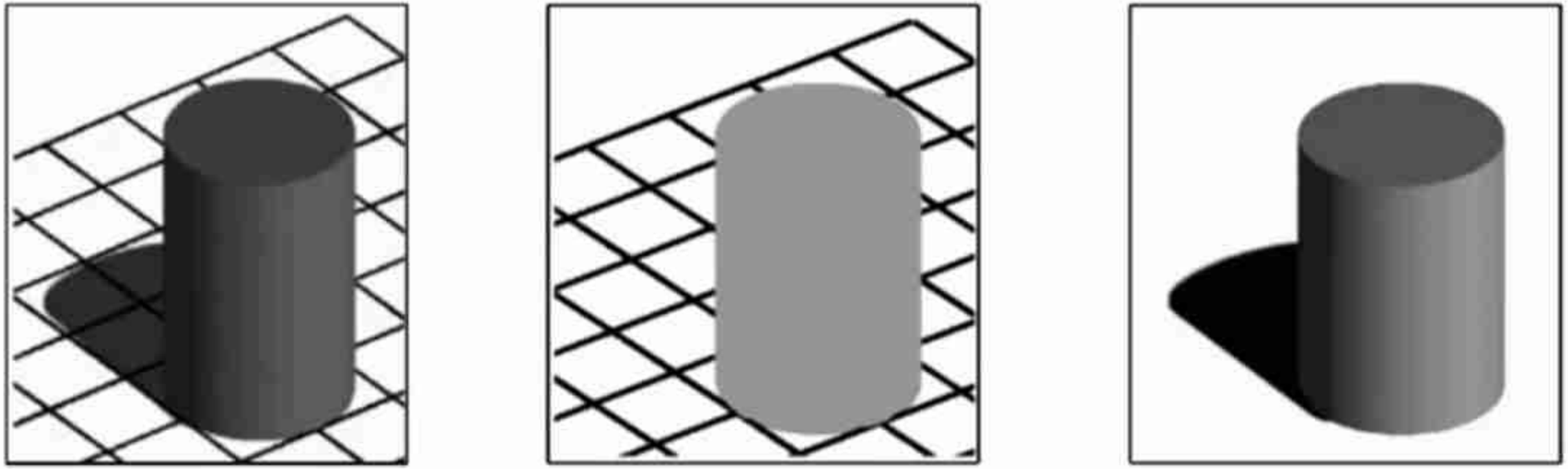


# Optical Flow

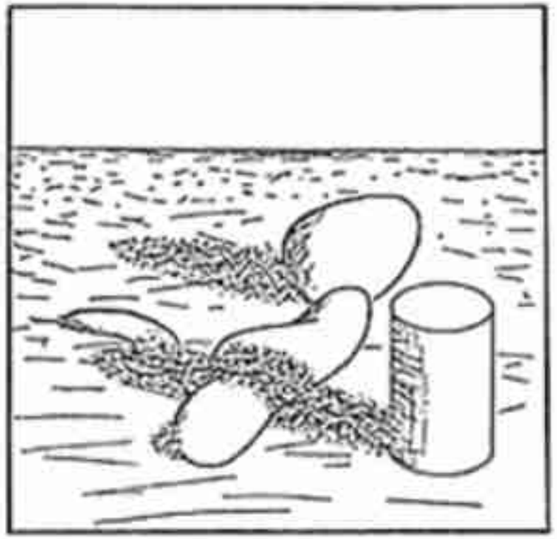


Source: C. Liu. Beyond Pixels: Exploring New Representations and Applications for Motion Analysis

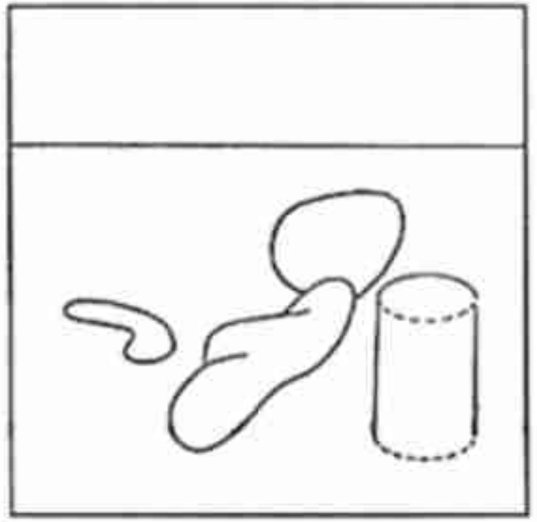
Input = reflectance × illumination



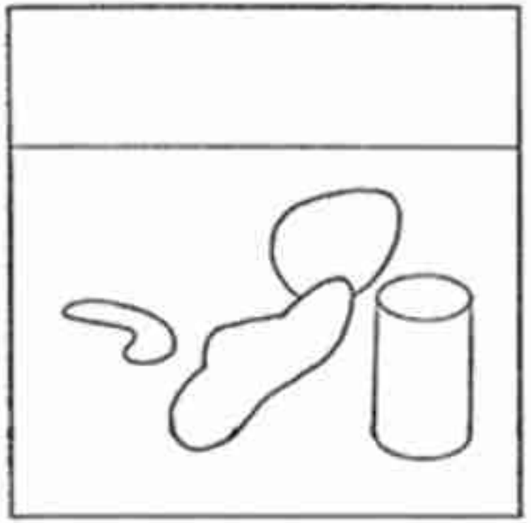
Source: Yair Weiss, Deriving Intrinsic Images from image sequences



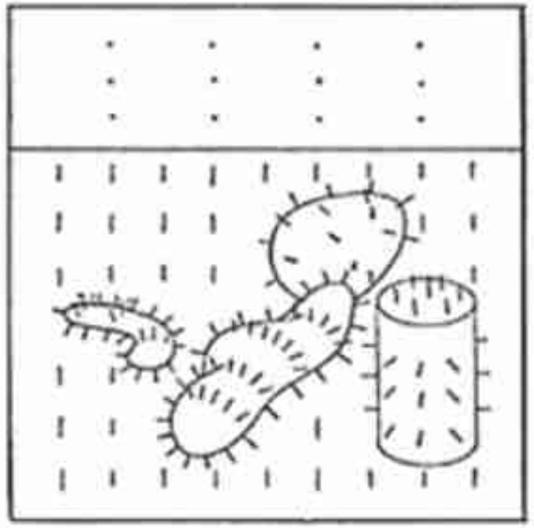
(a) ORIGINAL SCENE



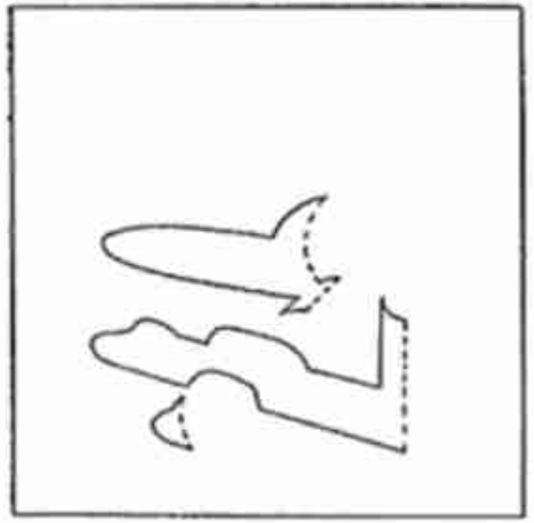
(b) DISTANCE



(c) REFLECTANCE

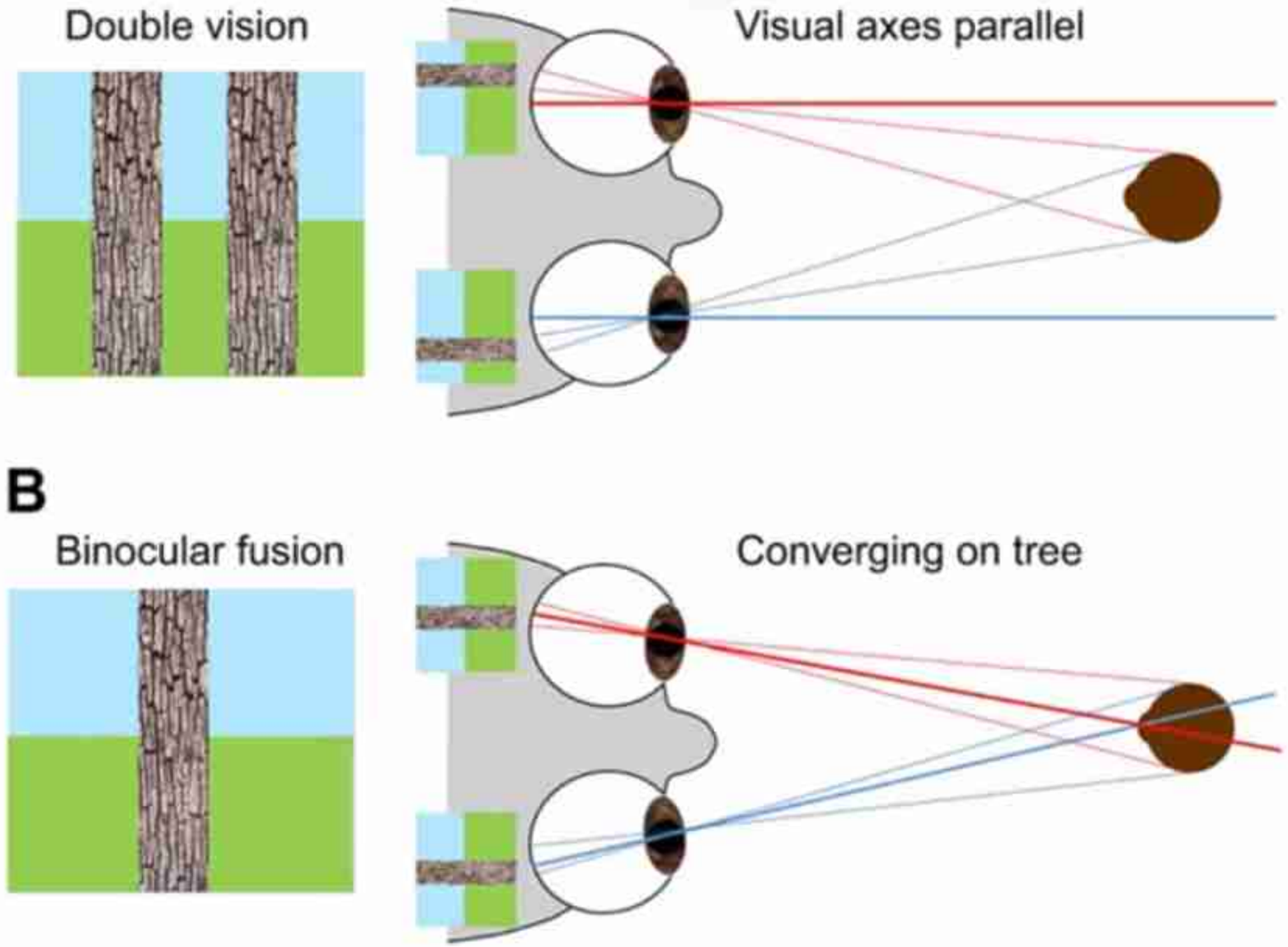


(d) ORIENTATION (VECTOR)

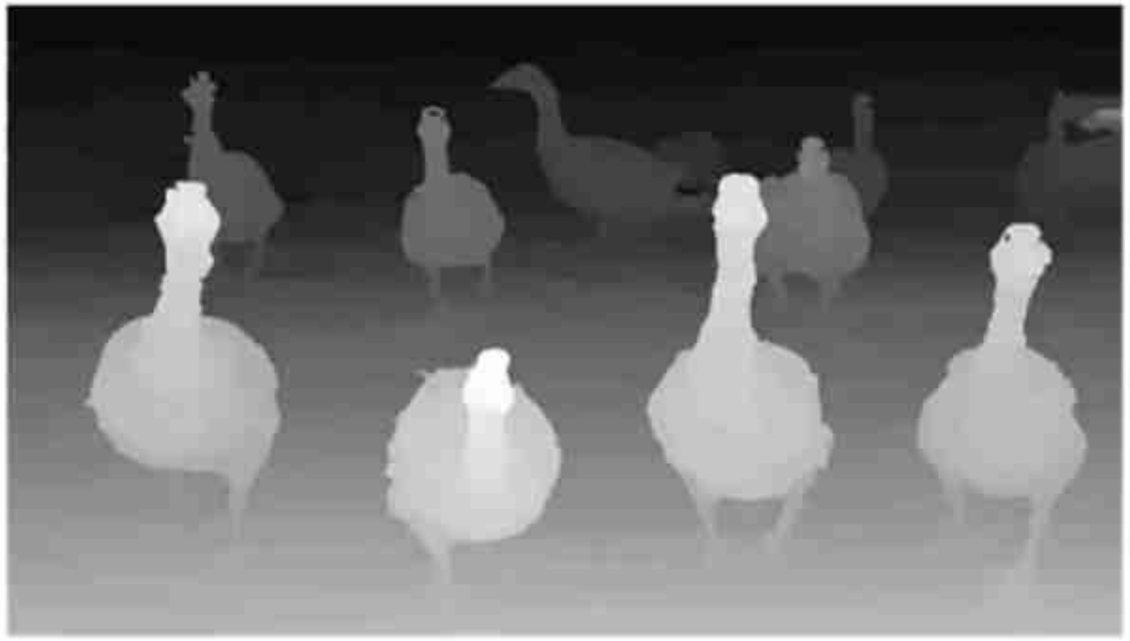


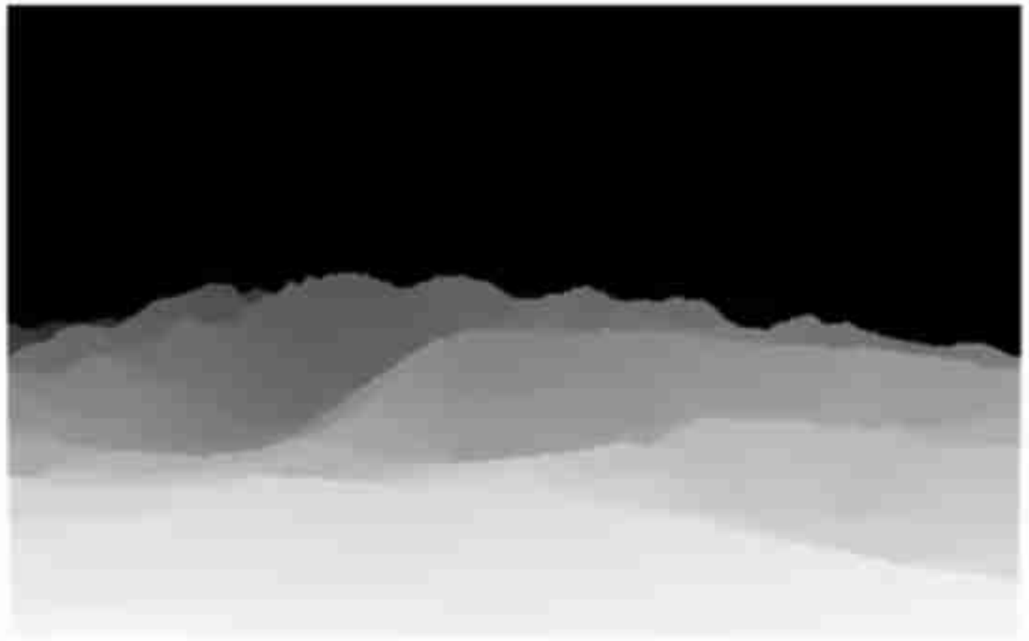
(e) ILLUMINATION

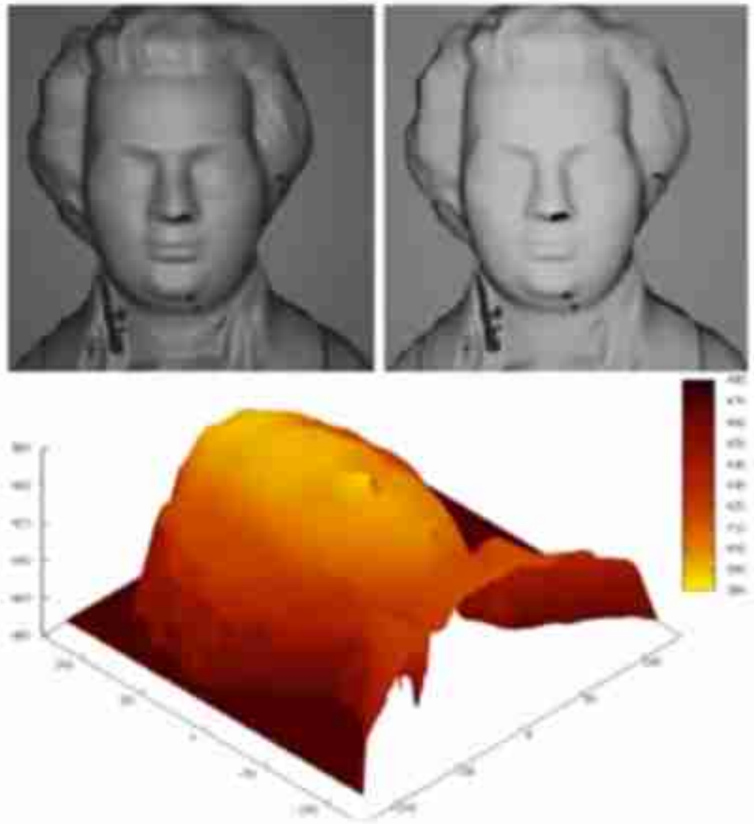
# Stereopsis




Source: Stereopsis in animals: evolution, function and mechanisms
Vivek Nityananda, Jenny C. A. Read







Source: Numerical schemes for advanced reflectance models for Shape from Shading.
Oliver Vogel, Emiliano Cristiani

Corners

1000 Strongest Feature Points from Scene Image

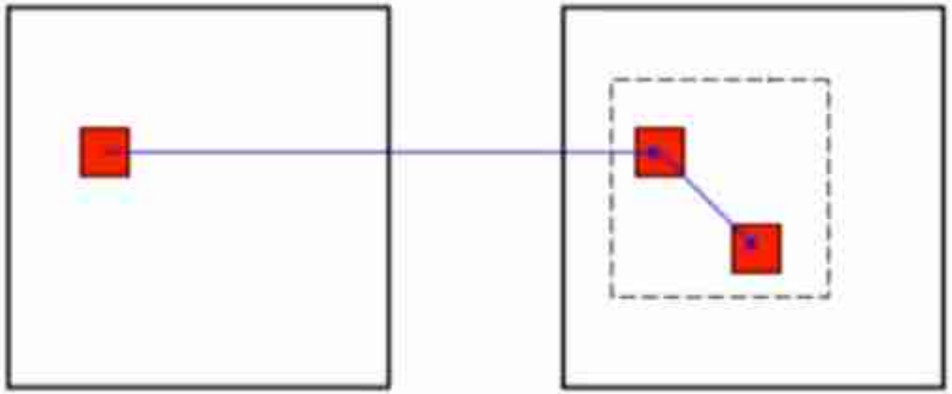
Share

Salient features

5:55 / 8:36

CRAYON

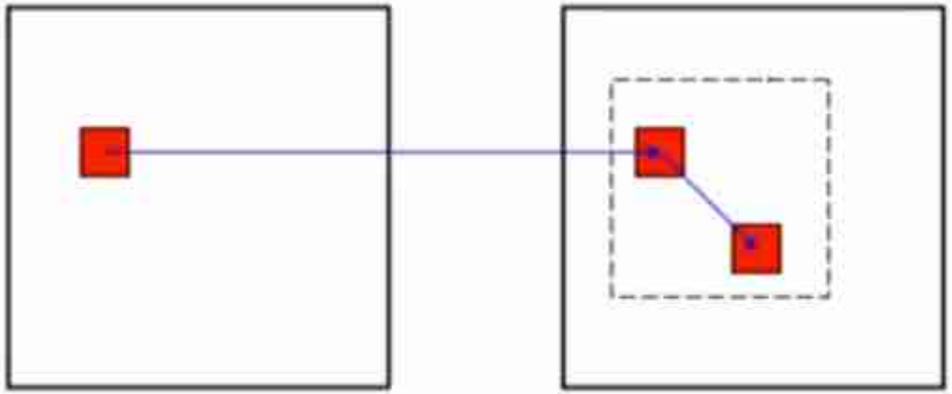

CRAYON



Block Matching

SSD (Sum of Squared Differences)

NCC (Normalized Cross Correlation)

