Computer Visions

Keith Reis & Justin Ta

Overview -

What we'll be discussing:

~ Computer Vision

~ Image Formation

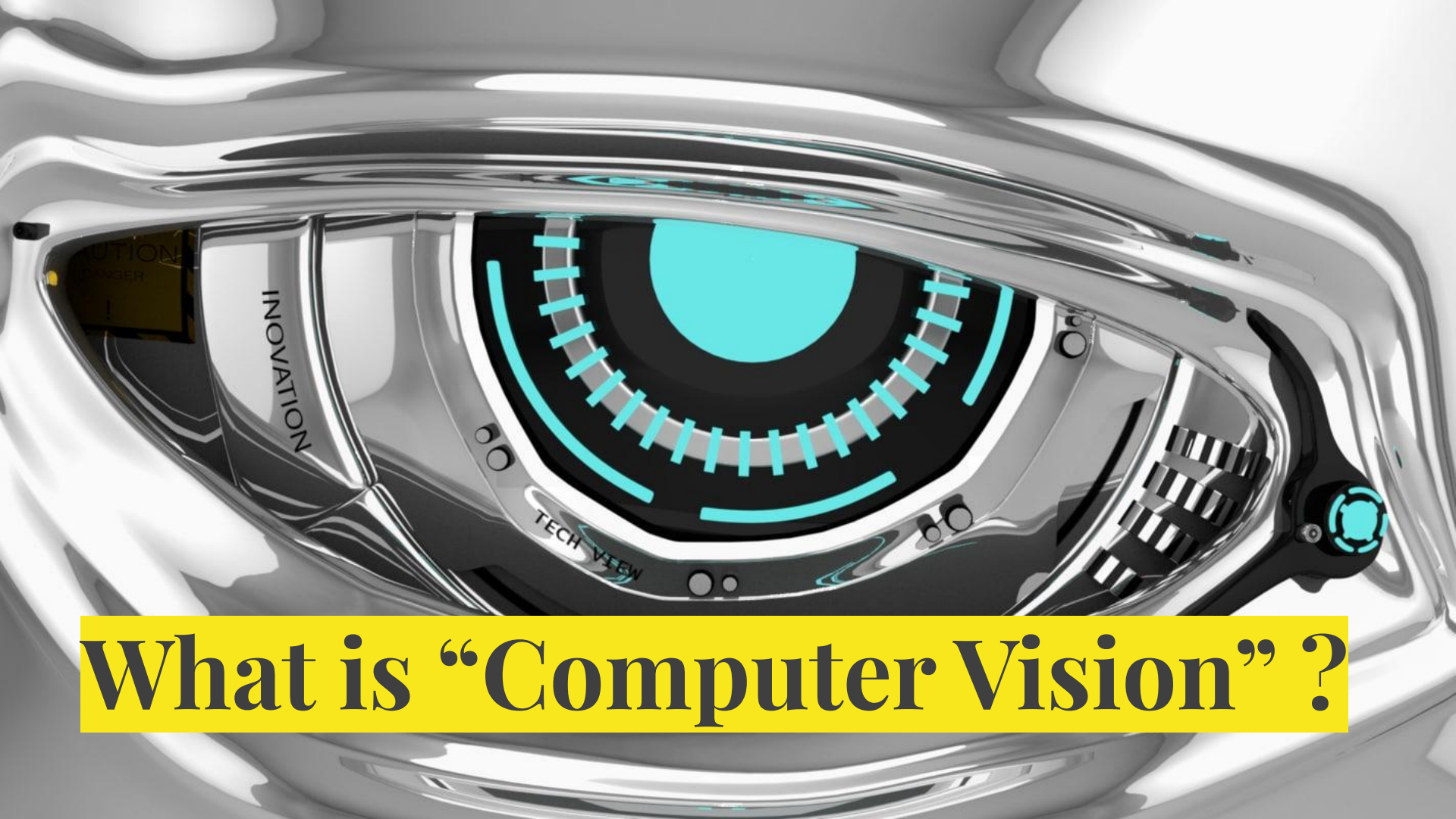
~ Simple Image Features

~ Classifying Images

~ Detecting Objects

~ The 3D World

~ Using Computer Vision



What is “Computer Vision” ?

Definition of Computer Vision -

Computer Vision, is a specific field of computer science that works to enable computers to see, identify, and process visuals in the same way a human would.

Examples of Interesting Implementations of Computer Vision:

- **Autonomous Vehicles (Self-Driving Cars)**
- **Advanced Facial Recognition Systems**
- **Advanced Agriculture**

2 Types of Computer Vision Sensing

Active Sensing:

Send out signal (radar,ultrasound,etc.) in order to identify & process a reflection.

Examples:

- Whales
- Submarine

Passive Sensing:

Used by most agents that have vision, do not need to output light in order to see.

Examples:

- Human
- Camera

2 Core Problems of Computer Vision

Reconstruction: Agent constructs replica of surrounding environment from image or set of images.

Examples:

- Image Upscaling
- Facial Reconstruction

Recognition: Agent showing understanding of the visual in front of them, then being able to identify what it is.

Examples:

- Snapchat Filters
- IPhone Face ID
- Camera Recognition
- Drone Target Lock

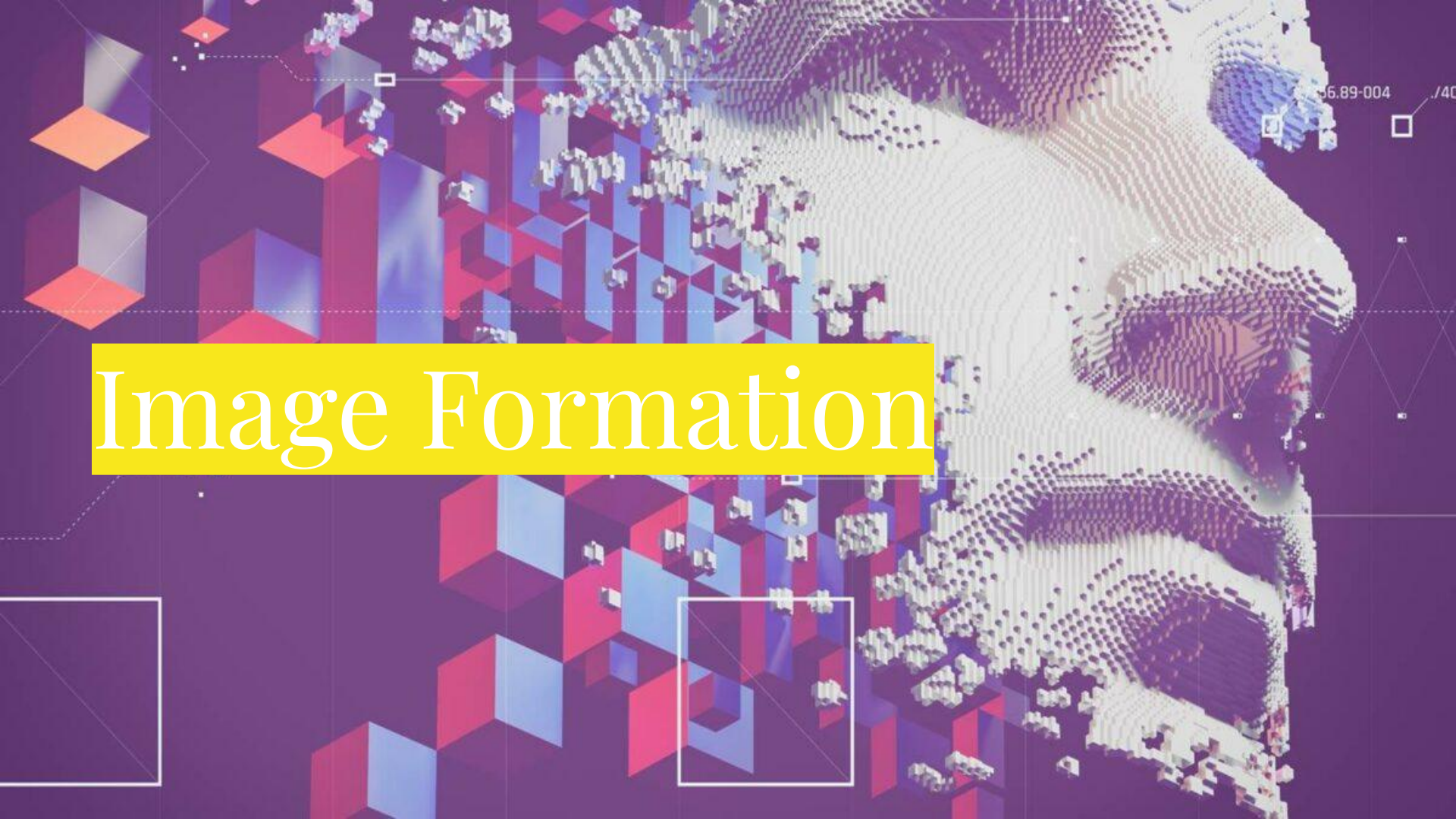


Image Formation

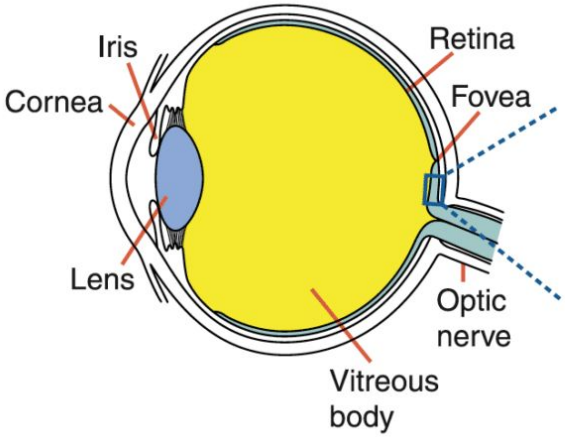
What is Image Formation?

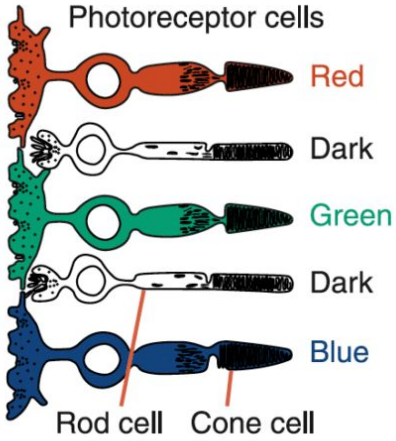
The system of compuer vision being able to recognize any perspective distortions(foreshortening) in an image,

Difference Between Human & Computer Image Formation

Humans:

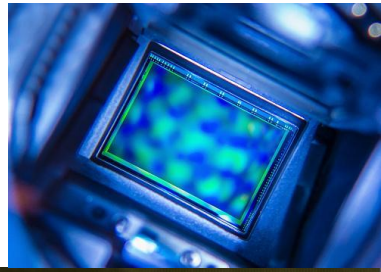
- **Process environments based on bio-chemical signals.**
- **Use rods(light/dark), and cones(color) for visualizing.**

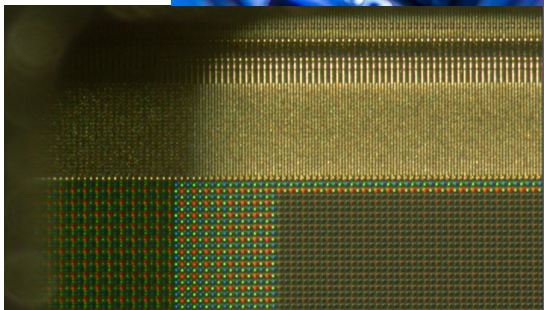




Computers:

- **Process environments based on tiny bits & bytes of pixels.**
- **Use digital sensors(pixel) and artificial eyes to capture images.**

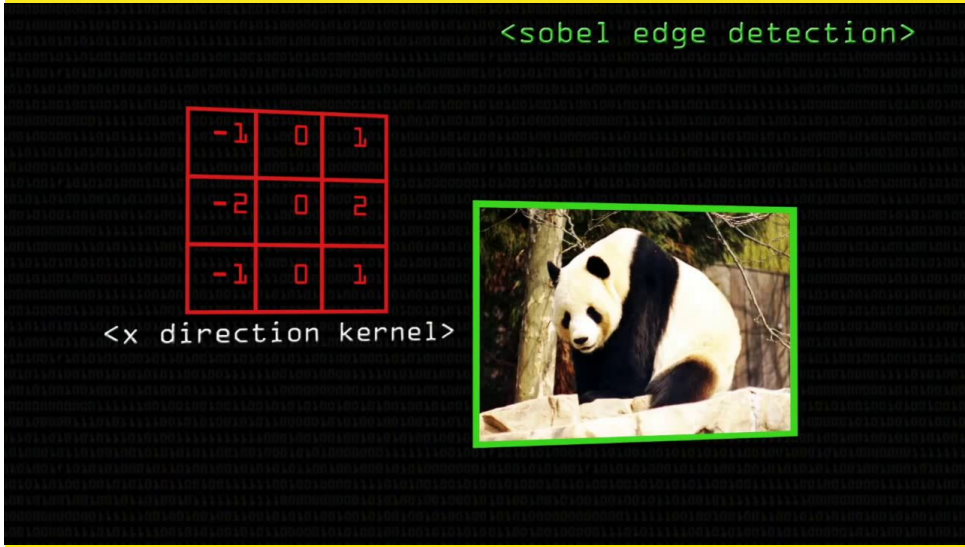





Simple Image Features

What is “Simple Image Features”?

The edges, optical flow, texture, and segmentation into regions of an image.



Edges

Edges are straight lines/bends in an image plane where there is significant difference in the value within it's neighboring pixels.

Four Types of Edges:

- Depth Discontinuities (Depth)
- Surface Orientation Discontinuities (Surfaces)
- Reflectance Discontinuities (Reflections)
- Illumination Discontinuities (Shadows)

Texture

Texture refers to the pattern on a surface that can be identified visually.

A rough model of texture is a repetitive pattern of elements, referred to as **‘texels’.**

Example:

A cheetah and a black panther have similar shapes/edges, but very distinct differing texels.

Optical Flow

The pattern of clearly visible object motions, surfaces, and edges in a moving scene; caused by the movement of the observer around the scene.

Optical Flow Vector Field: x & y field where corresponding points are found between one frame & the next.

2 Uses of Optical Flow:

- Video Stabilization
- Frame Interpolation(Estimating)

video rights courtesy @cebleague

Segmentation

The process of taking apart an image, and separating it into groups of similar pixels.

- Each pixel has specific visual properties attached to it: texture, brightness, color,etc.

2 Types of Segmentation:

- Boundary Detection
- Region Extraction

Classifying Images

What does it mean to ‘Classify Images’?

Analyzes an image and tells us which class it falls under by using pre-trained classifying models.

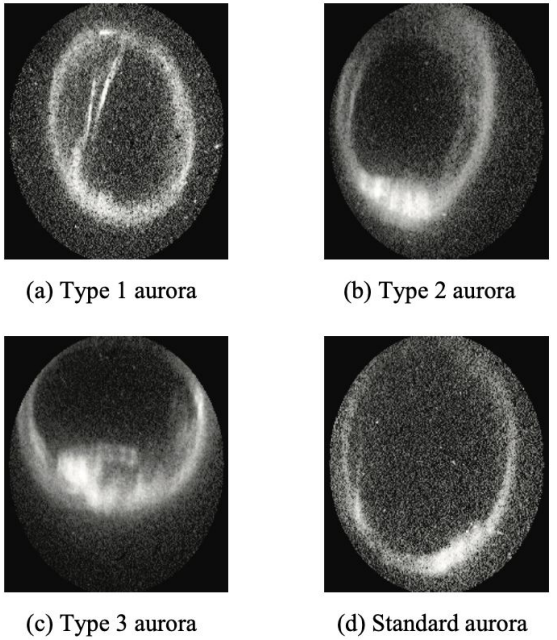
Two cases:

- Objects
- Scene

What does it mean to ‘Classify Images’?

What does it mean to ‘Classify Images’?

Example: Classifying different types of aurora using the geometric features of the aurora image.



(a) Type 1 aurora (b) Type 2 aurora

(c) Type 3 aurora (d) Standard aurora

Fig. 7. Three specific types of aurora

*Photo In courtesy of Dr. Soo Kim’s research

Factors that affects appearance of the image:

- Lighting
- Foreshortening
- Aspect
- Occlusion
- Deformation

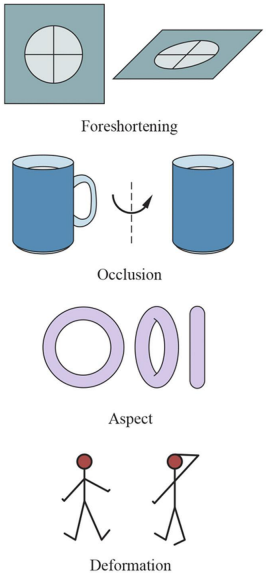


*Photo by Remy van der Winden

Convolutional Neural Networks for Classifying Images

- Spectacularly successful image classifiers.
- Much better results than other methods
- ImageNet dataset: 14 million training images, 30,000 fine-grained categories.