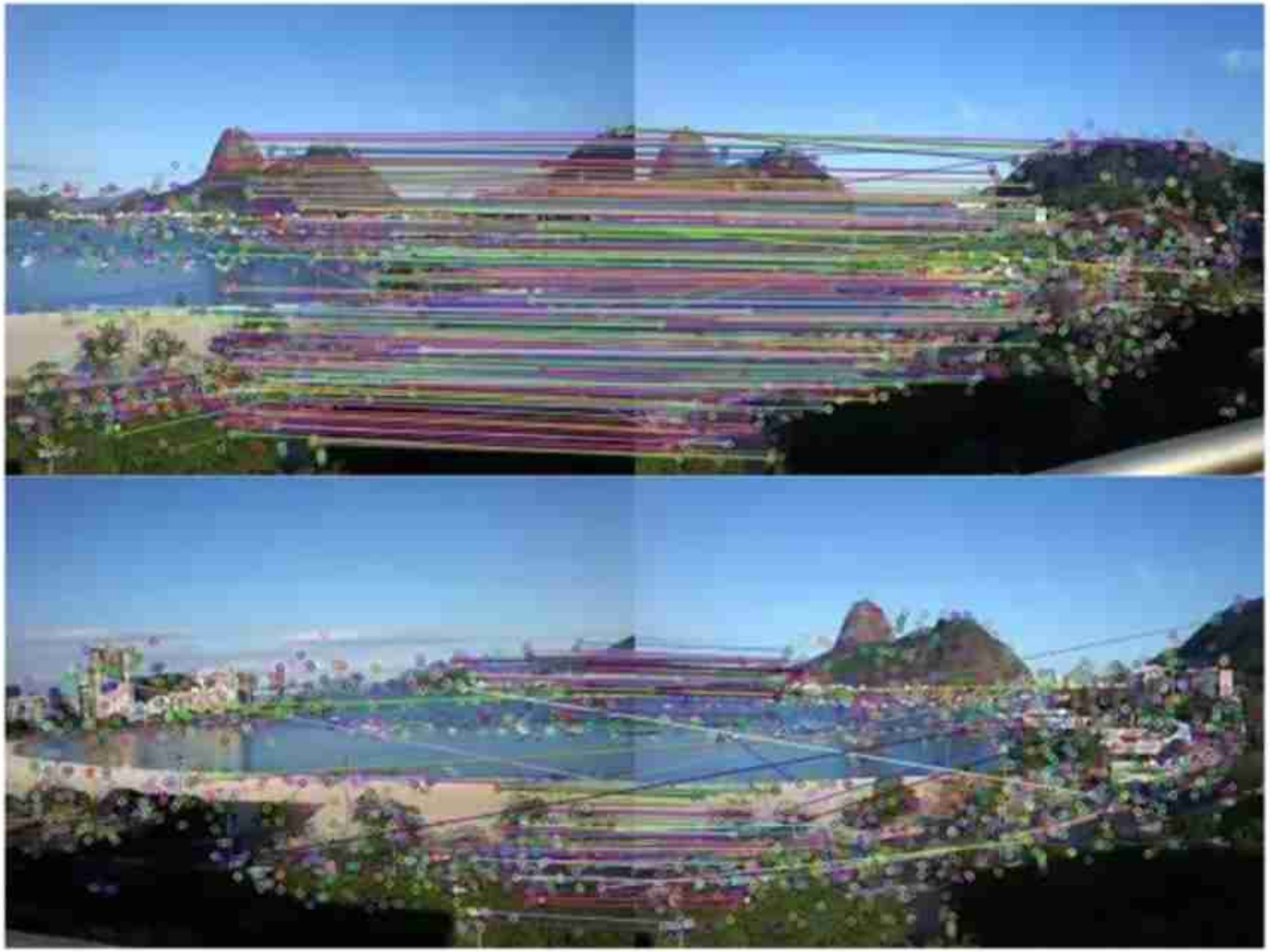


Block Matching

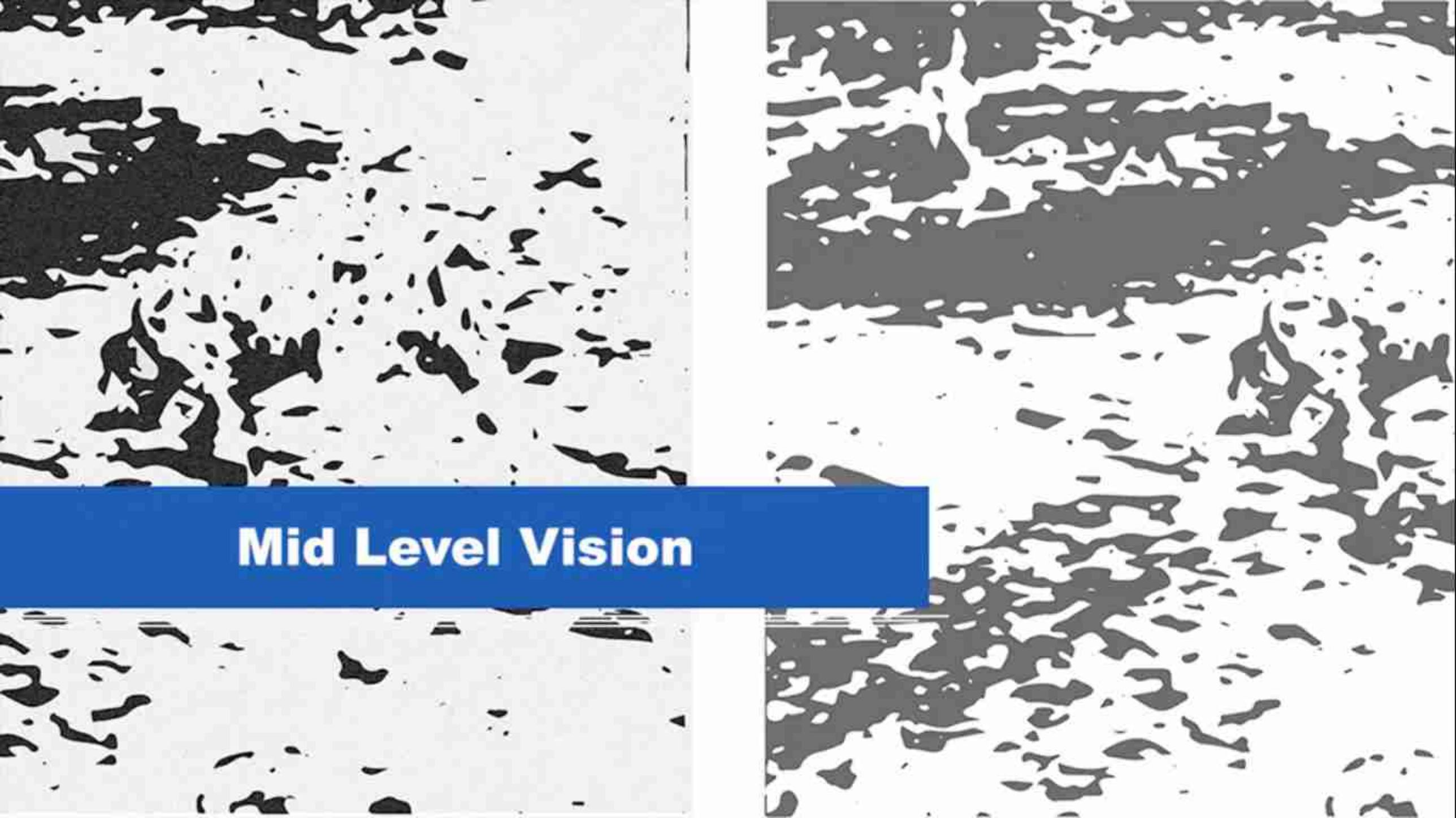
SSD (Sum of Squared Differences)

NCC (Normalized Cross Correlation)

Source: Mobile Panorama Stitching, Can Jin, Xiaopeng Peng, Philip Salvaggio



Source: Mobile Panorama Stitching, Can Jin, Xiaopeng Peng, Philip Salvaggio



# Mid Level Vision

# Mid Level Vision

scene geometry

camera and object motion

# Concepts of Geometric Vision

multiview geometry

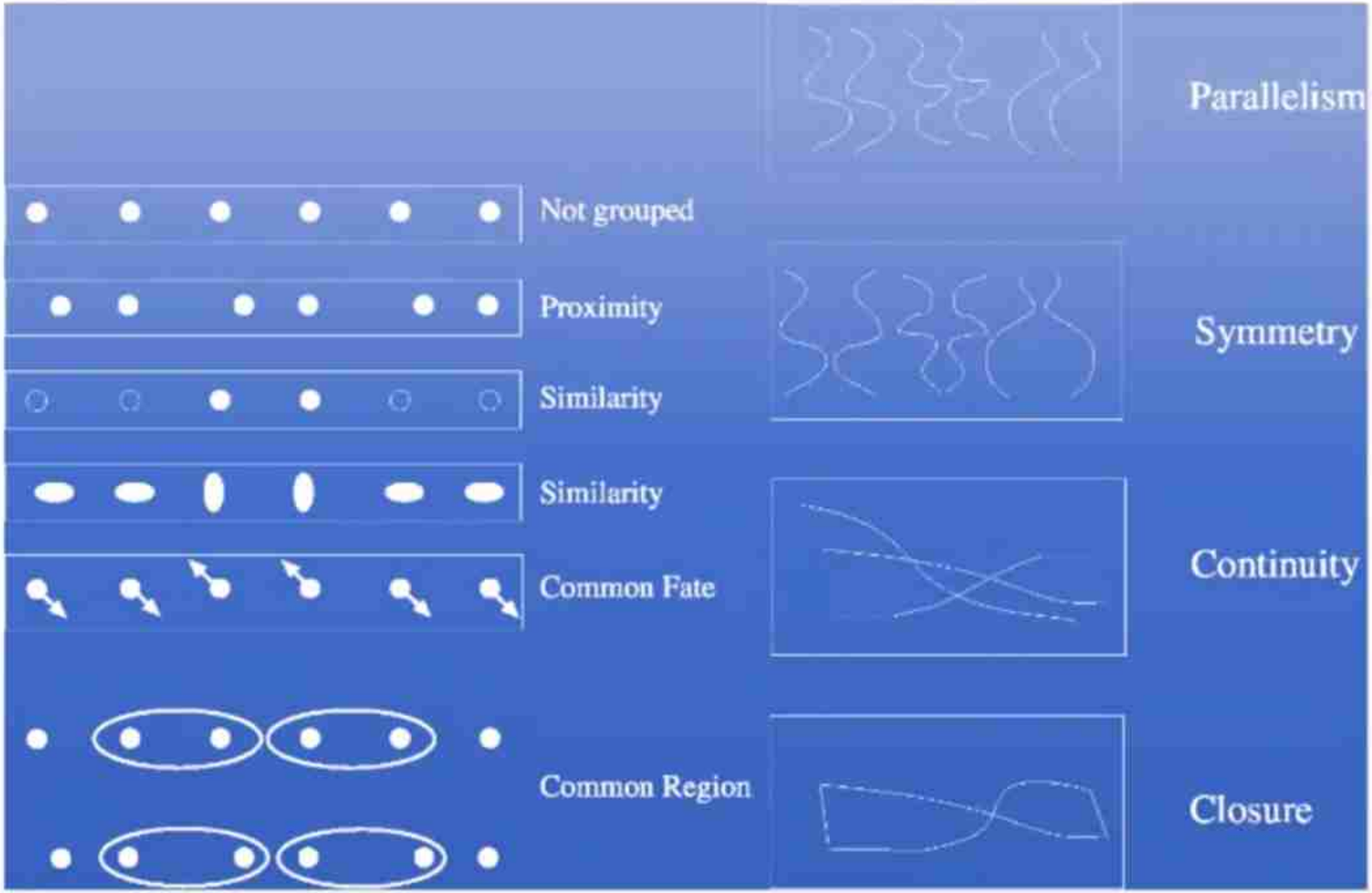
stereo and structure from motion

infer 3D scene information from 2D images

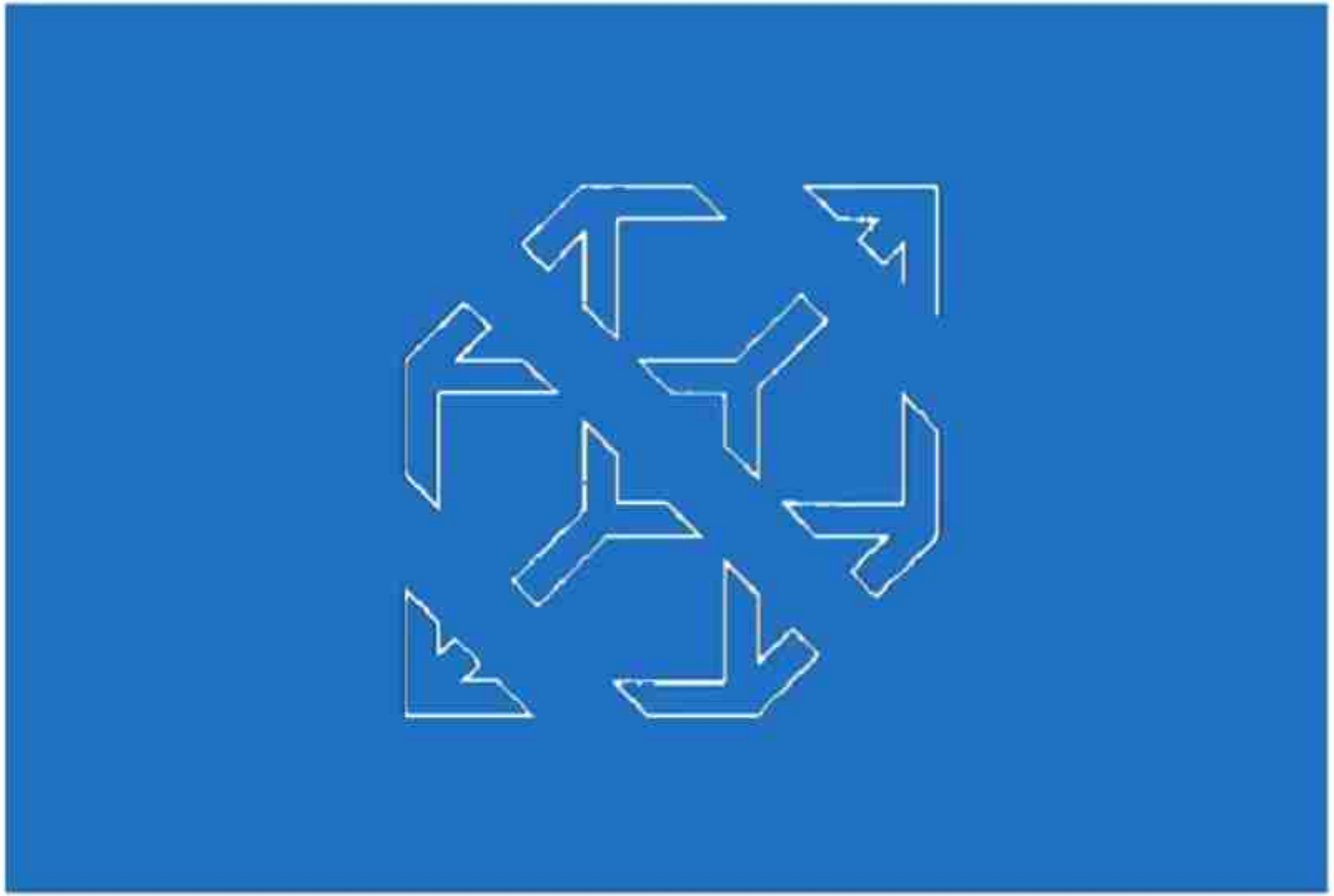
# Image Segmentation

has been a challenging fundamental problem

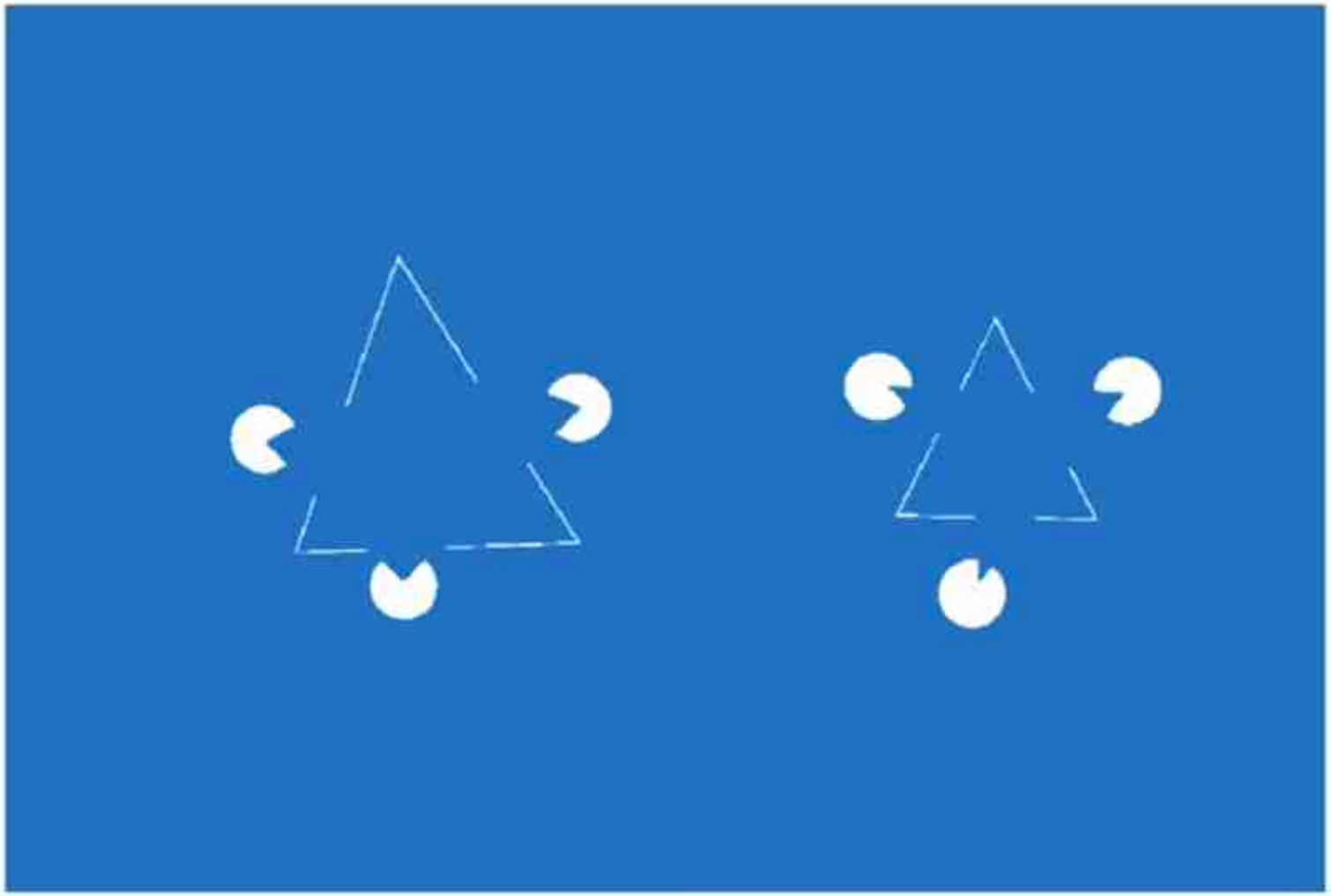
based on spatial similarities and continuities



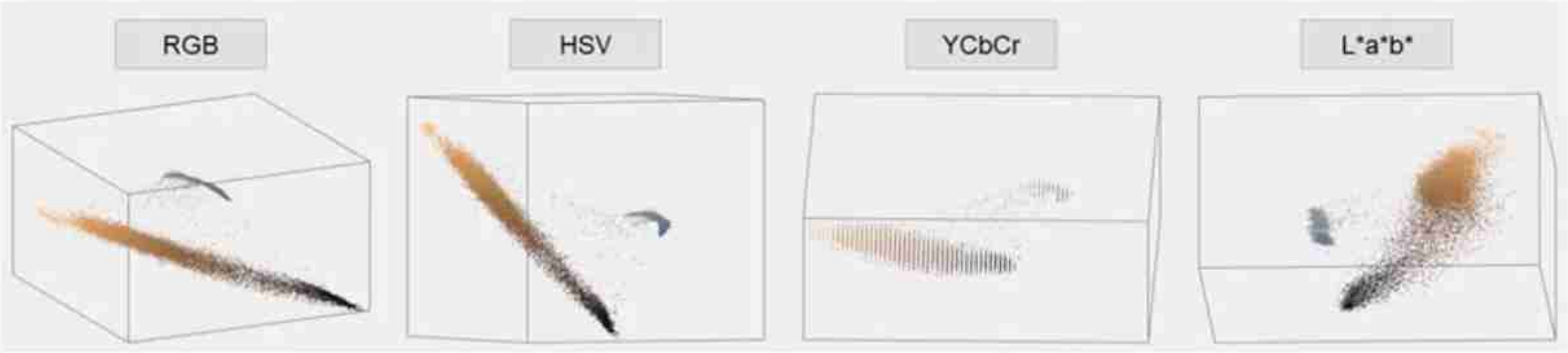

Credit: Gordon, Theories of Visual Perception

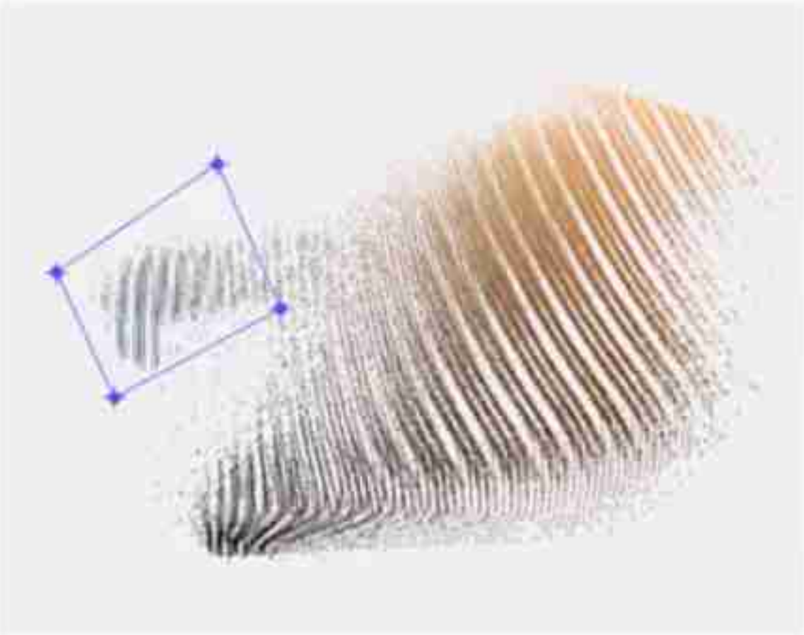


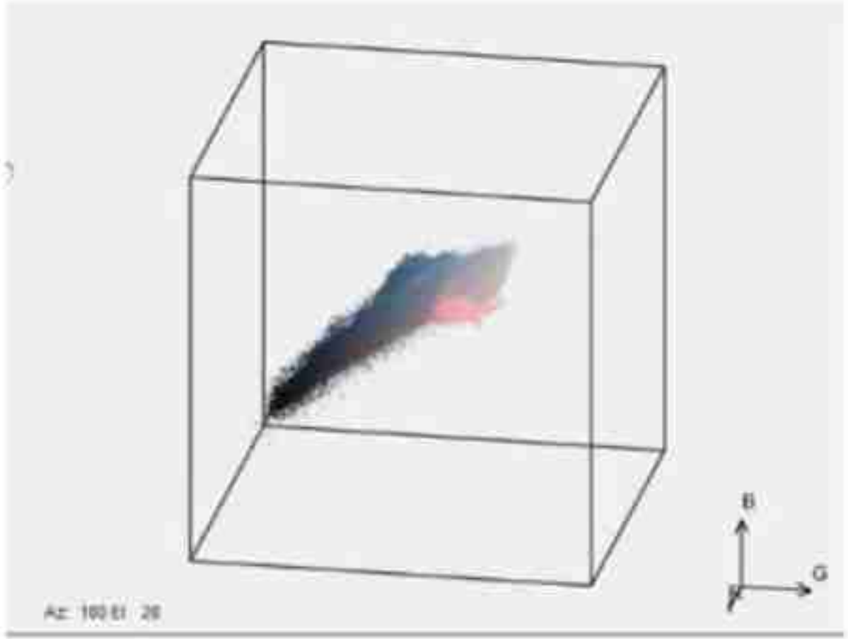
Credit: Gordon, Theories of Visual Perception

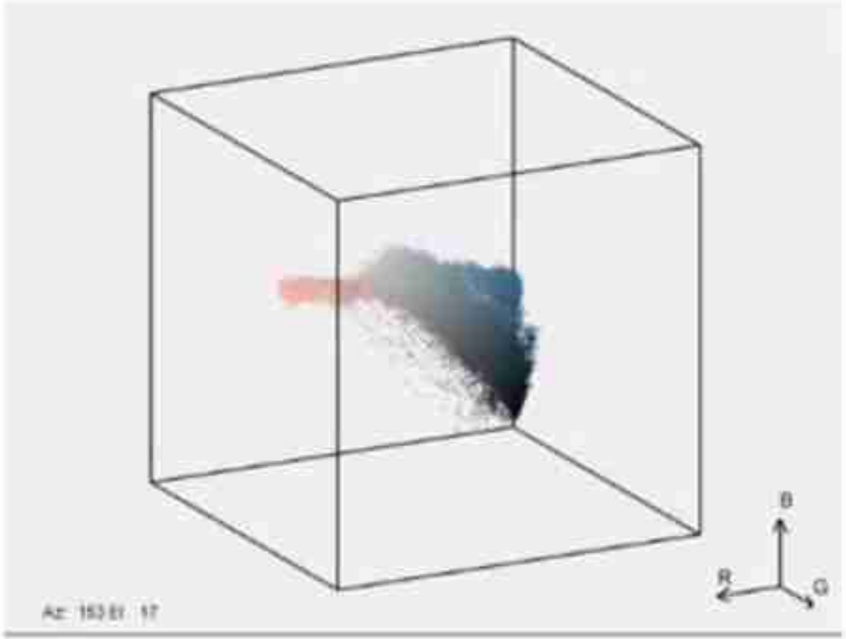


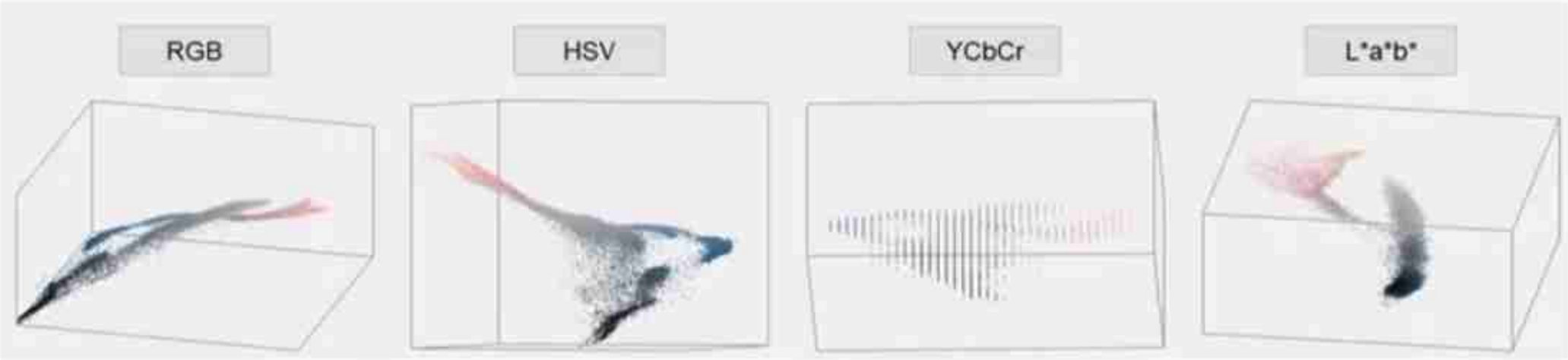
Credit: Gordon, Theories of Visual Perception