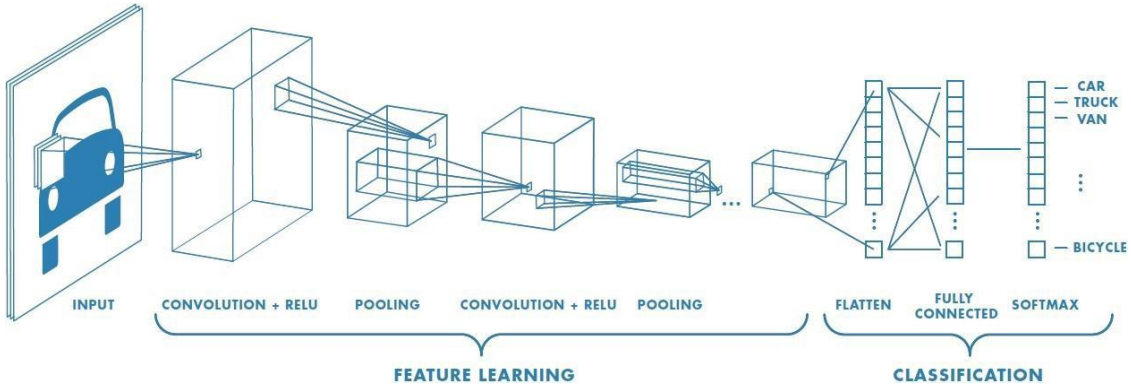


Image by Sumit Saha

Koala's eye? = Y

Koala's nose? = Y

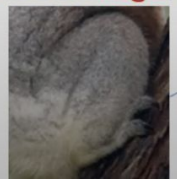
Koala's ears? = Y

Koala's head? = Y

Is it Koala? = Y

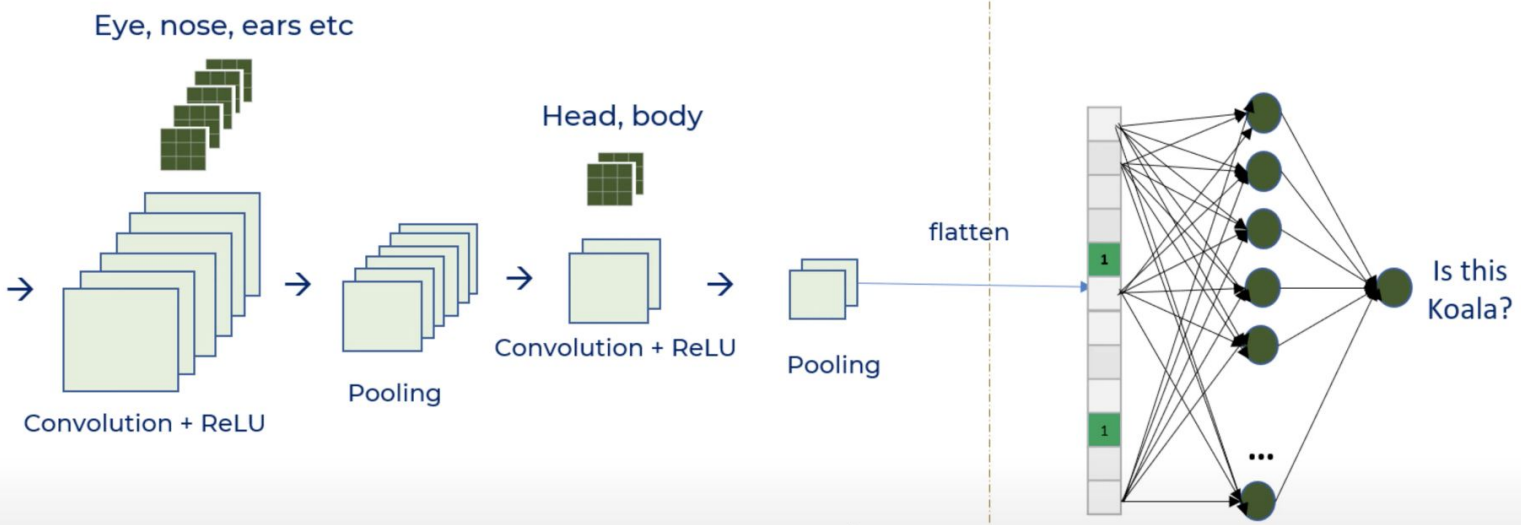
Koala's hands? = Y

Koala's legs? = Y

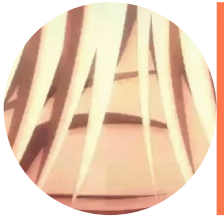


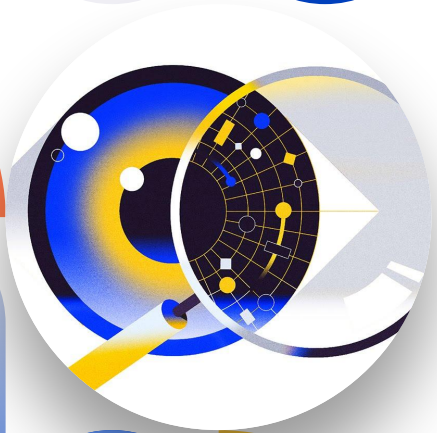
Koala's body? = Y

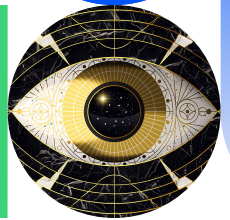
Convolutional Neural Networks for Classifying Images