RGB
HSV
YCbCr
L*a*b*

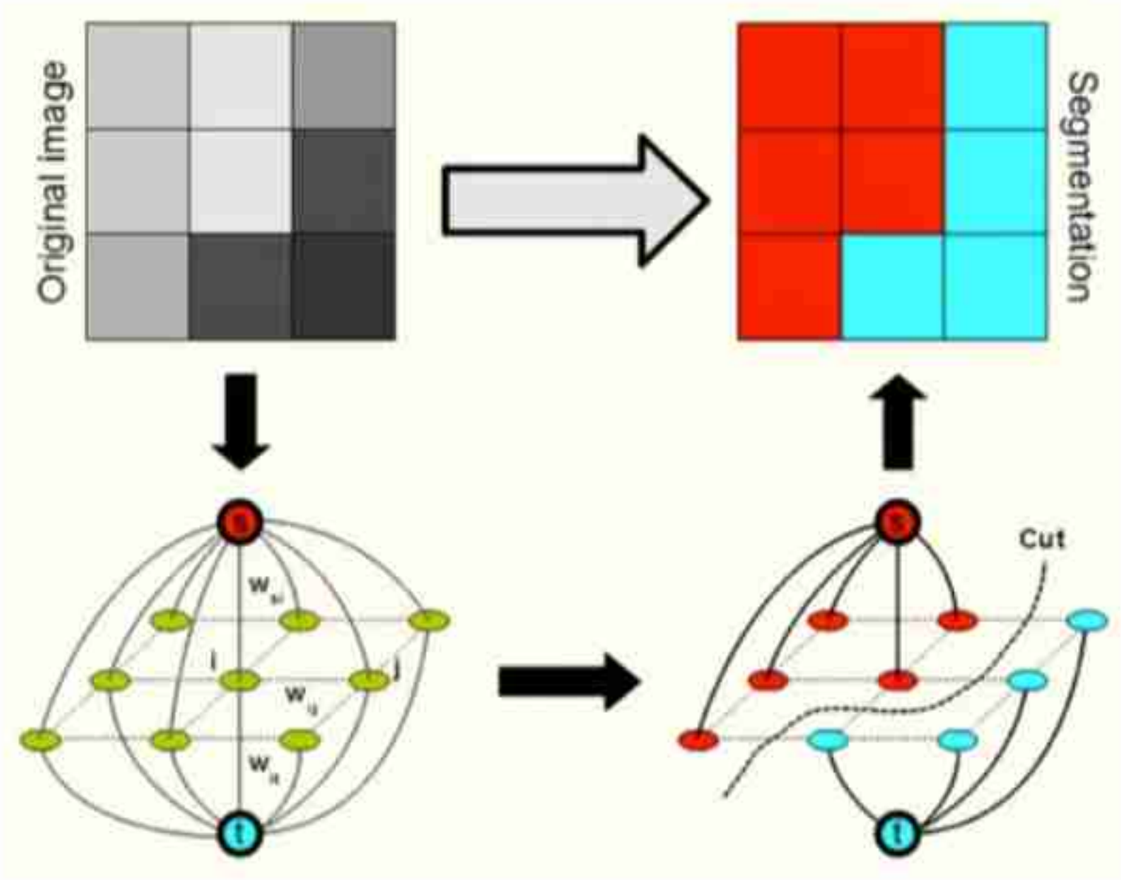



B
G
Az: 100 El: 26

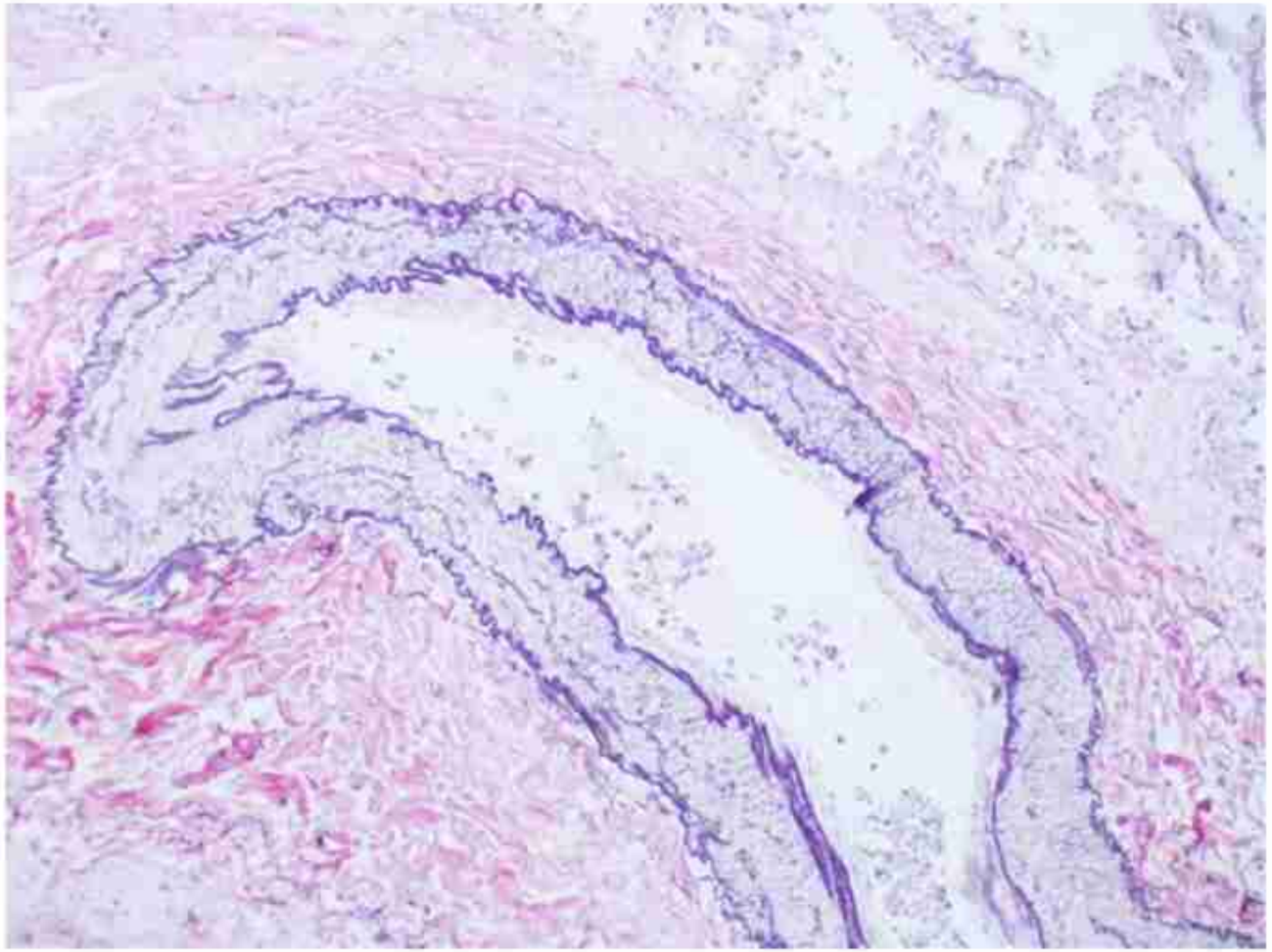

B
R
G
Az: 153 El: 17


RGB
HSV
YCbCr
L*a*b*

# Image as Graph

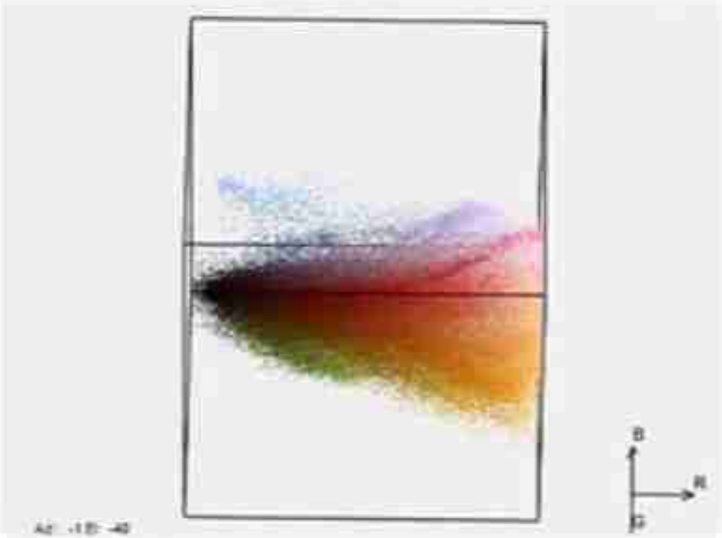


Graph Cuts and Efficient N-D Image Segmentation - Dr. Boykov and Dr. Kolmogorov

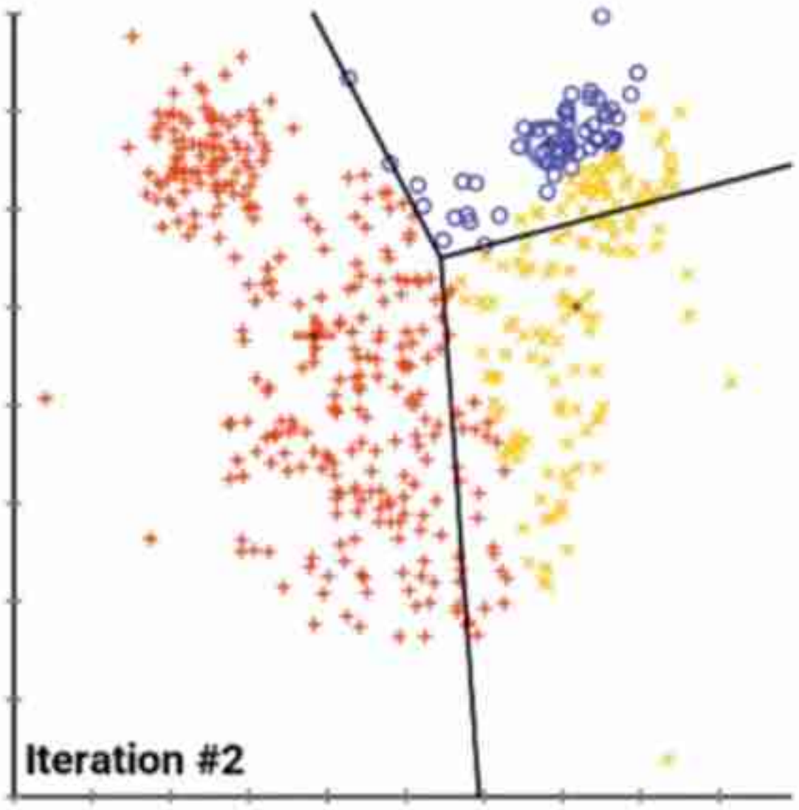


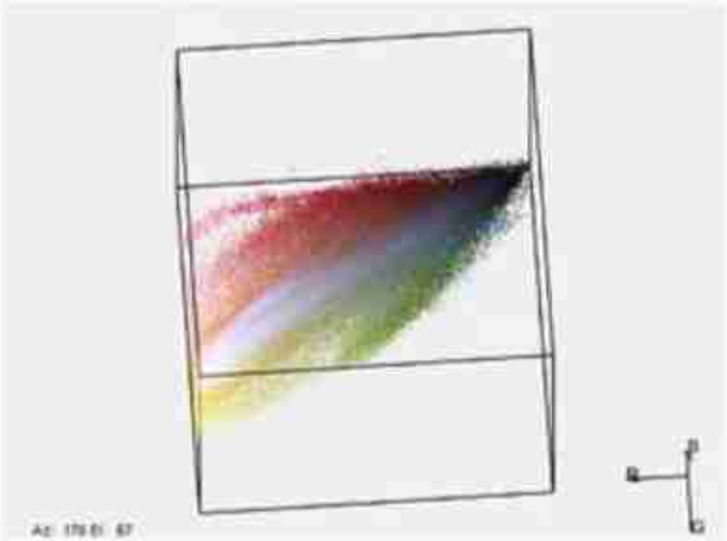
Pulmonary Artery

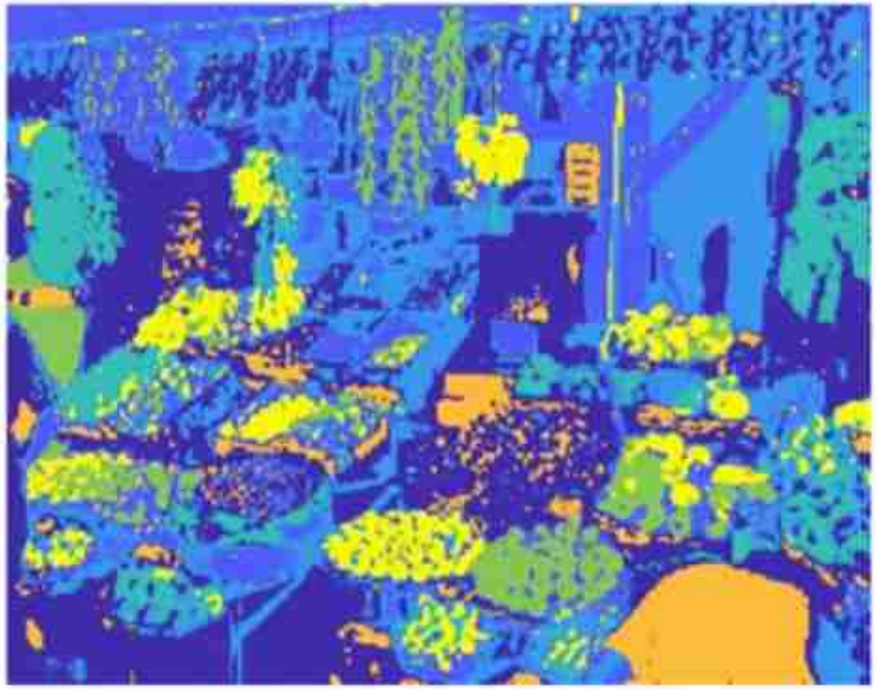
Active Contours
Share
3:50 / 7:14


B
R
G

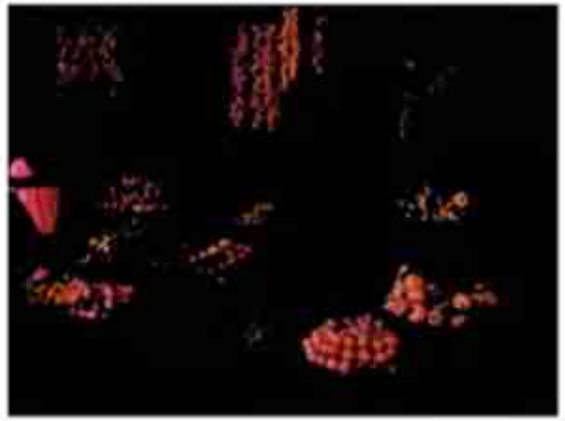
# K-Means Clustering

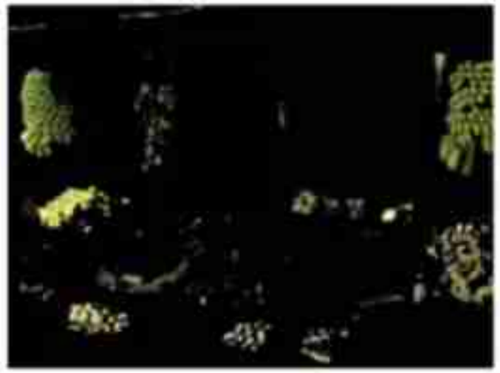


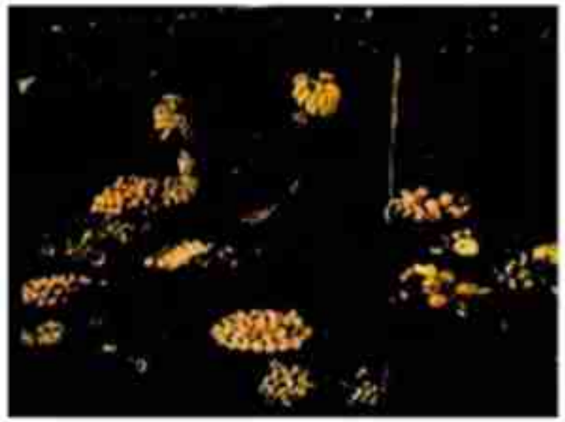



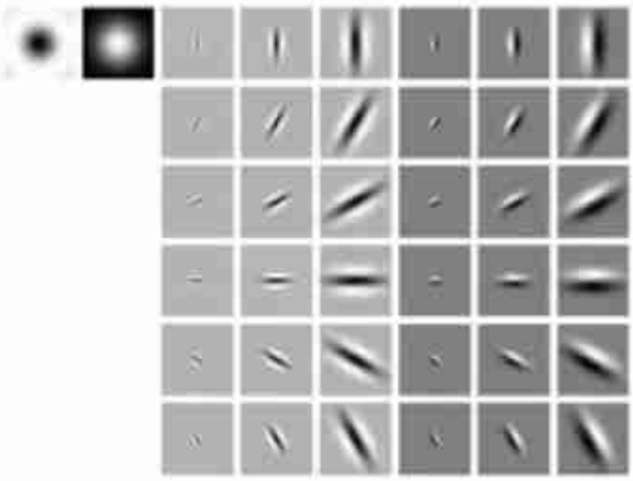


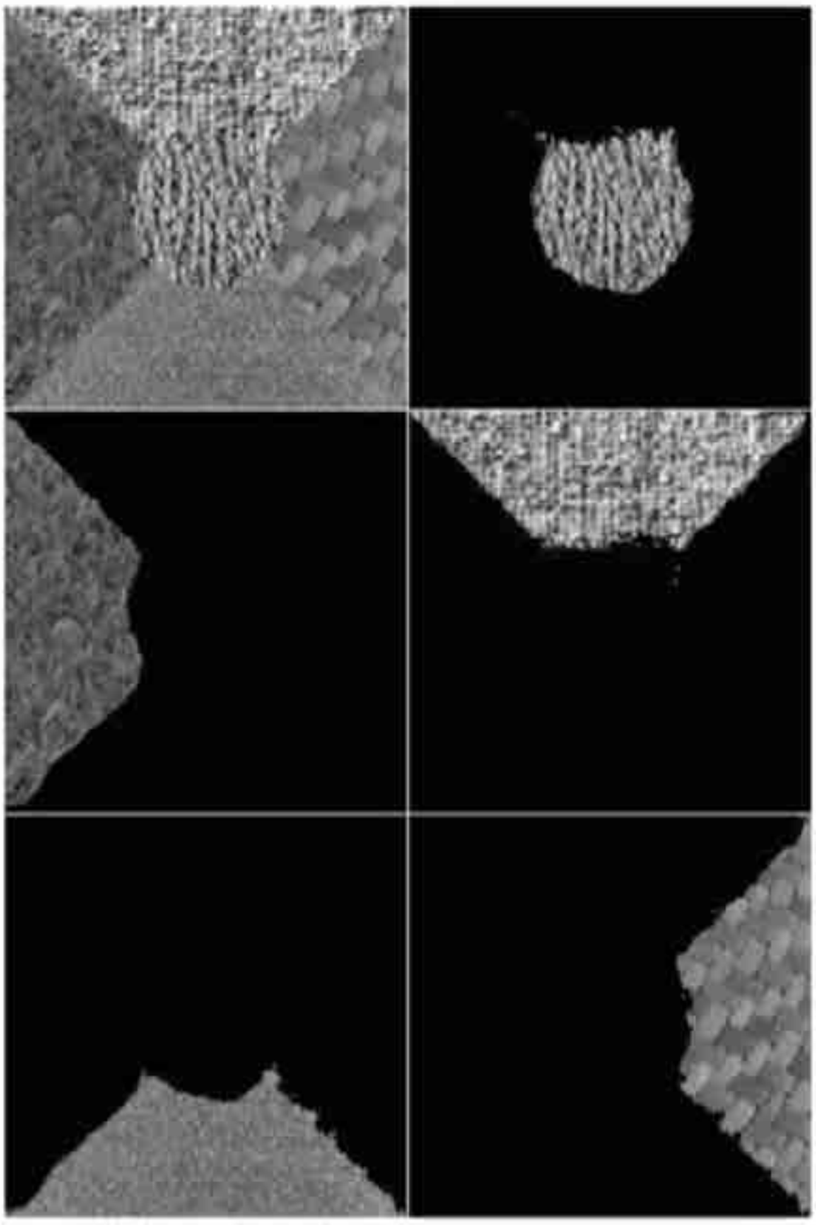
K-Means Clustering
K = 7

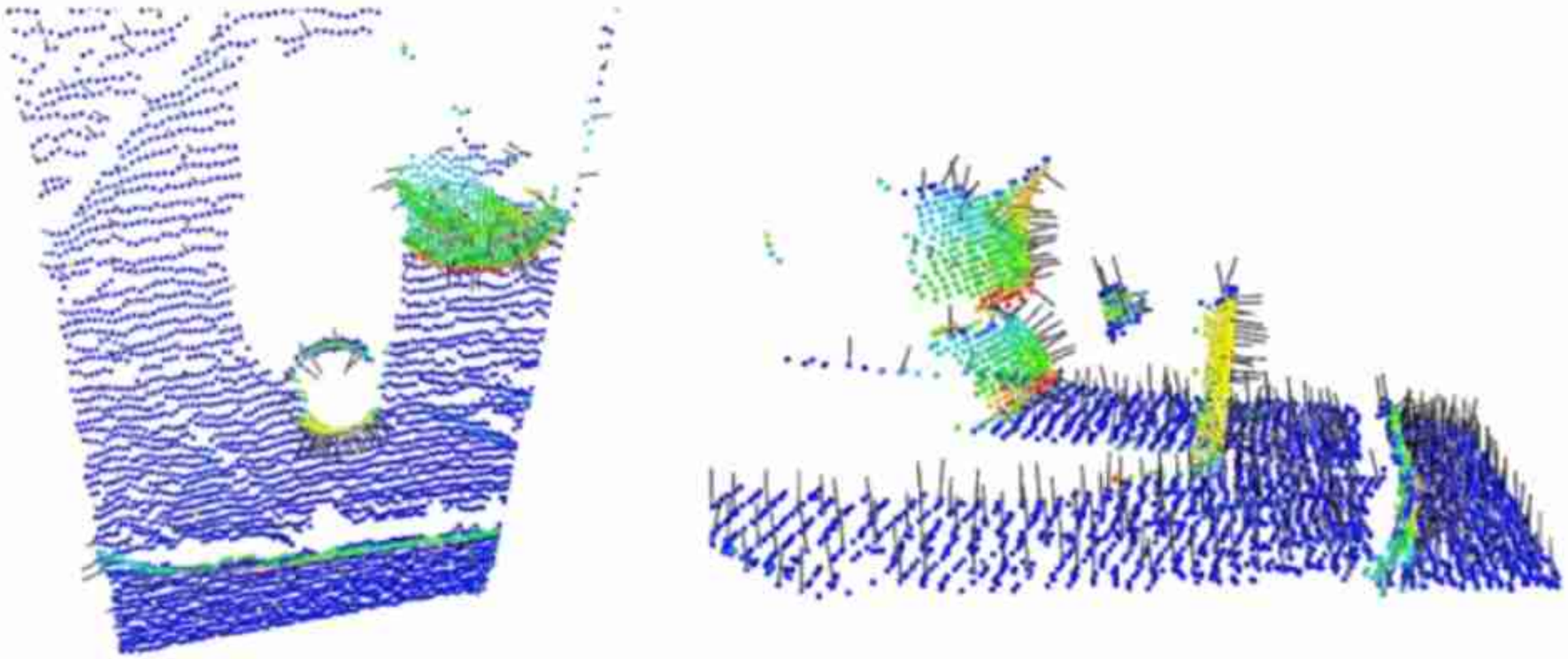












Source: Real-Time Plane Segmentation using RGB-D Cameras - Dirk Holz, StefanHolzer, Radu Bogdan Rusu and Sven Behnke

Source: Real-Time Plane Segmentation using RGB-D Cameras - Dirk Holz, StefanHolzer, Radu Bogdan Rusu and Sven Behnke

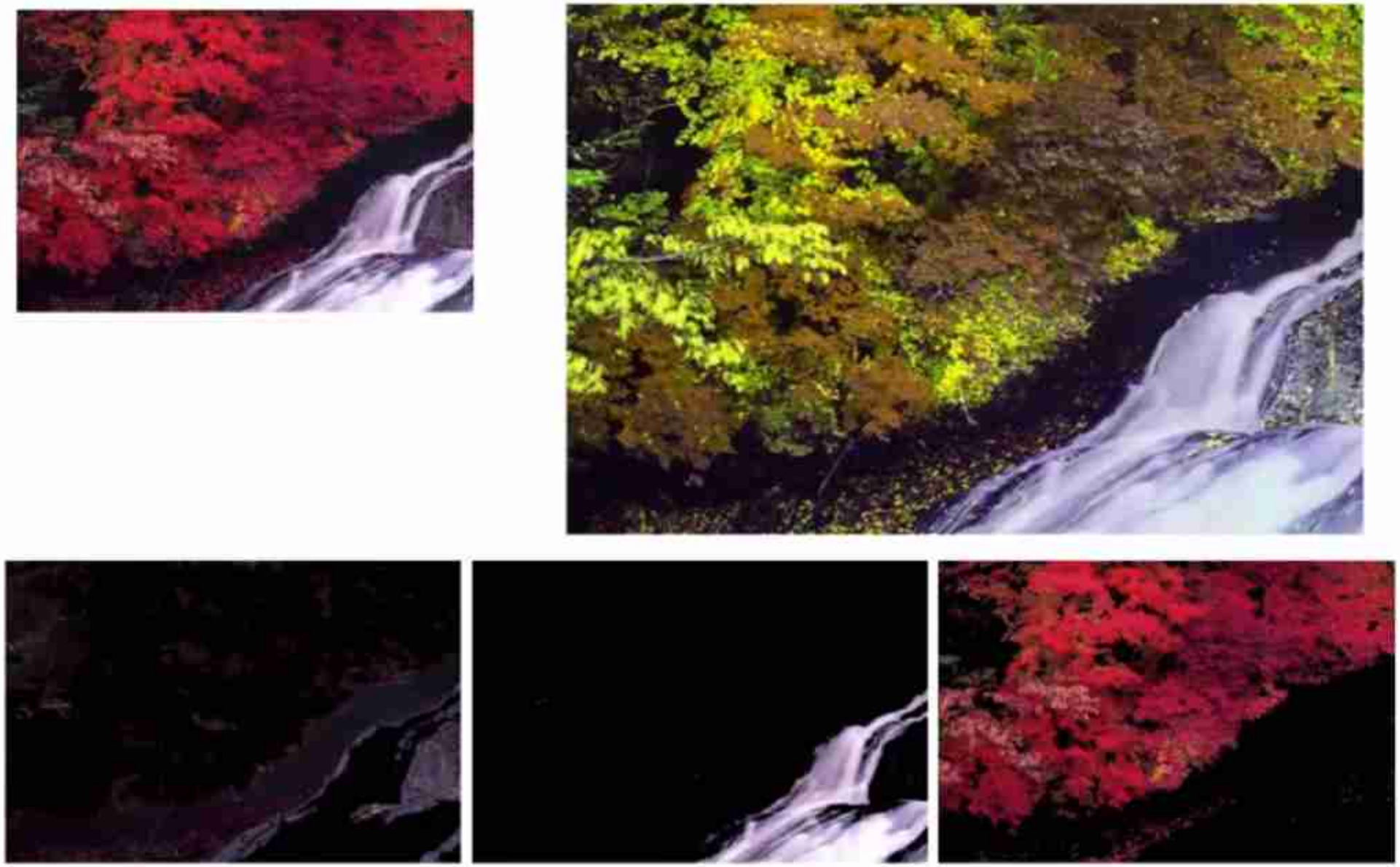
# Soft Segmentation

Source: Semantic Soft Segmentation, Yagiz Aksoy, Tae-Hyun Oh, Sylvain Paris, Marc Pollefeys and Wojciech Matusik

Share
5:50 / 7:14

# Soft Segmentation

Source: Semantic Soft Segmentation, Yagiz Aksoy, Tae-Hyun Oh, Sylvain Paris, Marc Pollefeys and Wojciech Matusik



Local Color Transfer via Probabilistic Segmentation by Expectation-Maximization
Yu-Wing Tai, Jiaya Jia, Chi-Keung Tang

Share

6:06 / 7:14

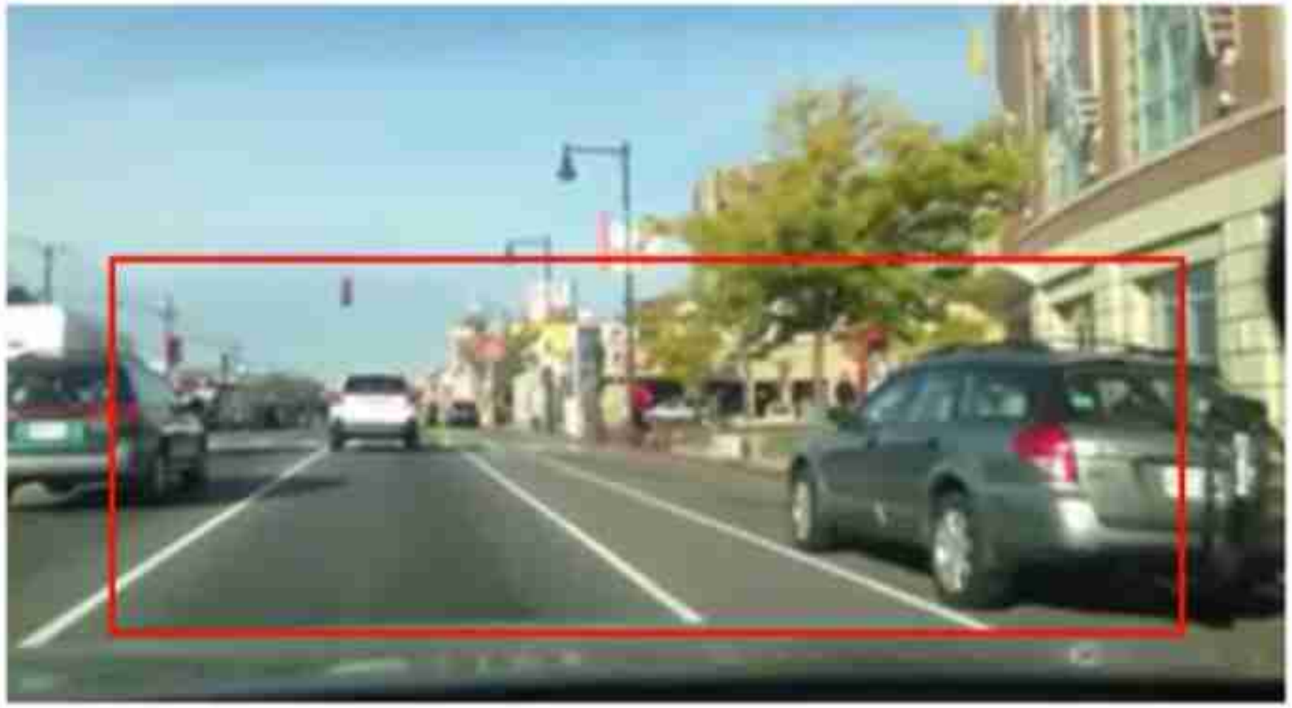
MATLAB - Computer Vision Toolbox

MATLAB - Computer Vision Toolbox

MATLAB - Computer Vision Toolbox

MATLAB - Computer Vision Toolbox

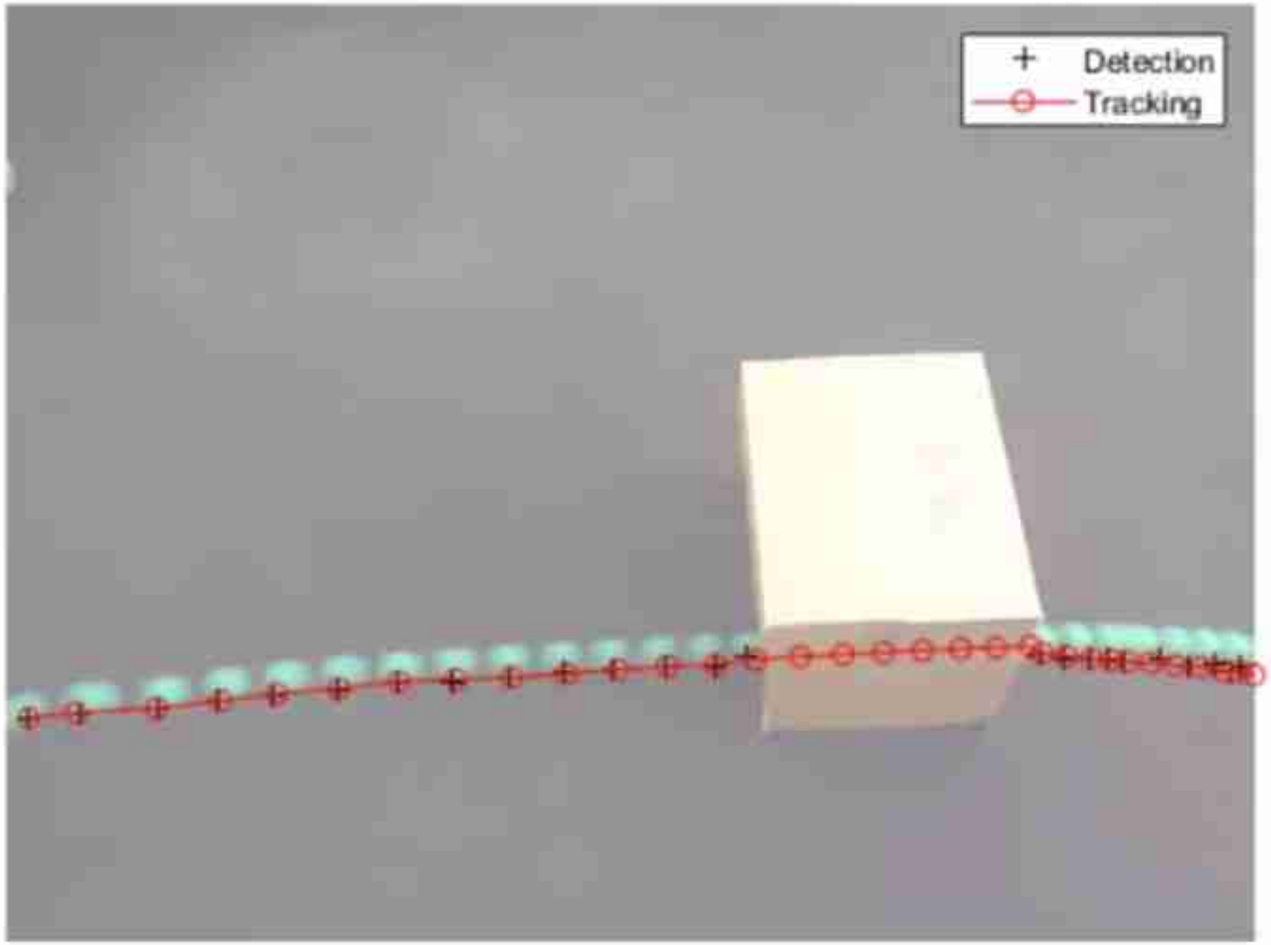
MATLAB - Computer Vision Toolbox



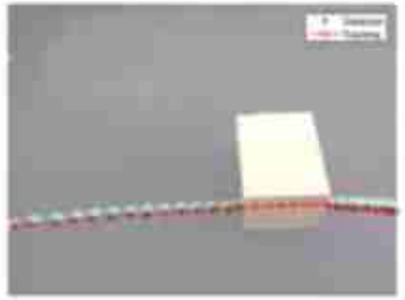
MATLAB - Computer Vision Toolbox

MATLAB - Computer Vision Toolbox

MATLAB - Computer Vision Toolbox

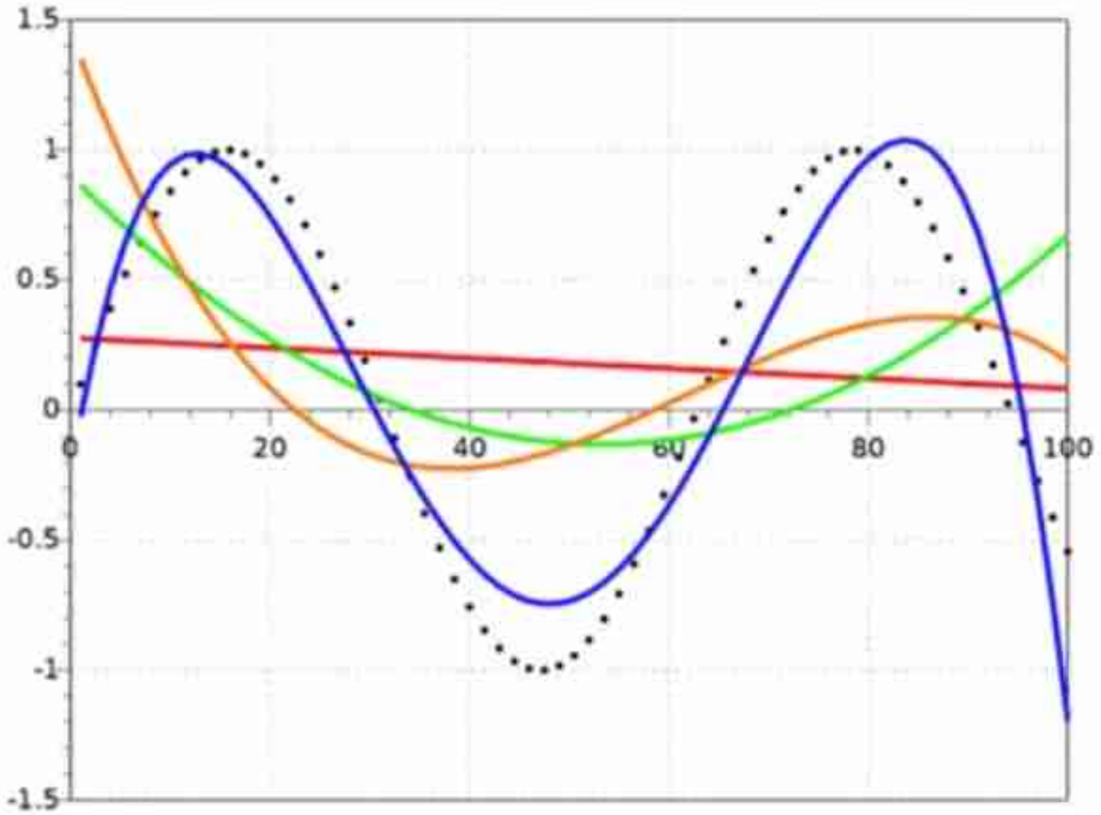



MATLAB - Computer Vision Toolbox



MATLAB - Computer Vision Toolbox

# Model Fitting



10
cents
20
TWENTY CENTS

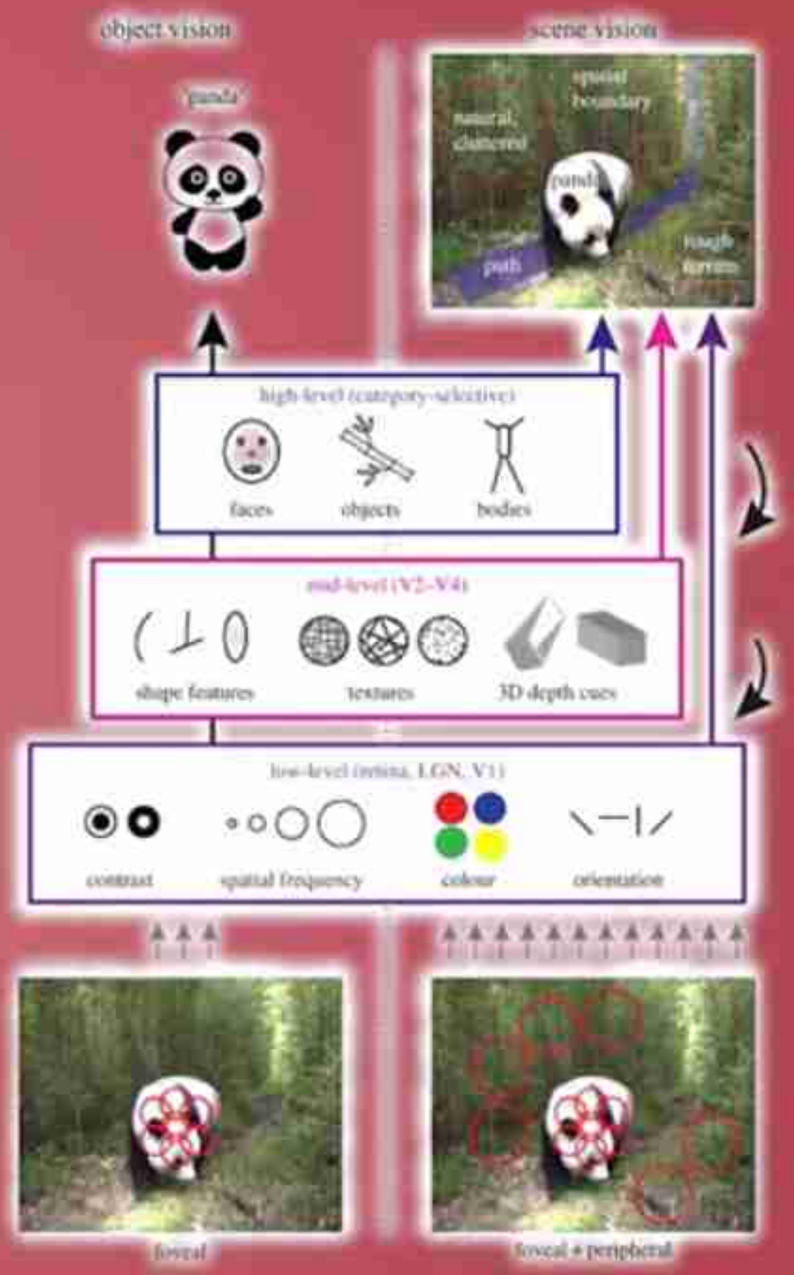
10
cents
20

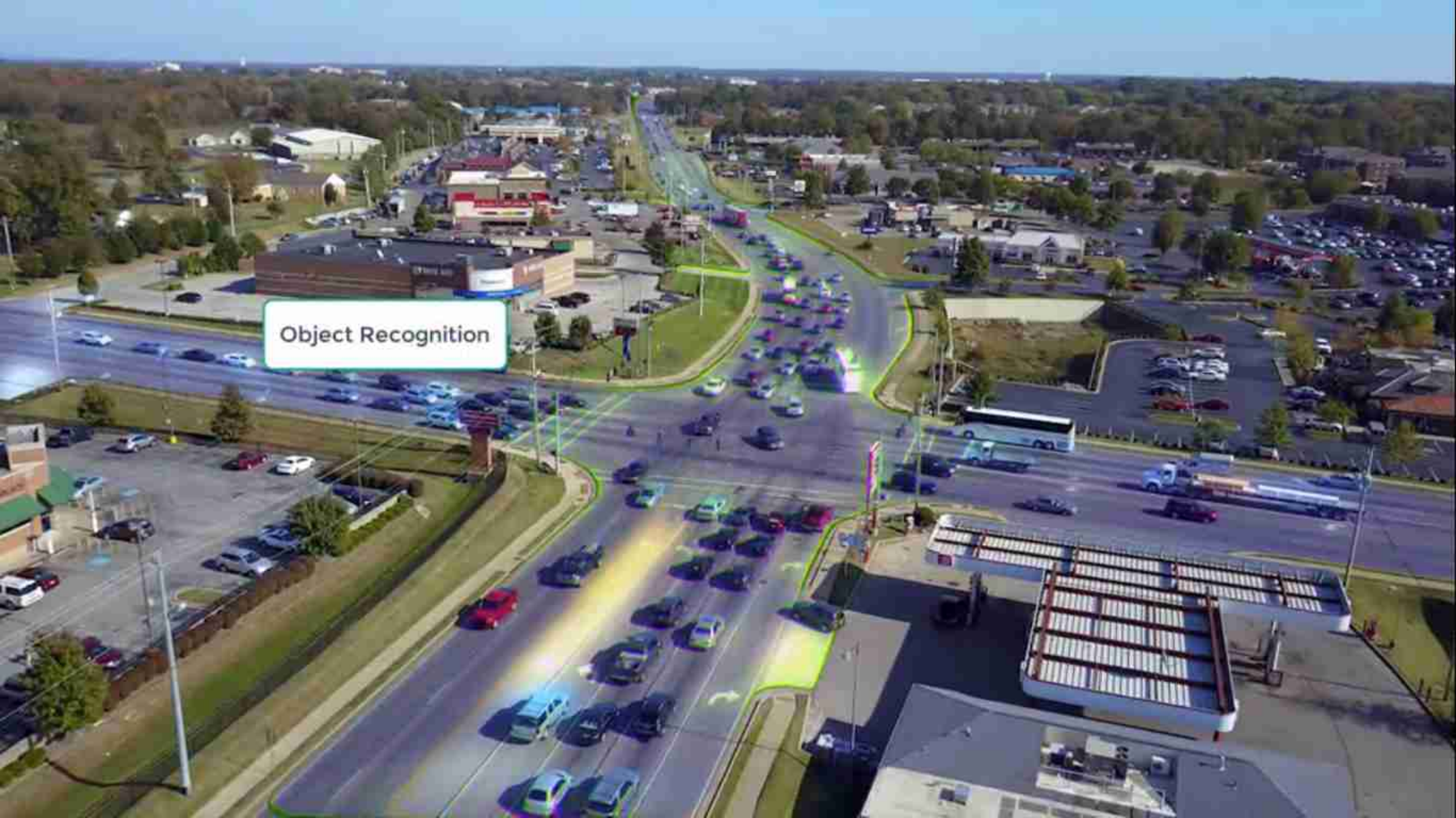
# High Level Vision

# High Level Vision

-It is the final stage in identifying objects

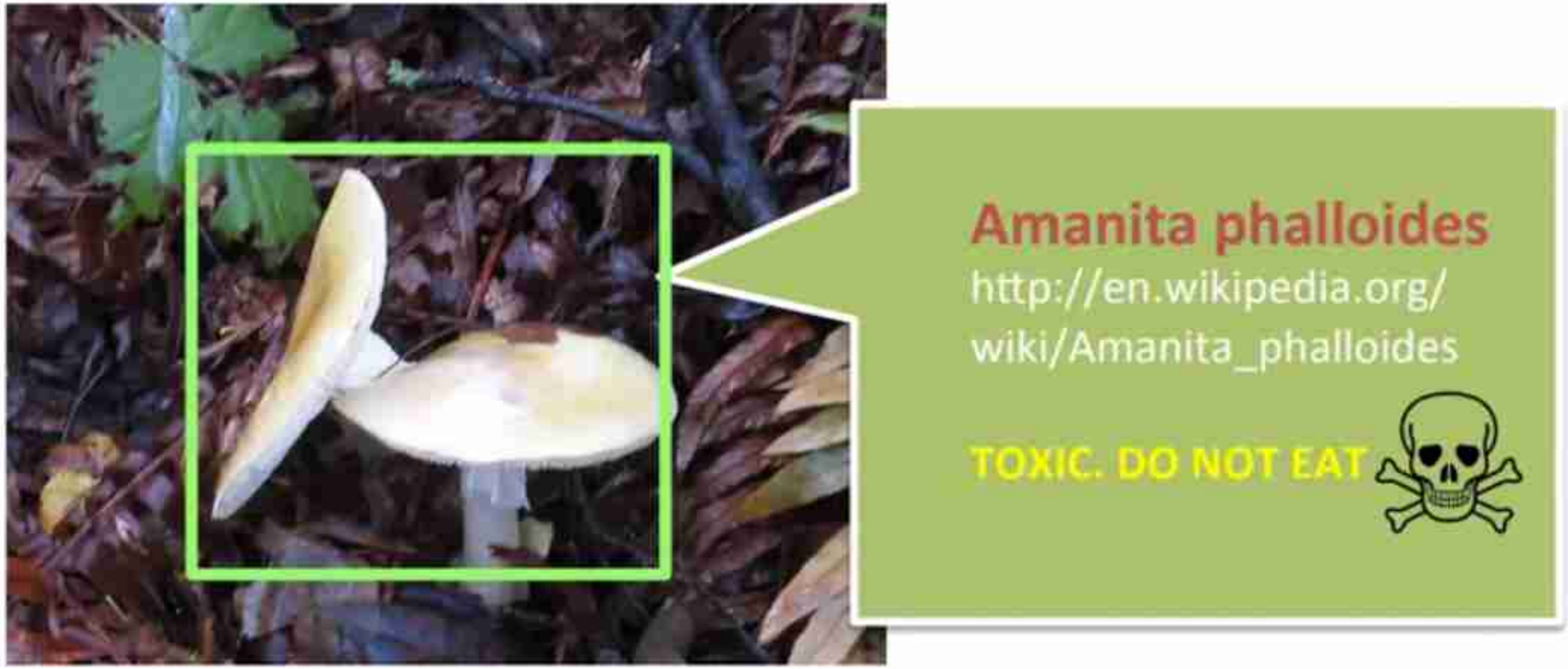
-It will identify the difference between a car and a road, or cat vs a dog




Object Recognition

120
MPB

Credit: Dr. Fei Fei Li




Credit: Dr. Fei Fei Li

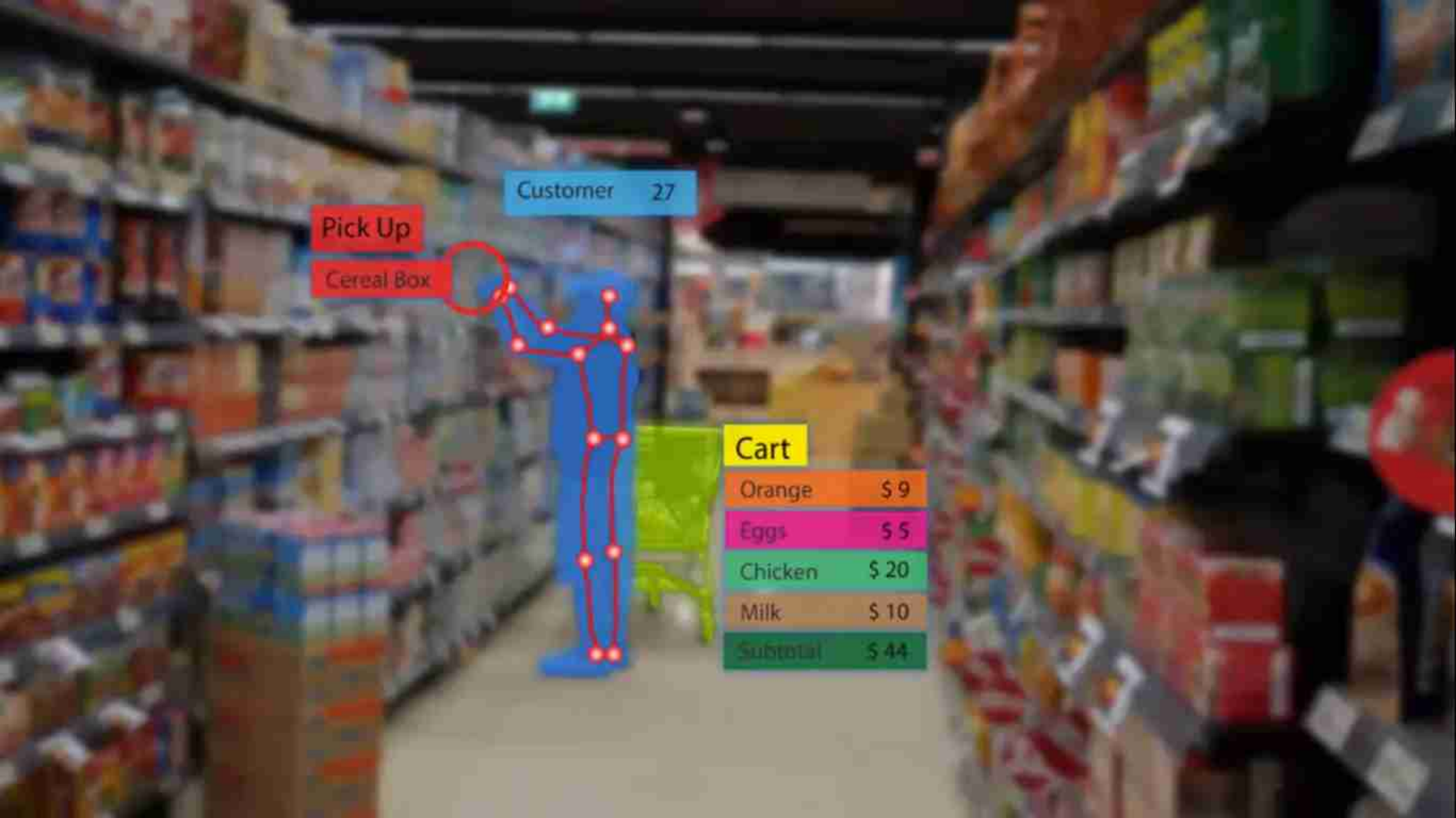
**Mountain Lion**

**DO NOT RUN**

Raise arms to appear larger.
Show your teeth

Credit: Dr. Fei Fei Li

Credit: Dr. Fei Fei Li


Customer 27
Pick Up
Cereal Box
Cart
Orange $ 9
Eggs $ 5
Chicken $ 20
Milk $ 10
Subtotal $ 44

Signal : OK
Human
Moving vehicle
40 km/h
Human

Signal : OK
Moving vehicle
40 km/h
Human

# Visual Recognition

## Classification: Does this video contain people?

# Visual Recognition

## Classification: Is this a football ground?

# Visual Recognition

## Detection: Where is the bicycle in this video?

# Visual Recognition

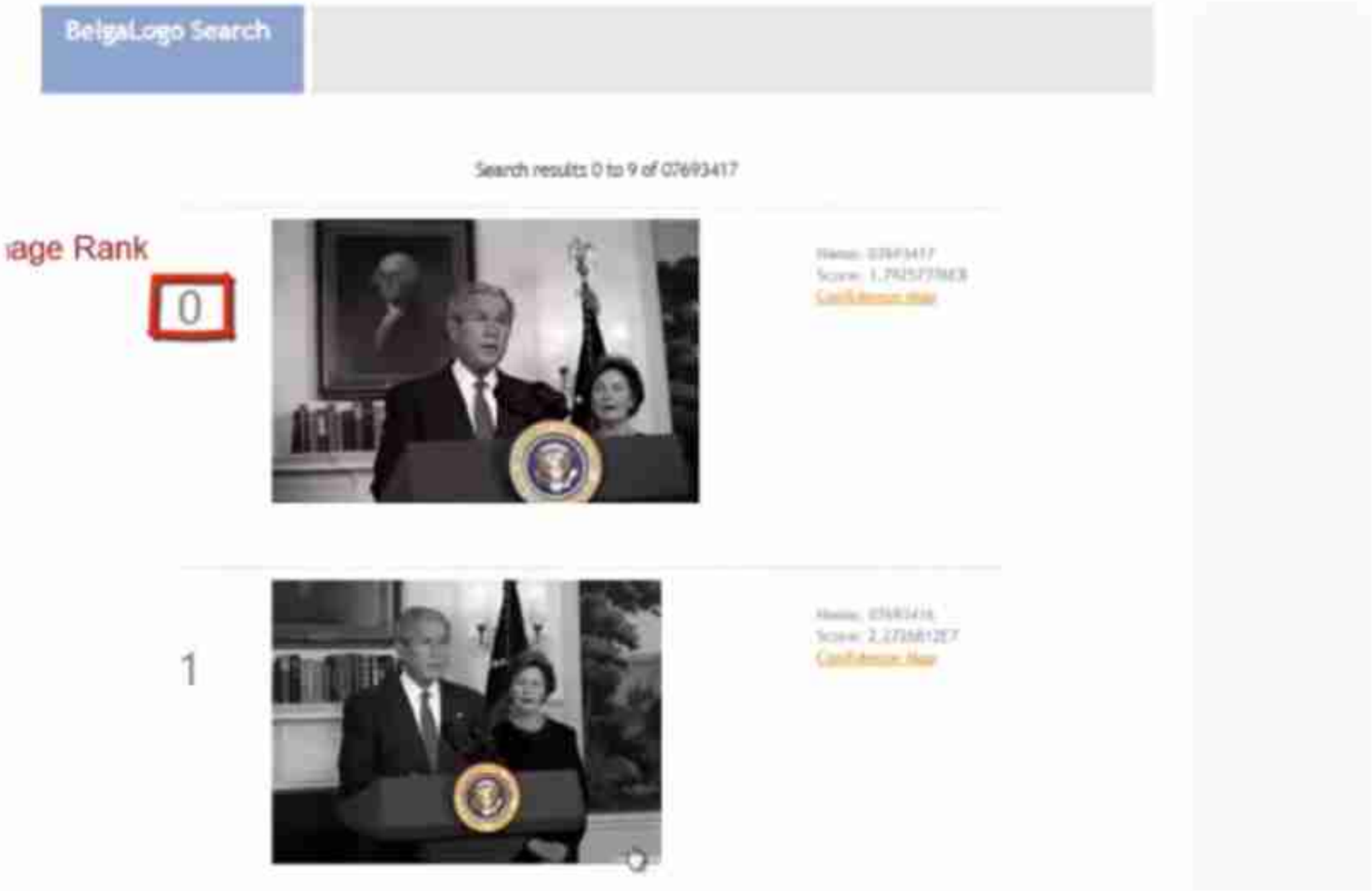
## Detection: Where is the bicycle in this video?

Share
1:47 / 6:51

# Visual Recognition

Detection: What objects does this Image contain?