Detecting Objects



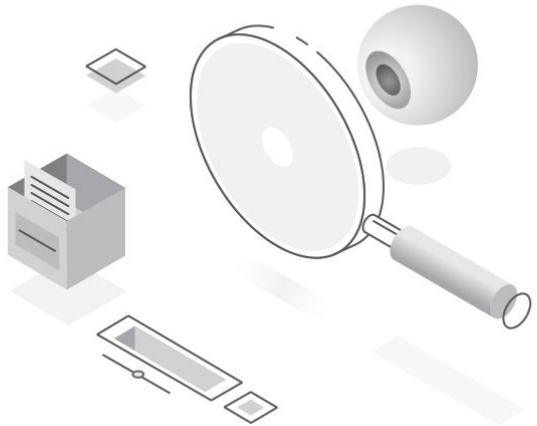


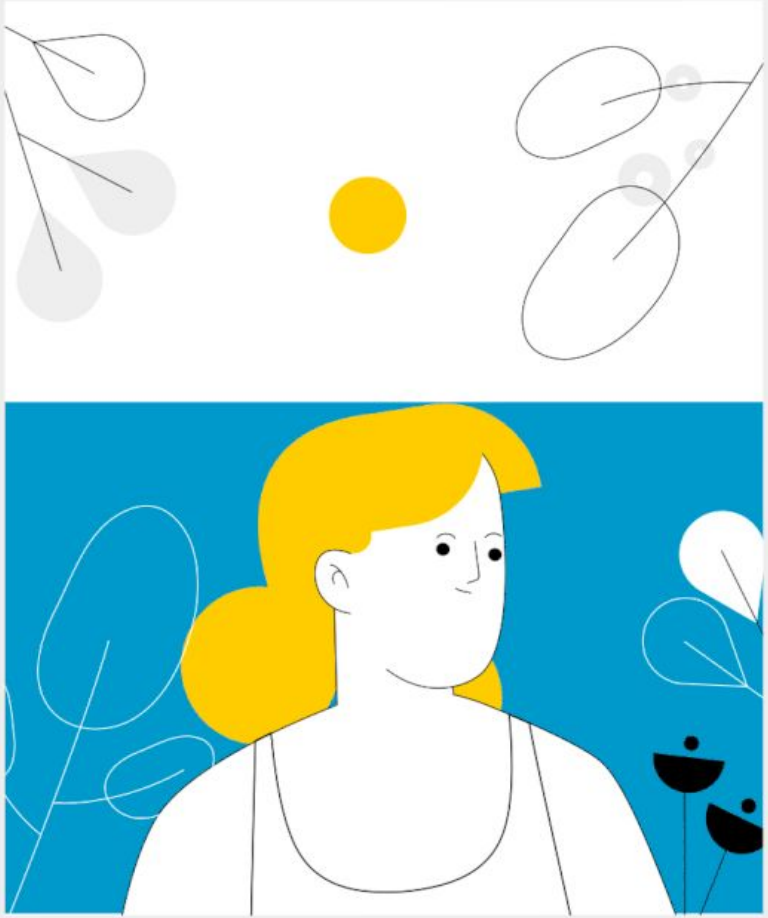


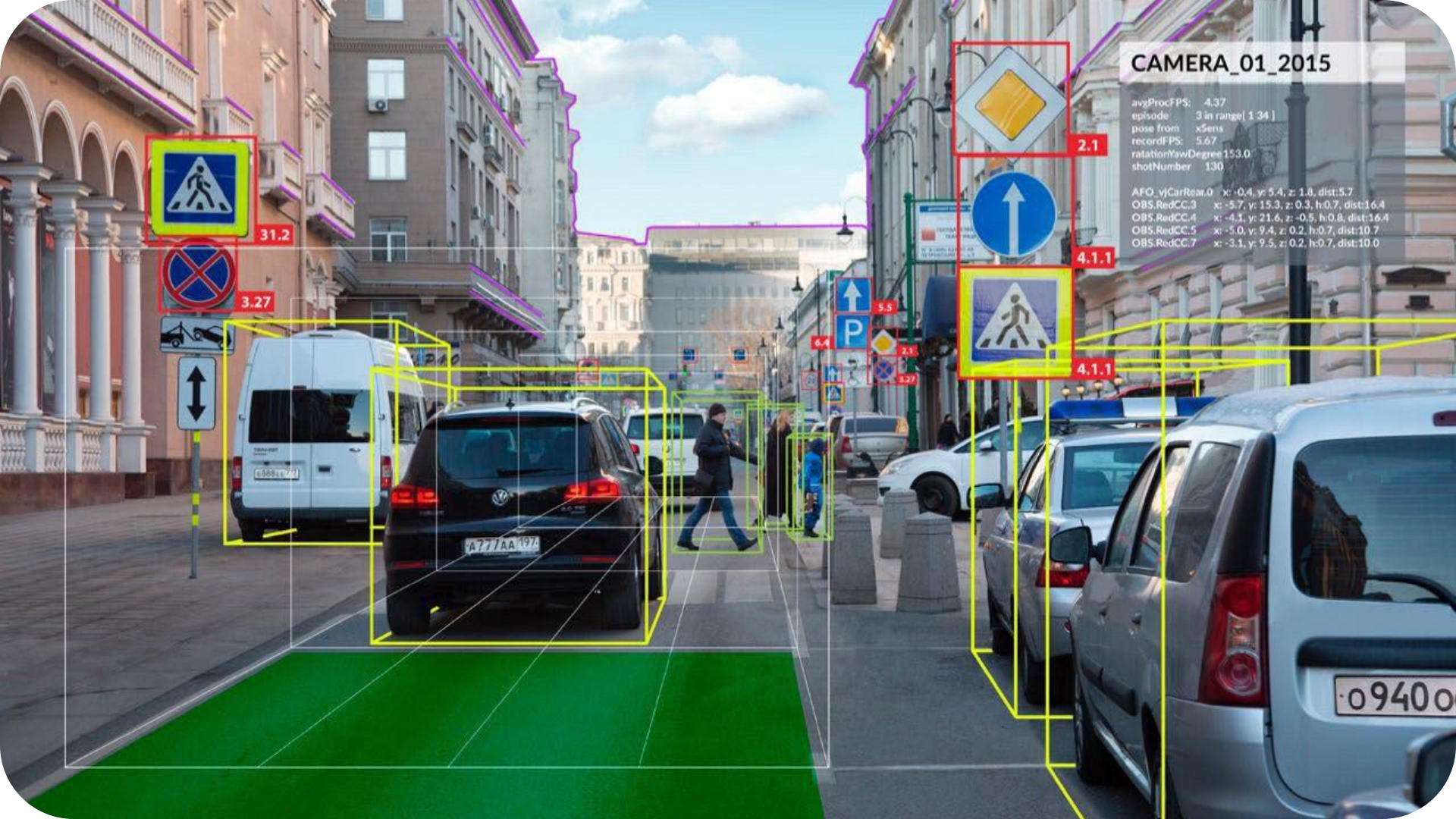
How Can Computer Vision Detect Objects?