# Visual Recognition

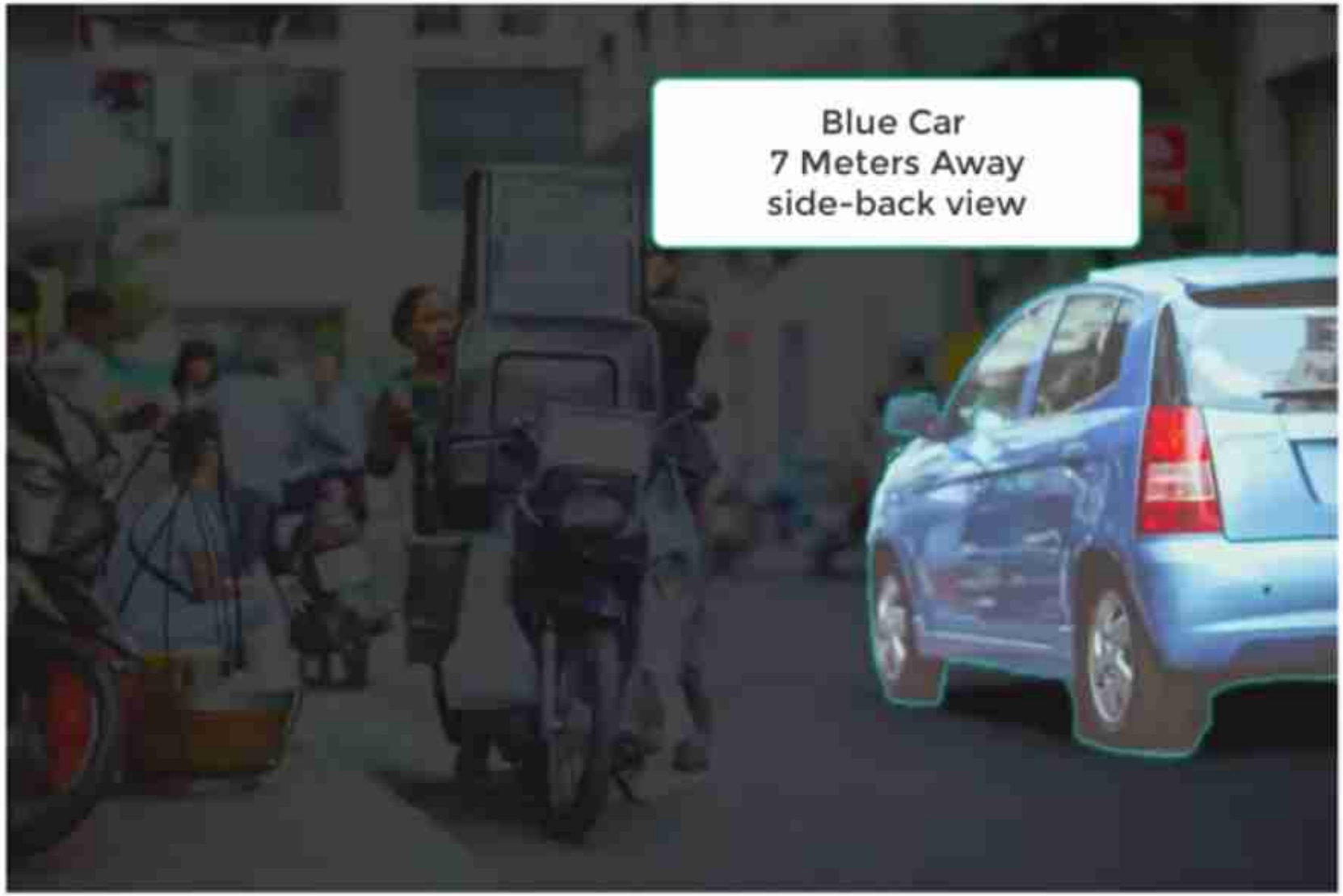


# Visual Recognition

## Semantic Segmentation - Accurate Localization and Detection

# Visual Recognition

## Semantic Segmentation - Accurate Localization and Detection



# Visual Recognition

## Semantic Segmentation - Accurate Localization and Detection

Building

# Visual Recognition

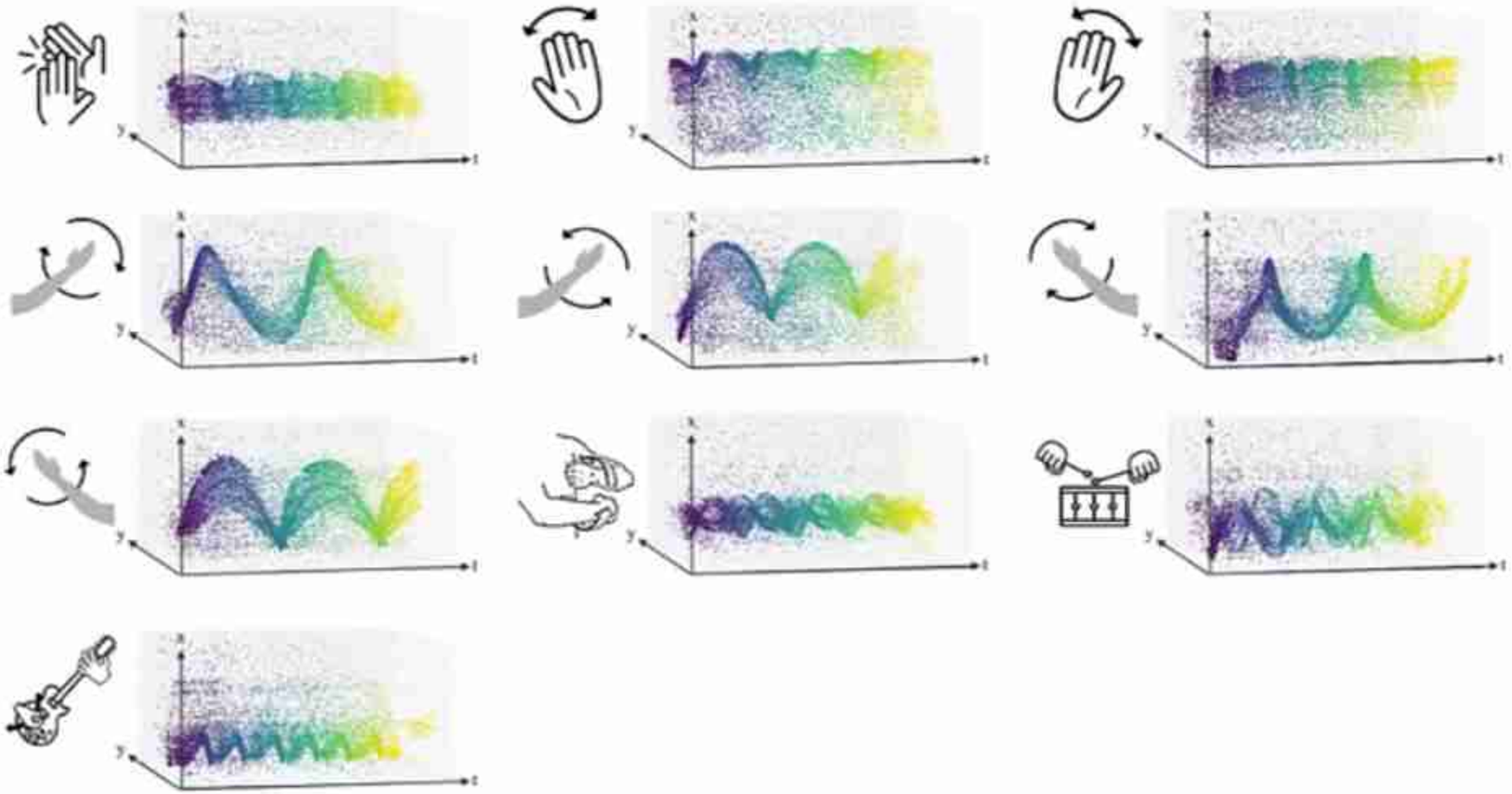
## Classification vs Single Instance Recognition

# Visual Recognition

What are these people doing?

# Visual Recognition

## Action Recognition



# Visual Recognition

# Visual Recognition

## Event Recognition

# Visual Recognition

Design Algorithms that are capable of

Classification of Images/Videos

Detect and Localize Objects

Estimate semantic and geometric attributes

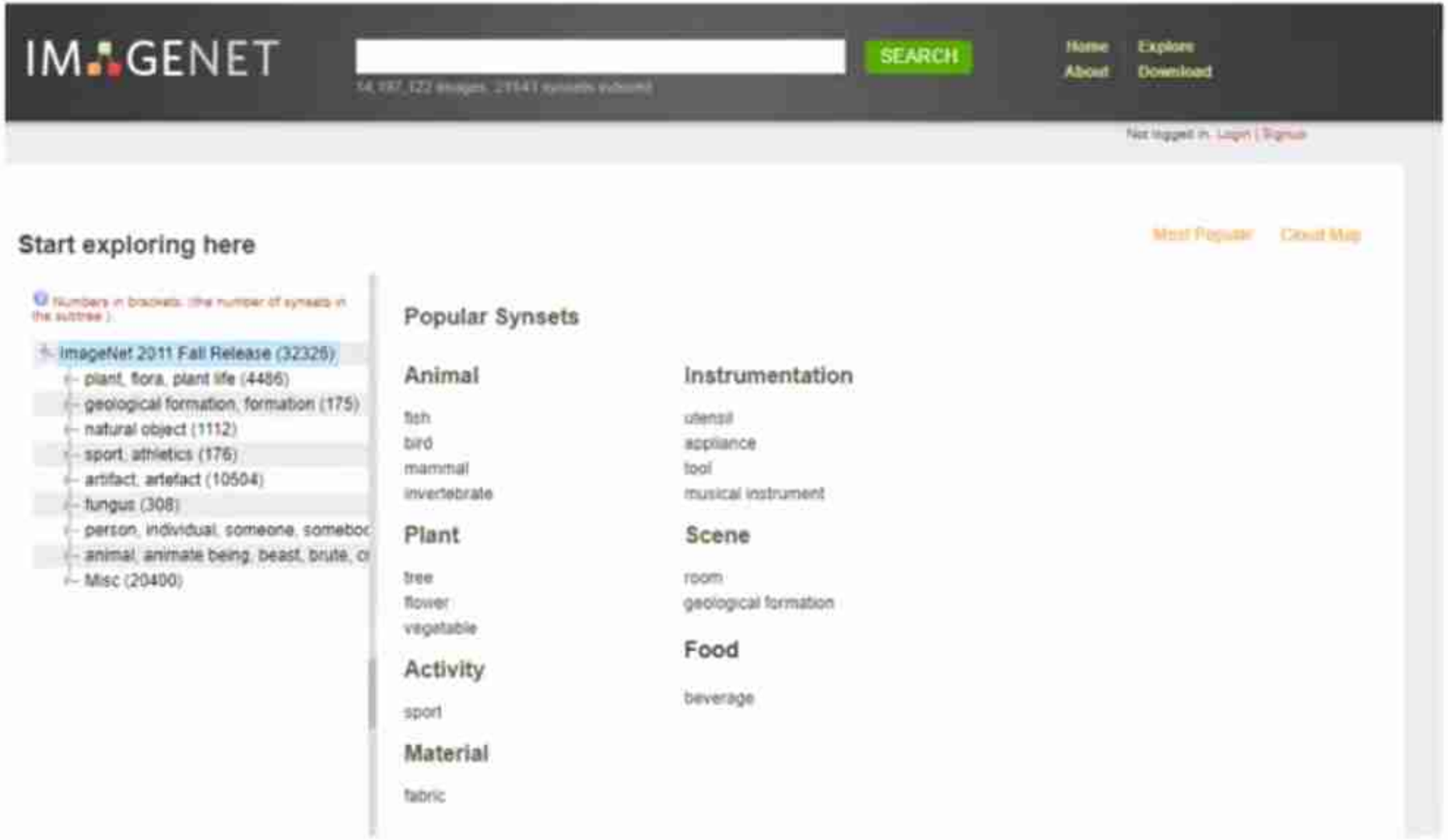
Classify human activities and videos

Credit: S Savarese.

Share

3:52 / 6:51

# Visual Recognition



# Visual Recognition

IMAGENET

SEARCH

14,197,122 images, 21841 synsets indexed

## ~30000 Object Categories

### Start exploring here

Numbers in brackets: (the number of synsets in the subtree ).

- ImageNet 2011 Fall Release (32326)
  - plant, flora, plant life (4486)
  - geological formation, formation (175)
  - natural object (1112)
  - sport, athletics (176)
  - artifact, artefact (10504)
  - fungus (308)
  - person, individual, someone, someboc

### Popular Synsets

**Animal**

fish
bird
mammal
invertebrate

**Instrumentation**

utensil
appliance
tool
musical instrument

**Plant**

**Scene**

# Visual Recognition

# Visual Recognition

## View point variation

# Visual Recognition

## Illumination

# Visual Recognition

## Illumination

# Visual Recognition

## Illumination

# Visual Recognition

## Scale variation

# Visual Recognition

## Deformation

# Visual Recognition

## Occlusion

# Visual Recognition

Press Esc to exit full screen

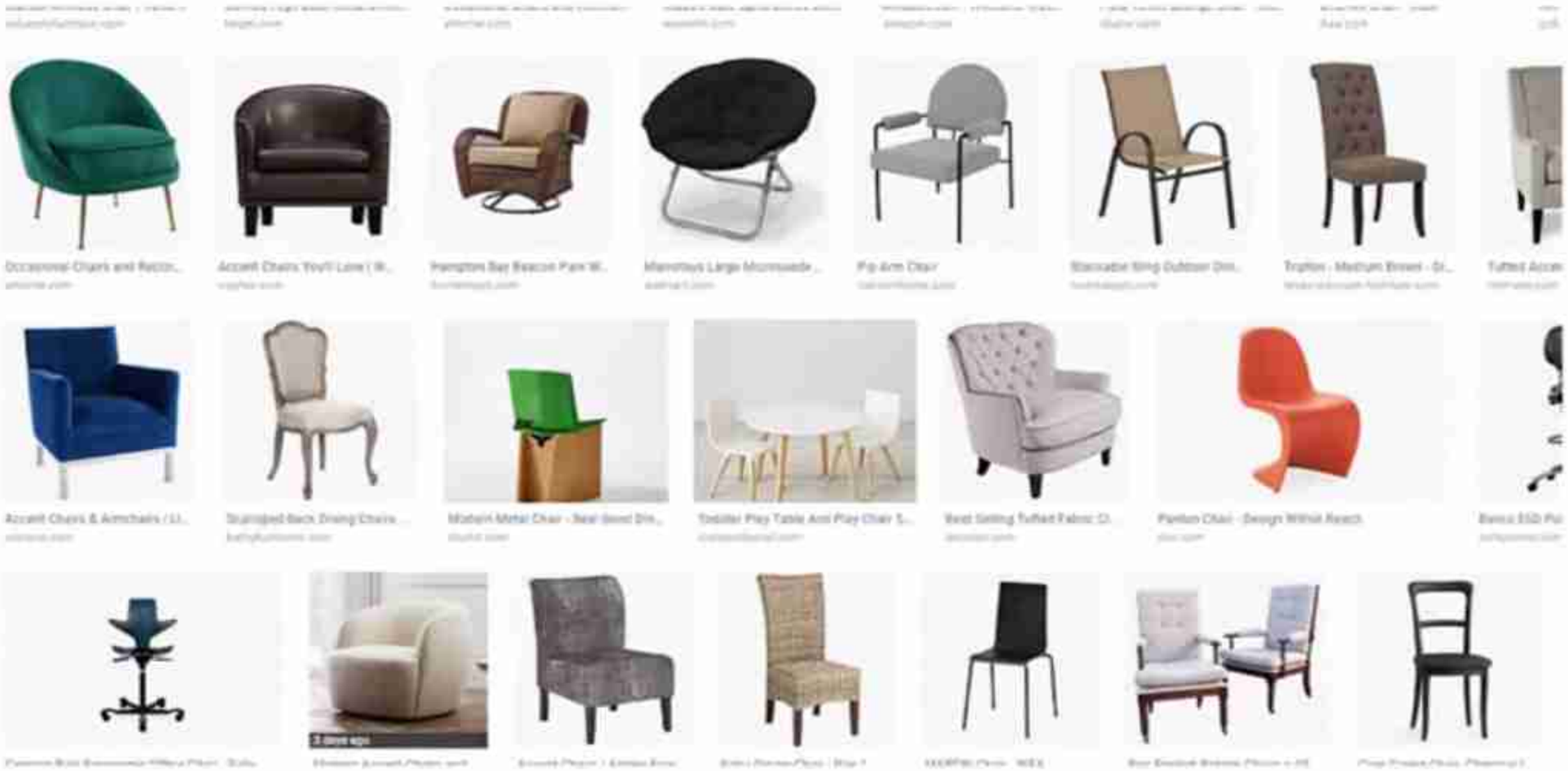
Share

## Background/Clutter

4:52 / 6:51

# Visual Recognition

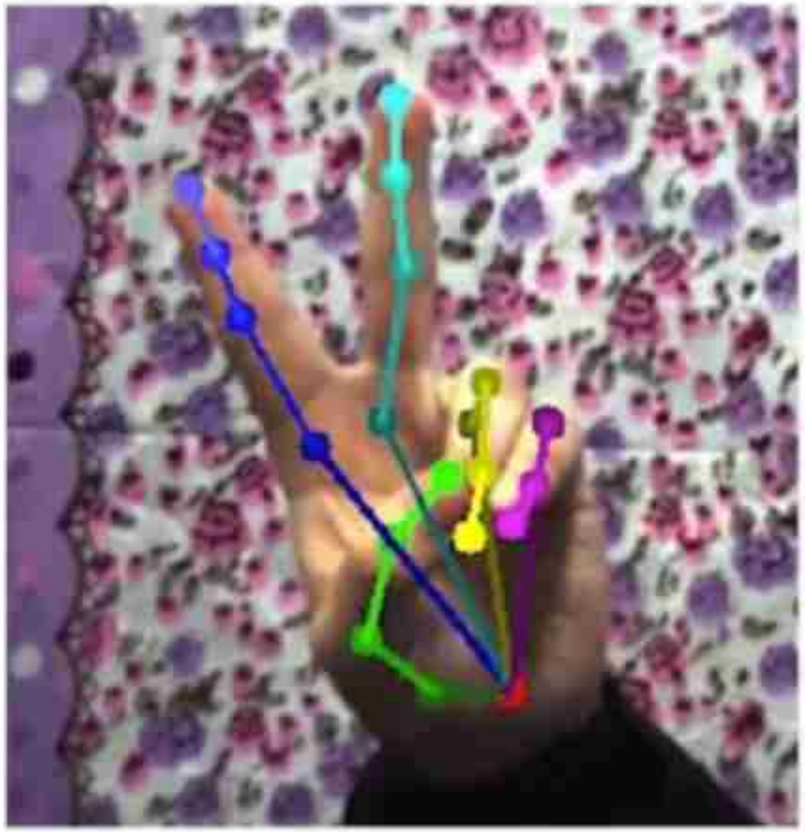
## Intra-Class Variation

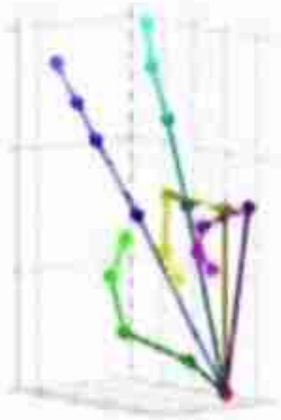


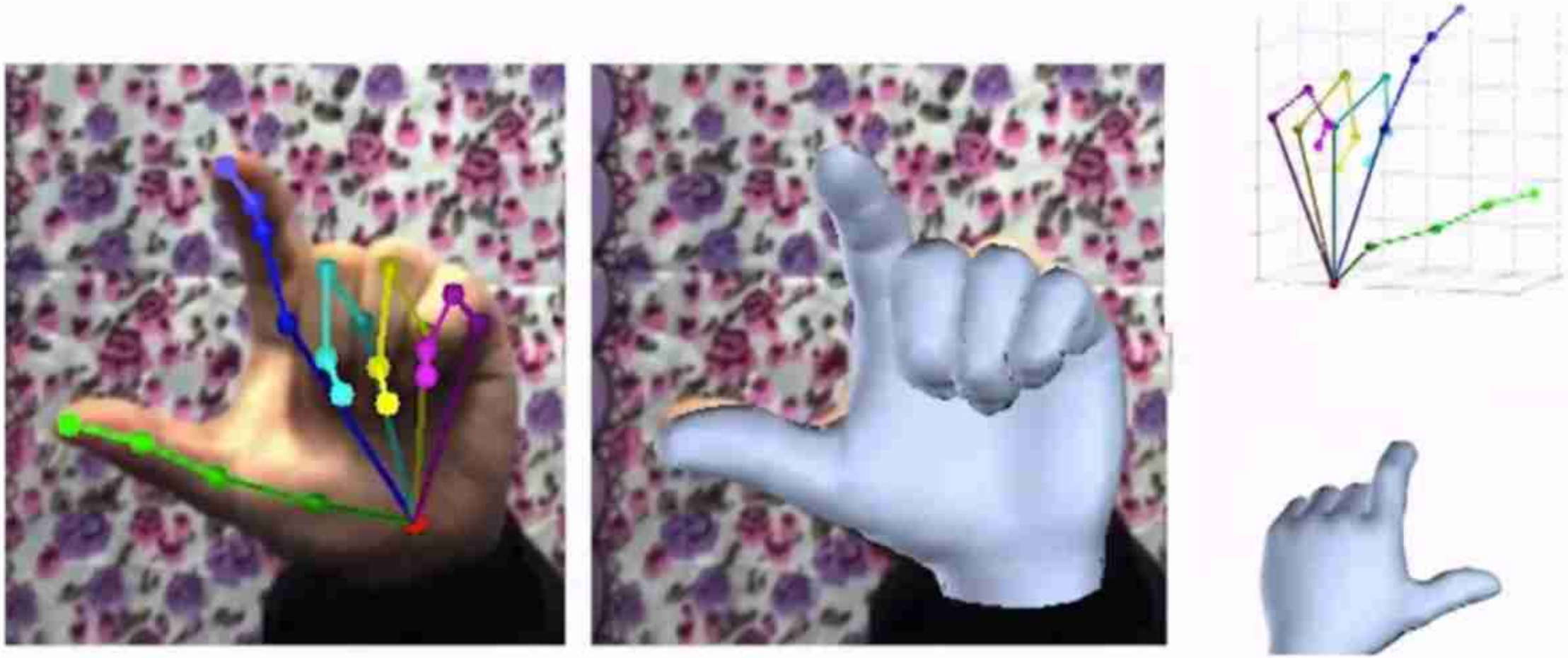
# Two Approaches - Recognition

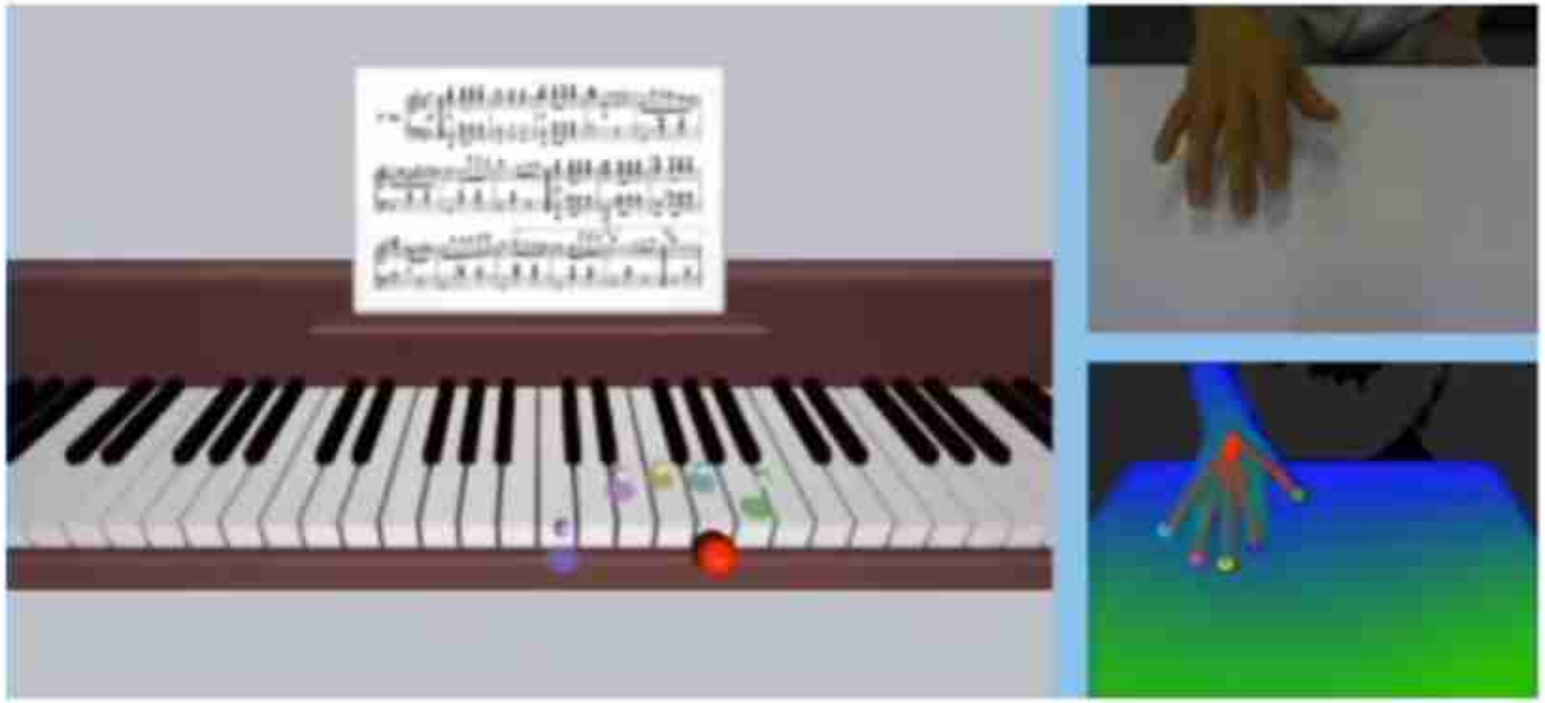
model-based recognition

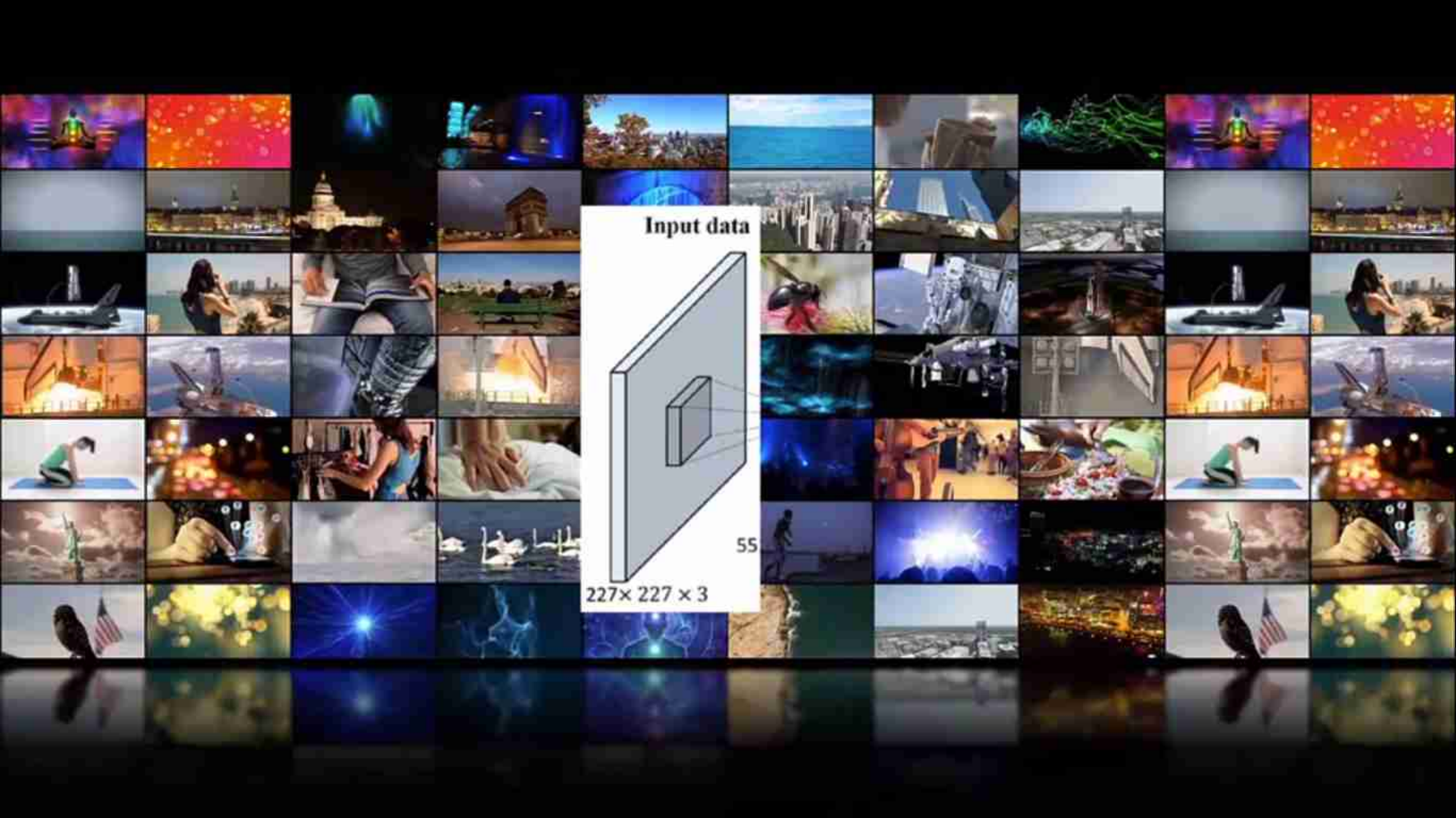
learning-based recognition

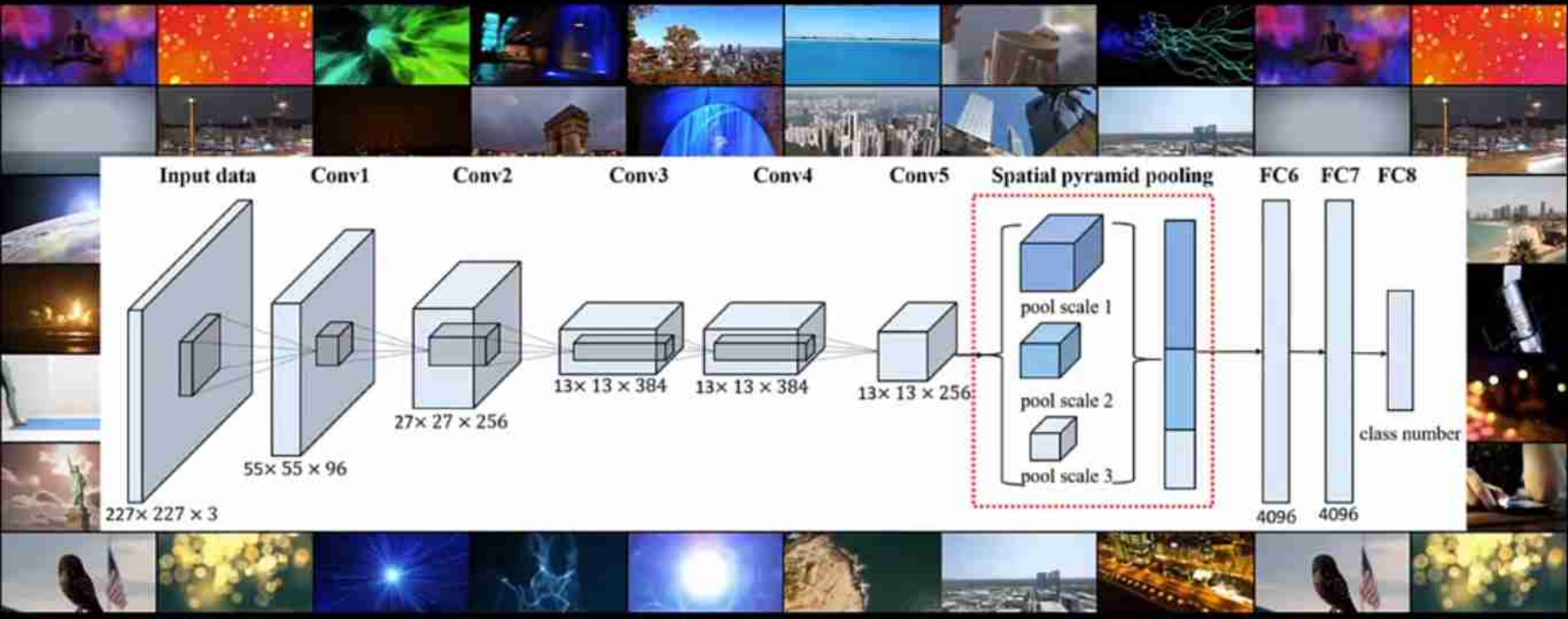









Input data
55
227× 227 × 3


Input data
Conv1
Conv2
Conv3
Conv4
Conv5
Spatial pyramid pooling
FC6
FC7
FC8
227× 227 × 3
55× 55 × 96
27× 27 × 256
13× 13 × 384
13× 13 × 384
13× 13 × 256
pool scale 1
pool scale 2
pool scale 3
4096
4096
class number

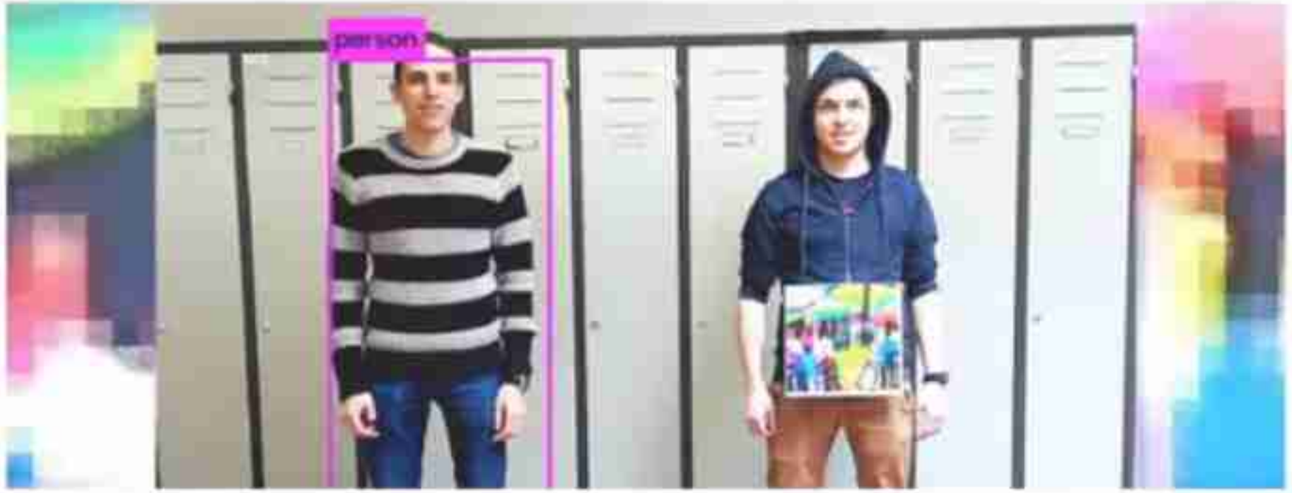
(a) Husky classified as wolf

(b) Explanation

MIT Technology Review

Artificial Intelligence Apr 23

## How to hide from the AI surveillance state with a color printout

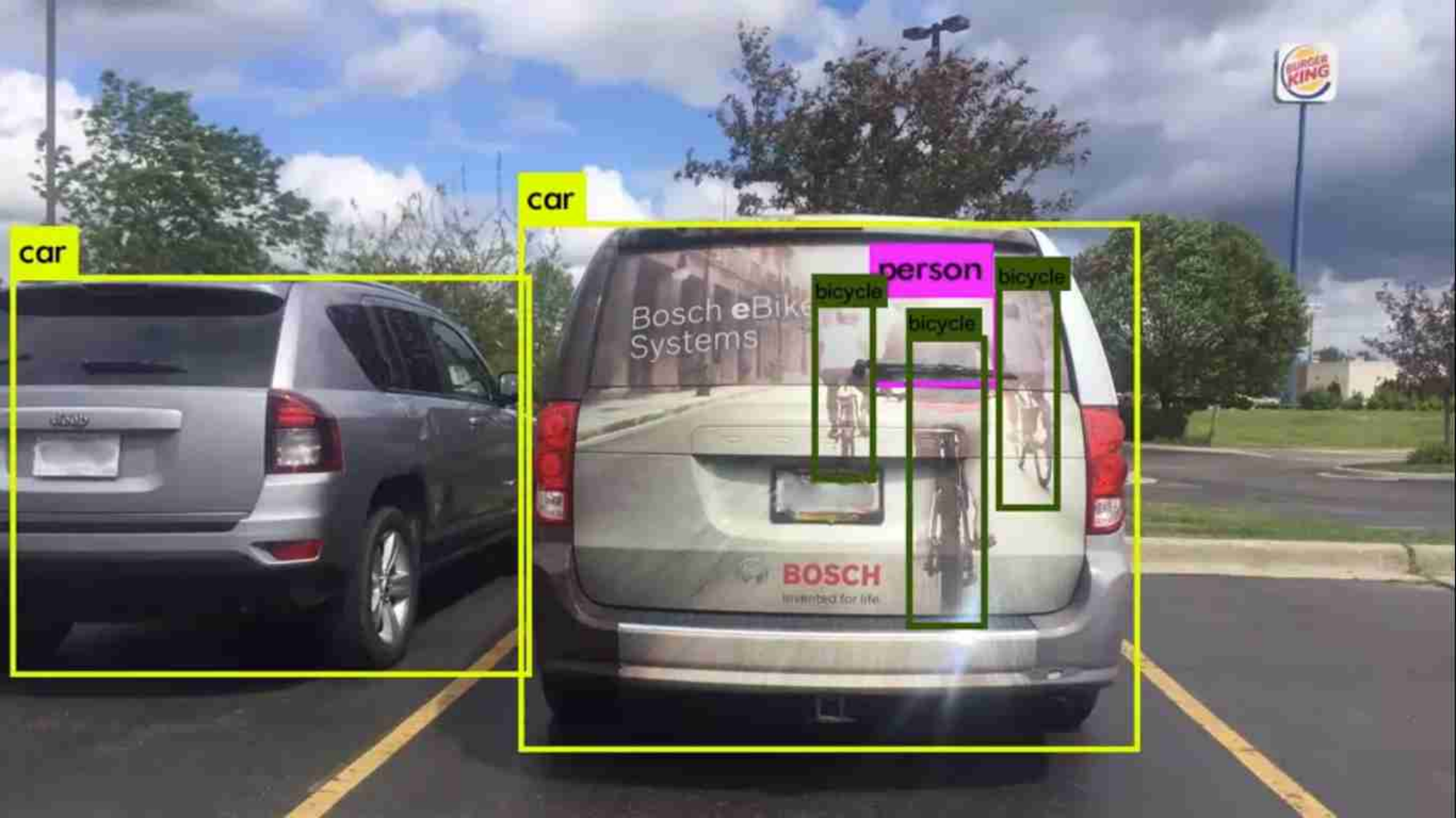


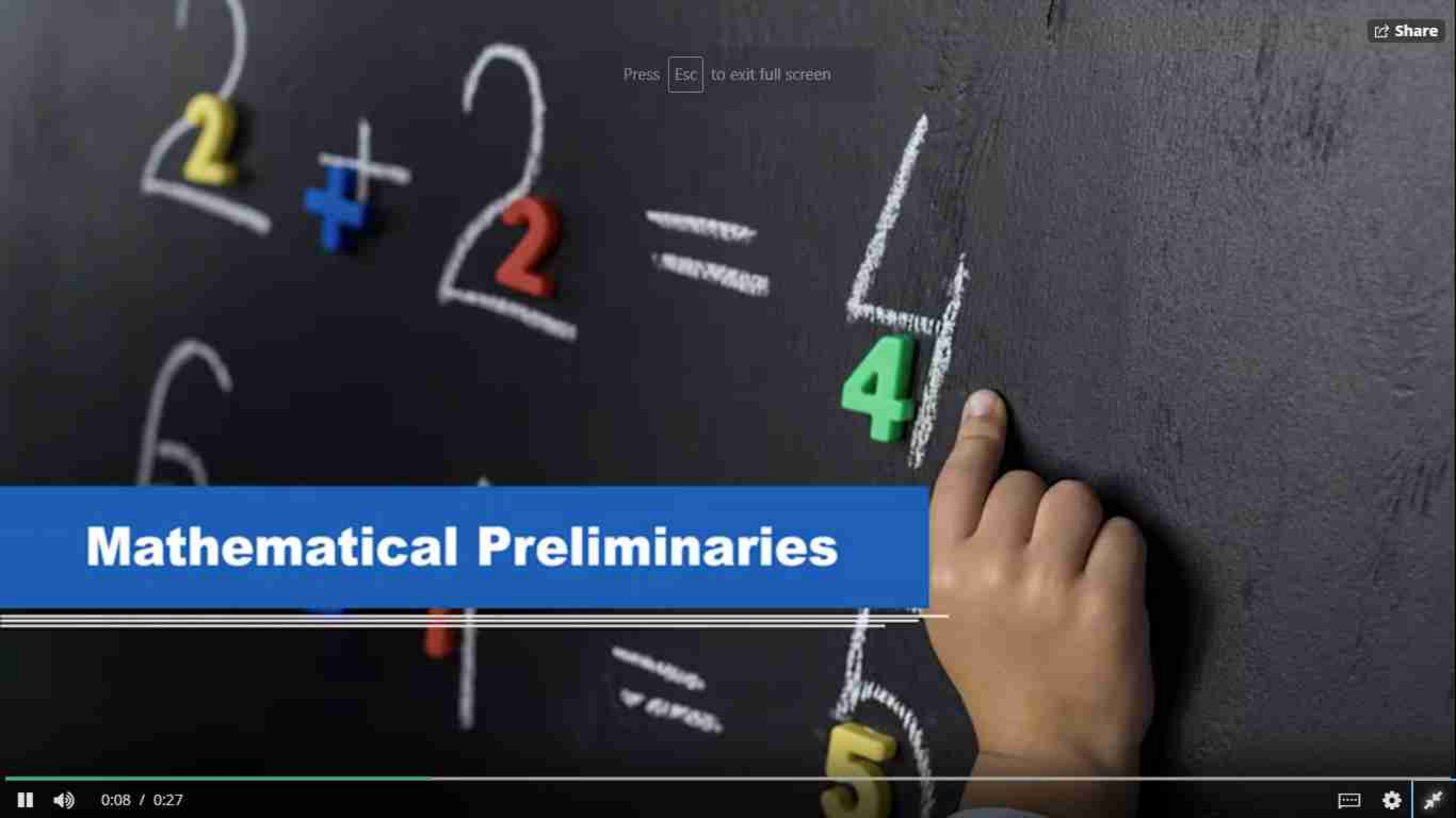

Fooling automated surveillance cameras: adversarial patches to attack person detection
Simen Thys, Wiebe Van Ranst, Toon Goedemé

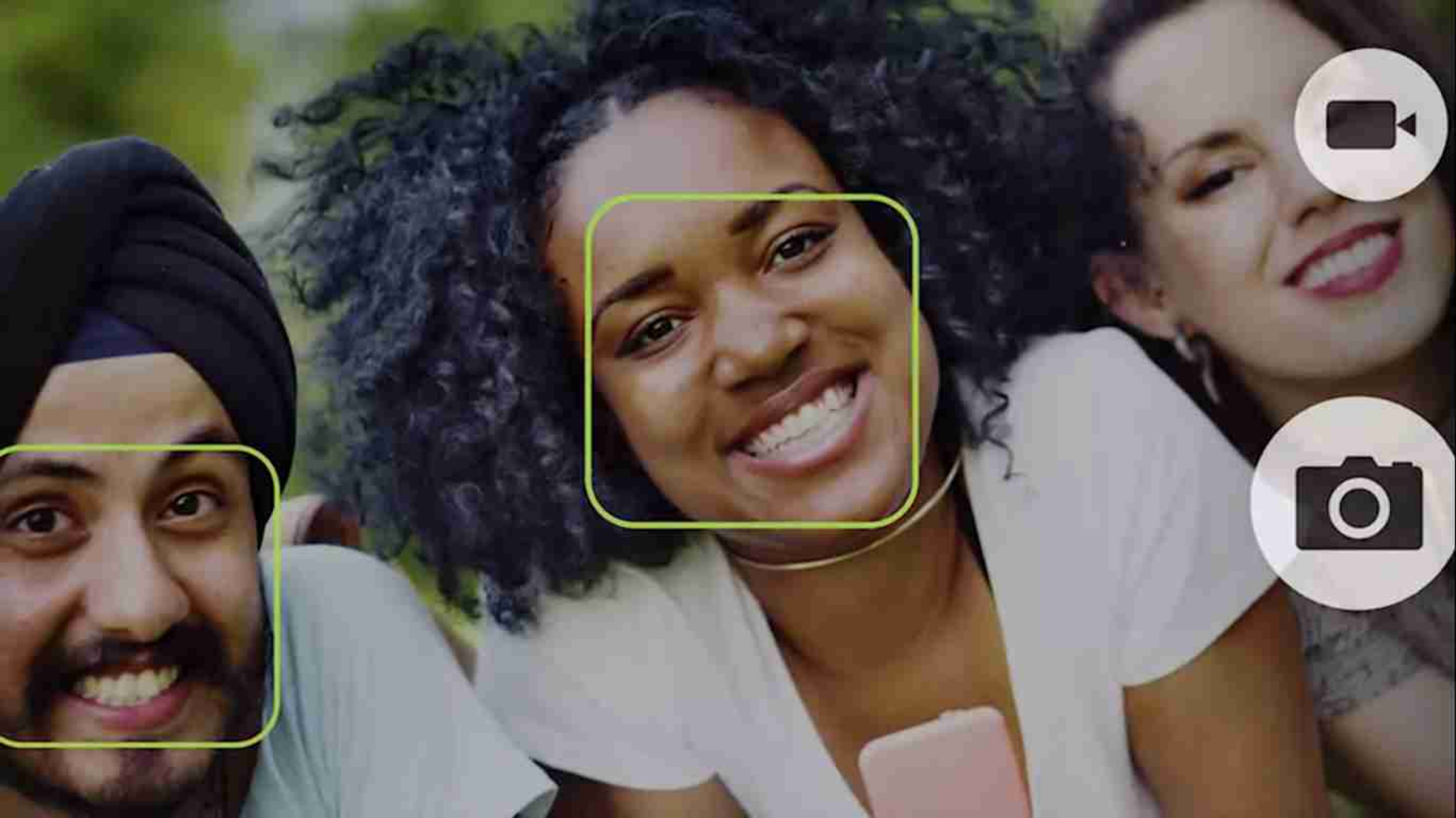
# How to hide from the AI surveillance state with a color printout

AI-powered video technology is becoming ubiquitous, tracking our faces and bodies through stores, offices, and public spaces. In some countries the technology constitutes a powerful new layer of policing and government surveillance

Fooling automated surveillance cameras: adversarial patches to attack person detection
Simen Thys, Wiebe Van Ranst, Toon Goedemé


car
car
person
bicycle
bicycle
bicycle
Bosch eBike
Systems
BOSCH
invented for life
BURGER KING


Share
Press Esc to exit full screen
Mathematical Preliminaries
0:08 / 0:27



White balance, exposure, dynamic range, and other features may be different.