Object Detector: Finds objects within an image, reports what specific class each object would be in, then reports where each object is by providing it with a bounding box.

Image Localization: identify the location of objects in the given image.






CAMERA_01_2015
avgProcFPS: 4.37
episode 3 in range[1 34]
pose from x5ens
recordFPS: 5.67
ratationYawDegree153.0
shotNumber 130
AFO_vjCarRear.0 x: -0.4, y: 5.4, z: 1.8, dist:5.7
OBS.RedCC.3 x: -5.7, y: 15.3, z: 0.3, h:0.7, dist:16.4
OBS.RedCC.4 x: -4.1, y: 21.6, z: -0.5, h:0.8, dist:16.4
OBS.RedCC.5 x: -5.0, y: 9.4, z: 0.2, h:0.7, dist:10.7
OBS.RedCC.7 x: -3.1, y: 9.5, z: 0.2, h:0.7, dist:10.0
31.2
3.27
2.1
4.1.1
4.1.1
5.5
6.4
2.1
3.27
A777AA197
o940o

Objects Detection Flow

Regional Proposal Network (RPN)

A network that finds regions with objects

Region of Interest (ROI)

Boxes with a good enough objectness score

Non-maximum Suppression

Discard overlapping windows

Bounding Box Regression

Trim the window down to a proper bounding box

The 3D World

We Can Build 3D World From

MULTIPLE VIEWS CUES

By using information from multiple images of a scene.

BINOCULAR STEREOPSIS

Device with two eyes or two cameras in the front.

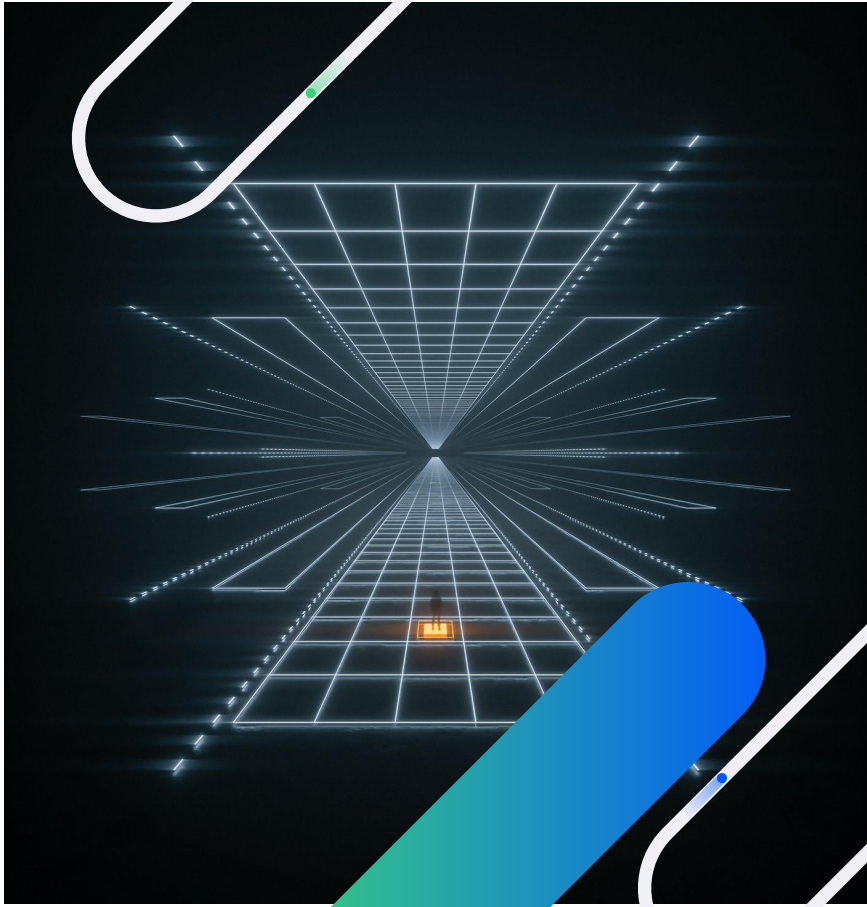
MOVING CAMERA

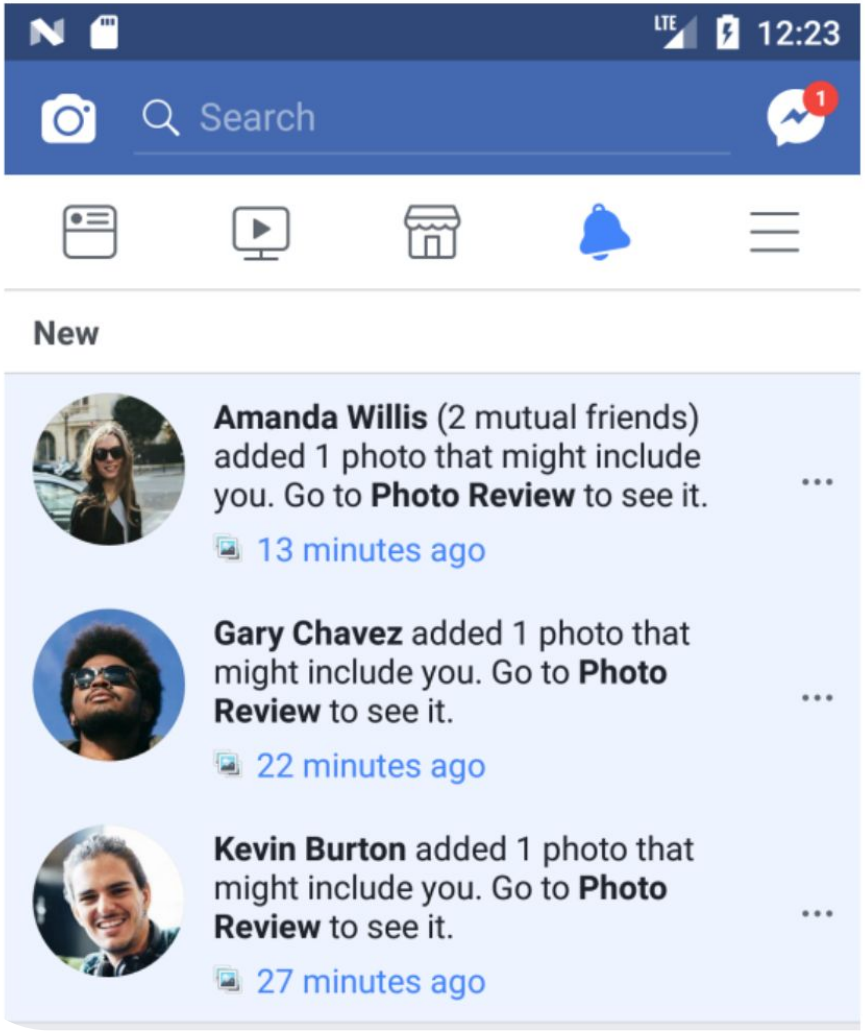
Using time and disparity in camera motion.

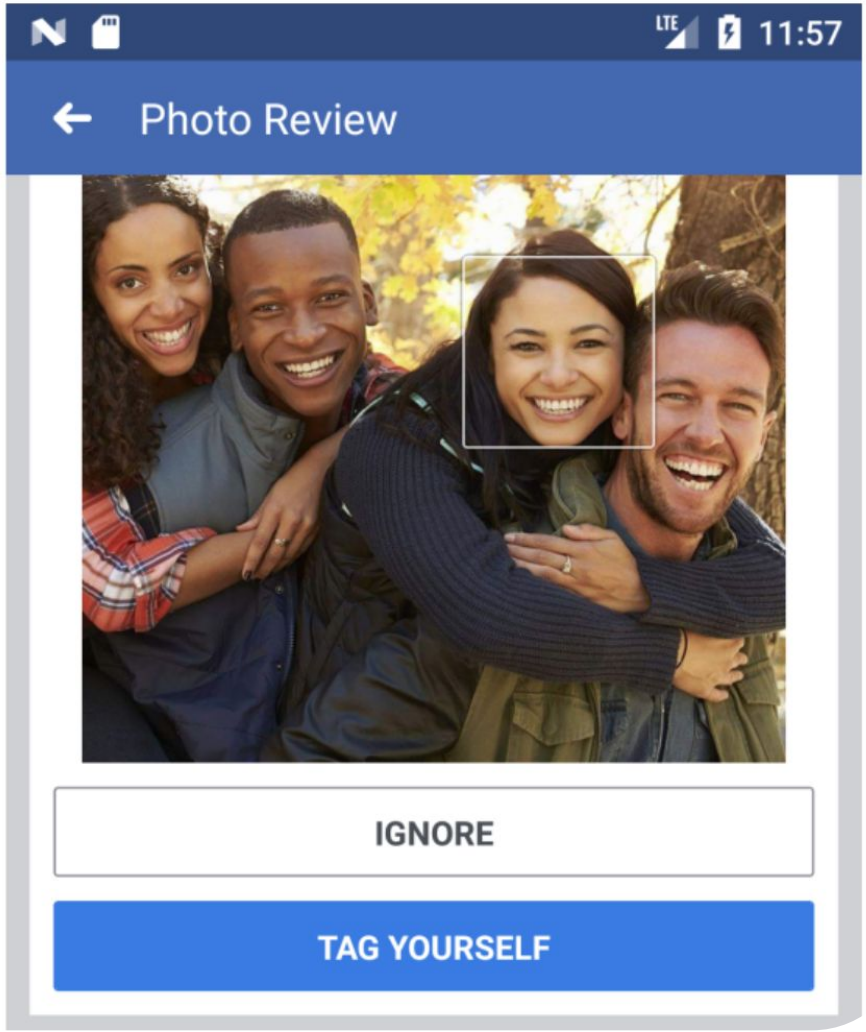
3D Cues from One View can also be possible, but it may be not as efficient as the other sources.