Image matching is still able to identify matching features.

Share

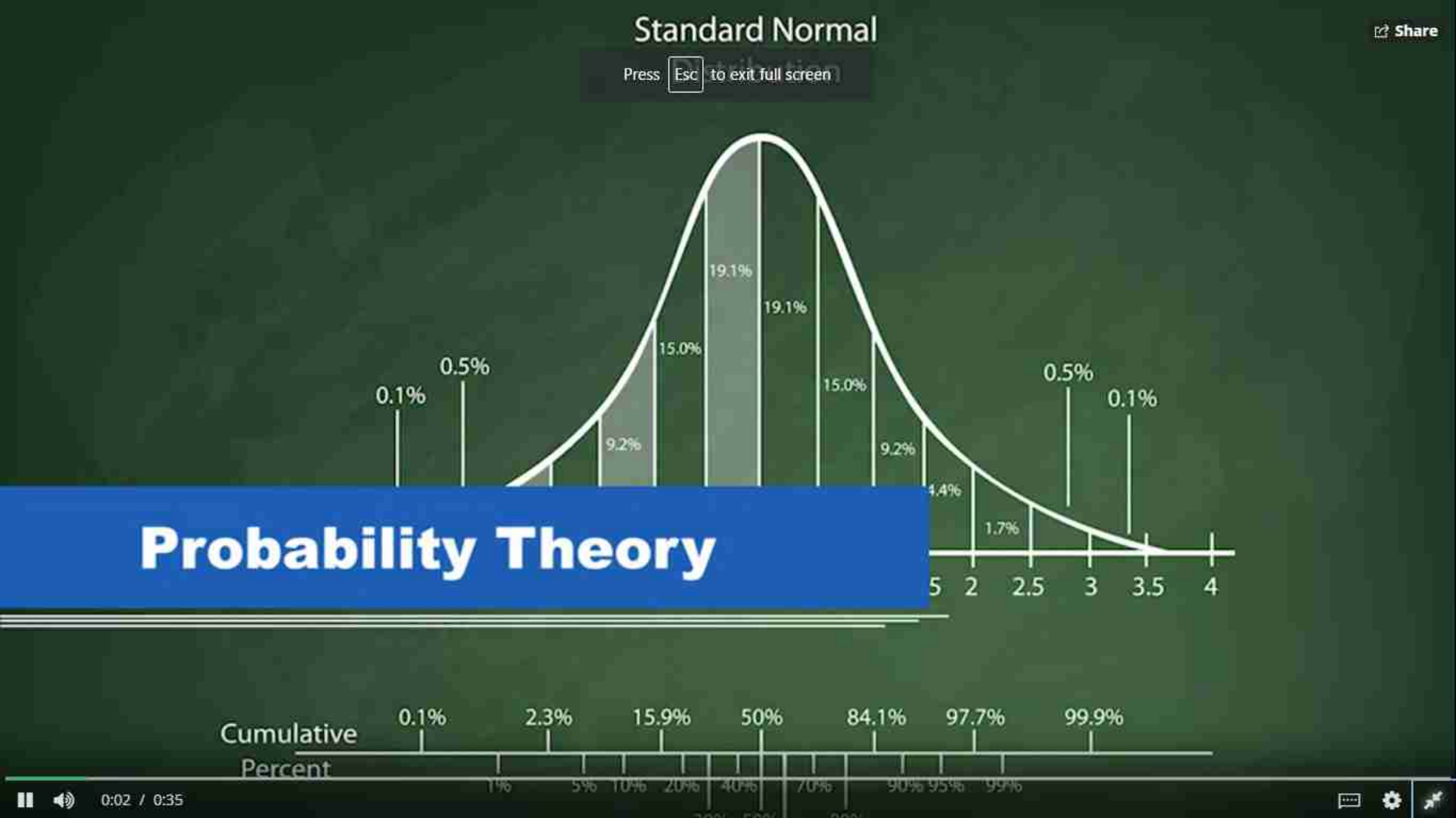
# Calculus - Two Major Branches

differential calculus

slopes of curves

integral calculus

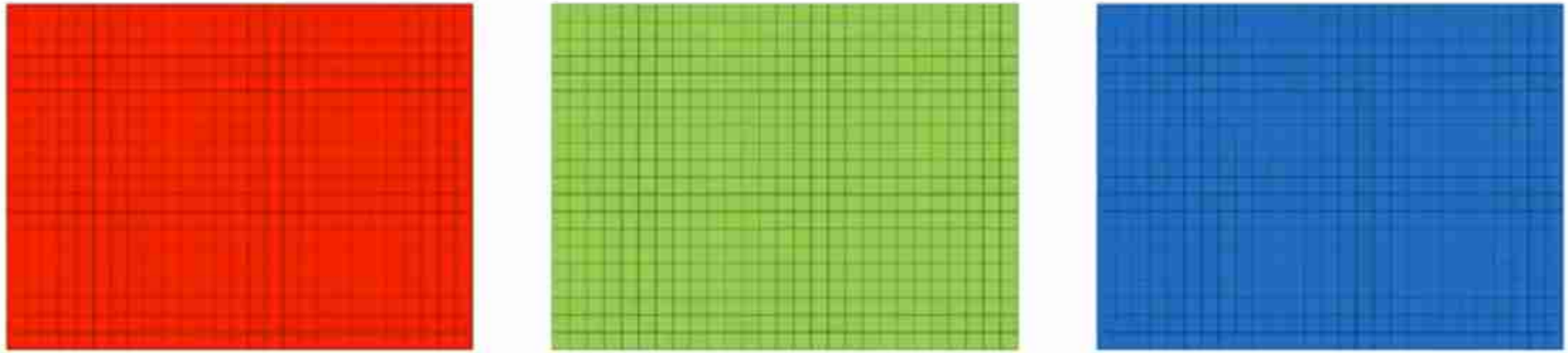
0:26 / 0:45


Standard Normal
Share
Press Esc to exit full screen
0.1%
0.5%
9.2%
15.0%
19.1%
19.1%
15.0%
9.2%
4.4%
1.7%
0.5%
0.1%
2
2.5
3
3.5
4
Probability Theory
Cumulative
Percent
0.1%
2.3%
15.9%
50%
84.1%
97.7%
99.9%
1%
5%
10%
20%
40%
70%
90%
95%
99%
0:02 / 0:35

# Dynamic Programming

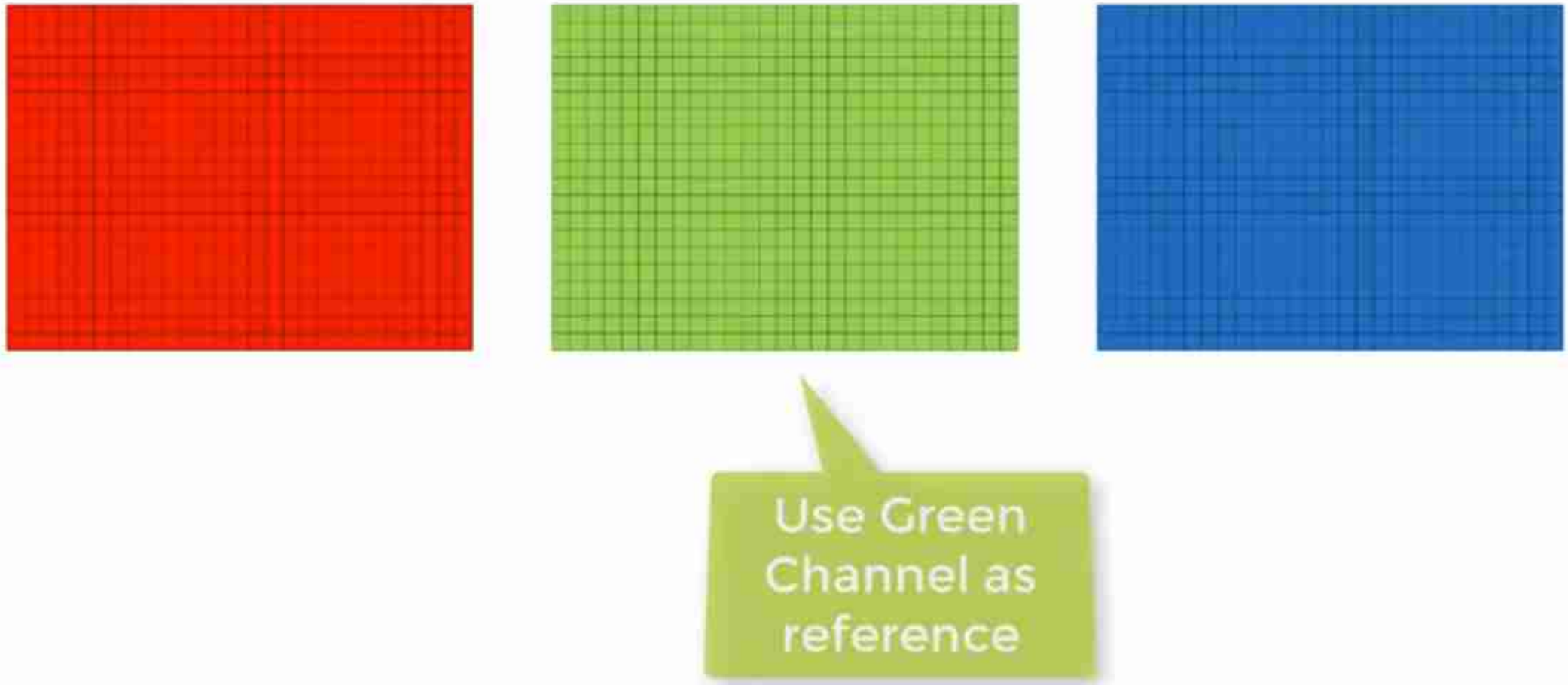
used in applications pertaining to
stereo matching and seam carving

Share



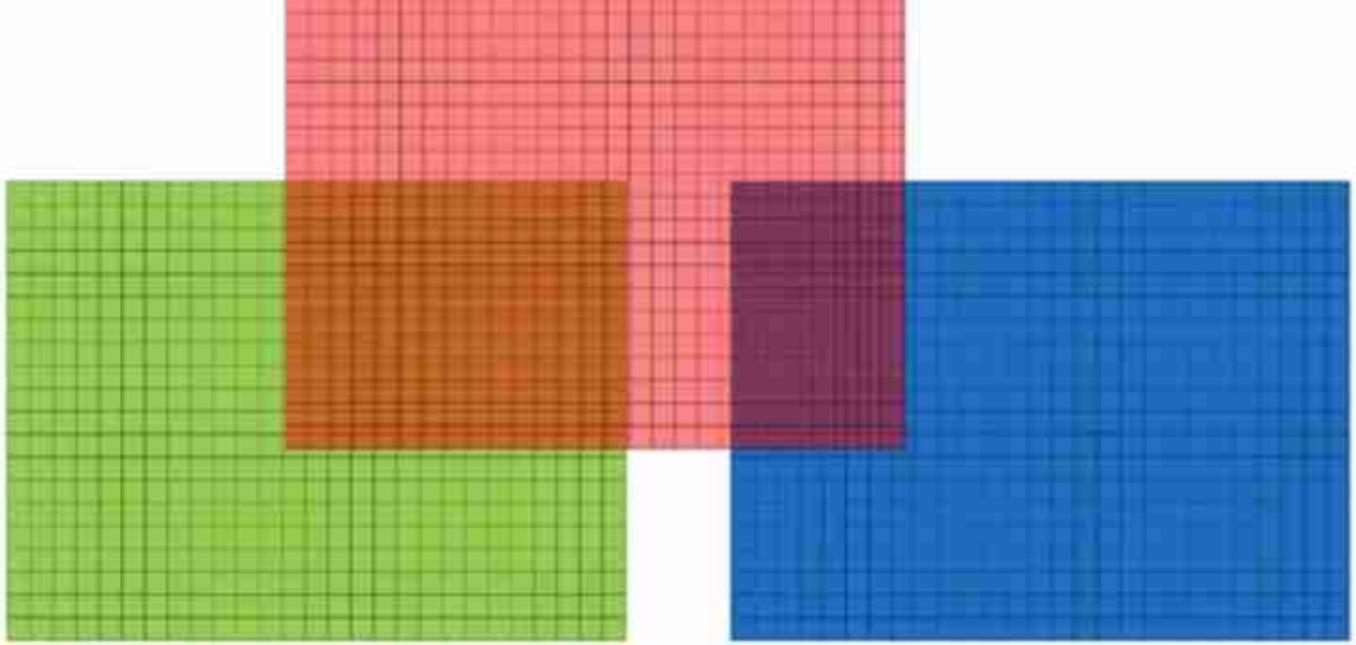
0:05 / 2:01

Share



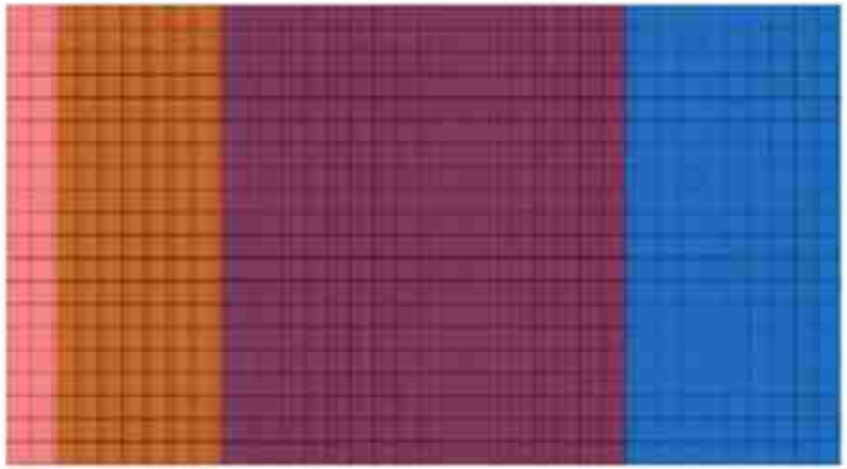

0:08 / 2:01

Use SSD as Matching metric



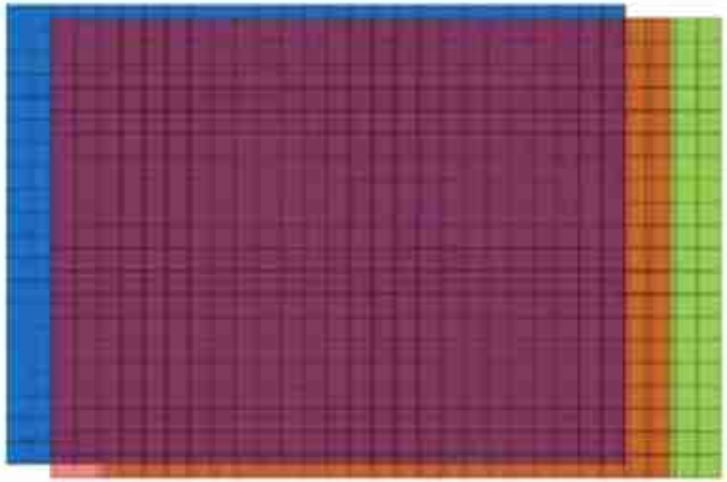
Slide Red channel over the Green Channel and settle at the best match

Use SSD as Matching metric

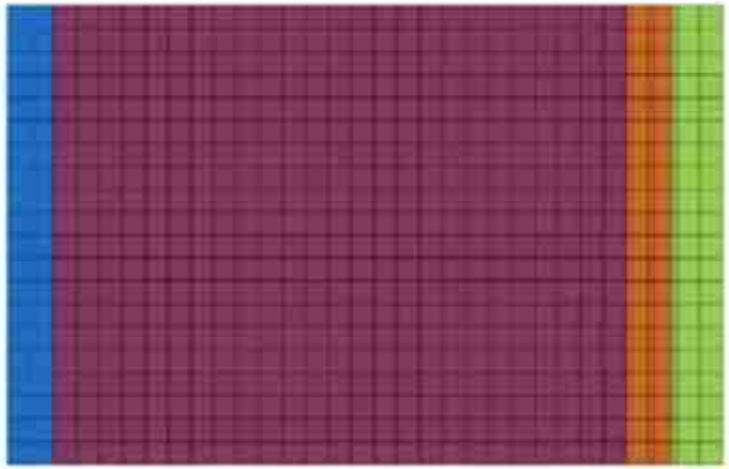


Slide Red channel over the Green Channel and settle at the best match

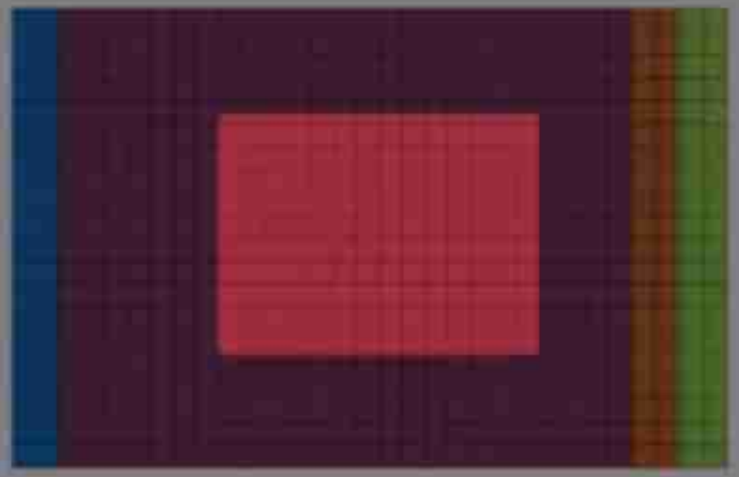
Use SSD as
Matching metric



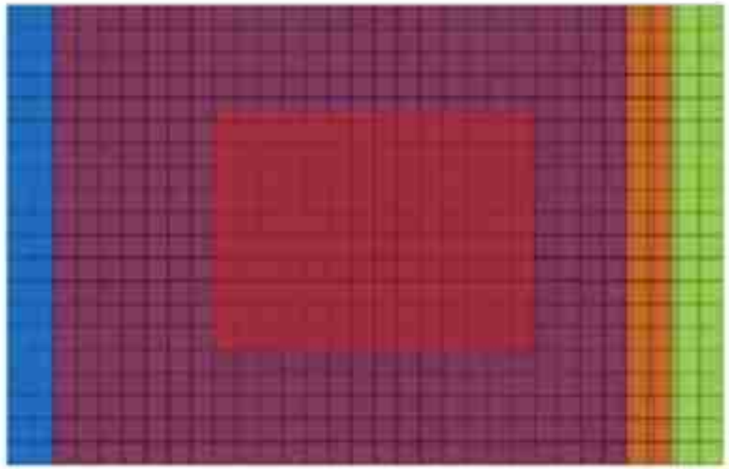
Repeat the
same with the
Blue Channel



Use a smaller region rather than the whole image for SSD Matching



Use a smaller region rather than the whole image for SSD Matching



Select a small region and use Matlab circshift function

circshift

# Sum of Squared Differences (SSD)

Share

$$\begin{bmatrix} 2 & 4 & 1 \\ 5 & 3 & 2 \\ 4 & 2 & 1 \end{bmatrix} - \begin{bmatrix} 5 & 2 & 5 \\ 7 & 1 & 2 \\ 1 & 5 & 4 \end{bmatrix} = \begin{bmatrix} 2-5 & 4-2 & 1-5 \\ 5-7 & 3-1 & 2-2 \\ 4-1 & 2-5 & 1-4 \end{bmatrix}$$

A B C= A - B

1:27 / 2:01

| | | |
|---|---|---|
| -3 | 2 | -4 |
| -2 | 2 | 0 |
| 3 | -3 | -3 |

uint8 does not support negative values

| | | |
|---|---|---|
| -3 | 2 | -4 |
| -2 | 2 | 0 |
| 3 | -3 | -3 |

uint8 does not support negative values

Make sure A, B and C are double type

SSD =

| | | |
|---|---|---|
| $(-3)^2$ | $(2)^2$ | $(-4)^2$ |
| $(-2)^2$ | $(2)^2$ | 0 |
| $(3)^2$ | $(-3)^2$ | $(-3)^2$ |

SSD = SUM

| | | |
|---|---|---|
| $(-3)^2$ | $(2)^2$ | $(-4)^2$ |
| $(-2)^2$ | $(2)^2$ | 0 |
| $(3)^2$ | $(-3)^2$ | $(-3)^2$ |