Using Computer Vision

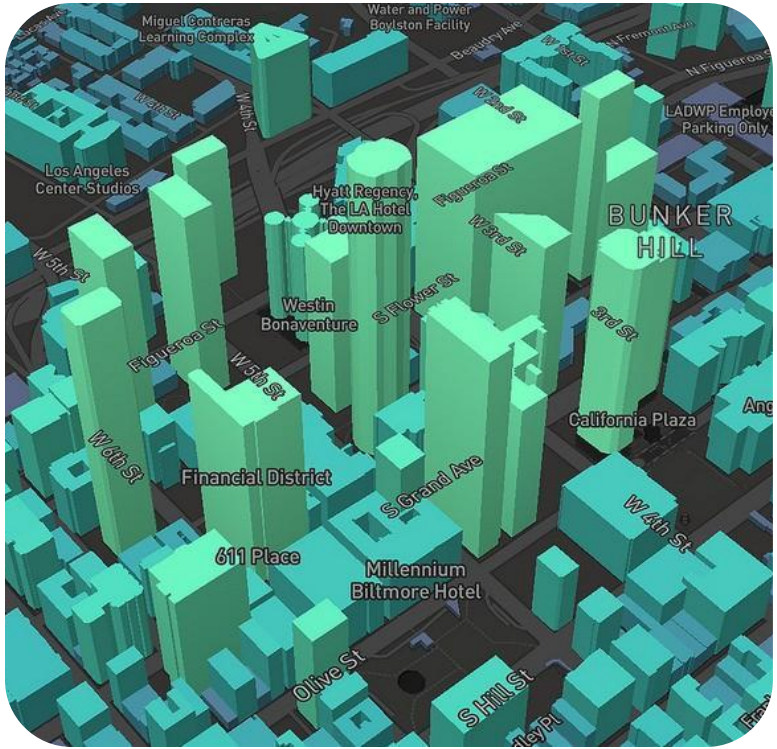
- To understand what people are doing
- To link pictures and words
- To reconstruct picture from many views
- To find geometry from a single view
- To make computer-generated pictures




12:23
Search
New
Amanda Willis (2 mutual friends) added 1 photo that might include you. Go to Photo Review to see it.
13 minutes ago
Gary Chavez added 1 photo that might include you. Go to Photo Review to see it.
22 minutes ago
Kevin Burton added 1 photo that might include you. Go to Photo Review to see it.
27 minutes ago

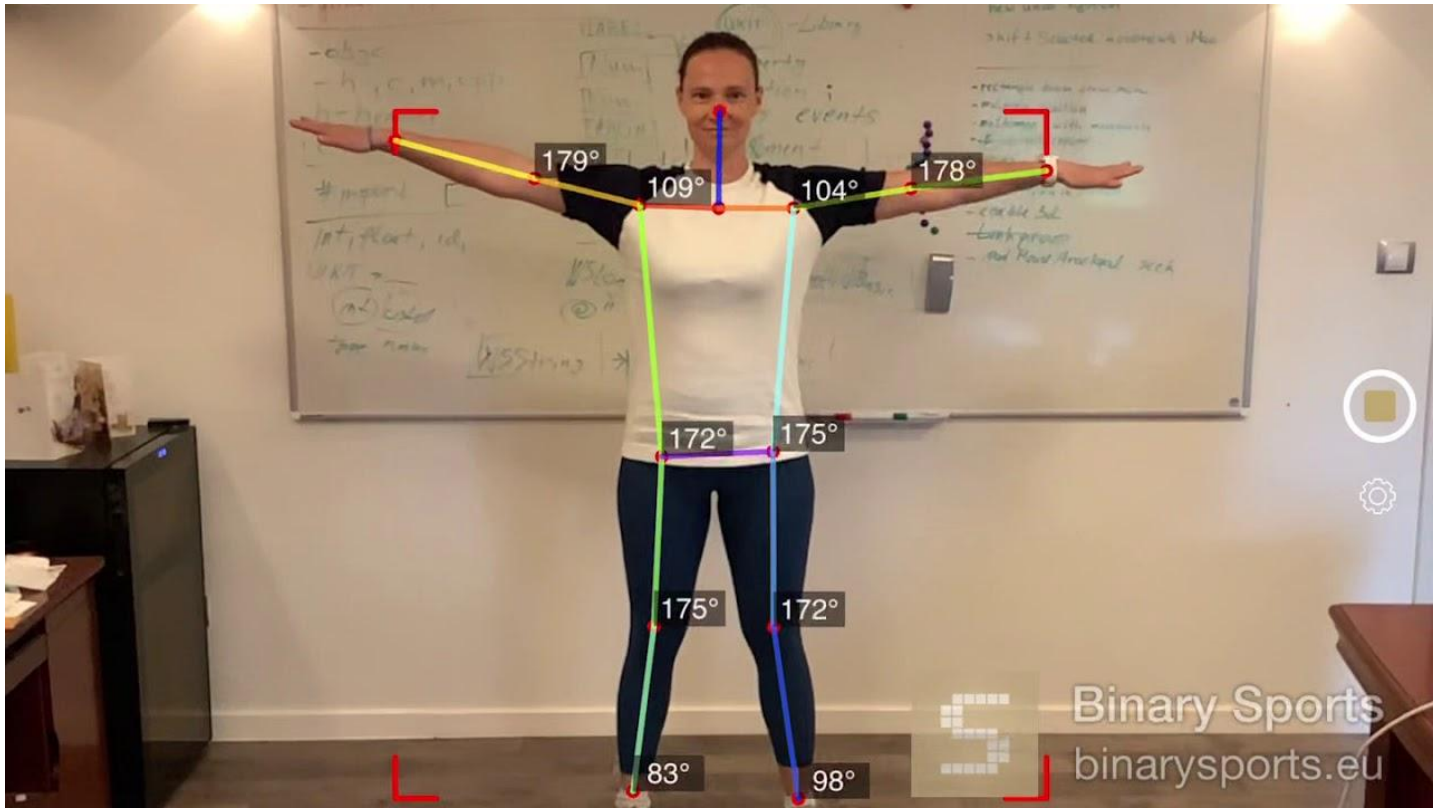

11:57
Photo Review
IGNORE
TAG YOURSELF

Build 3D maps.



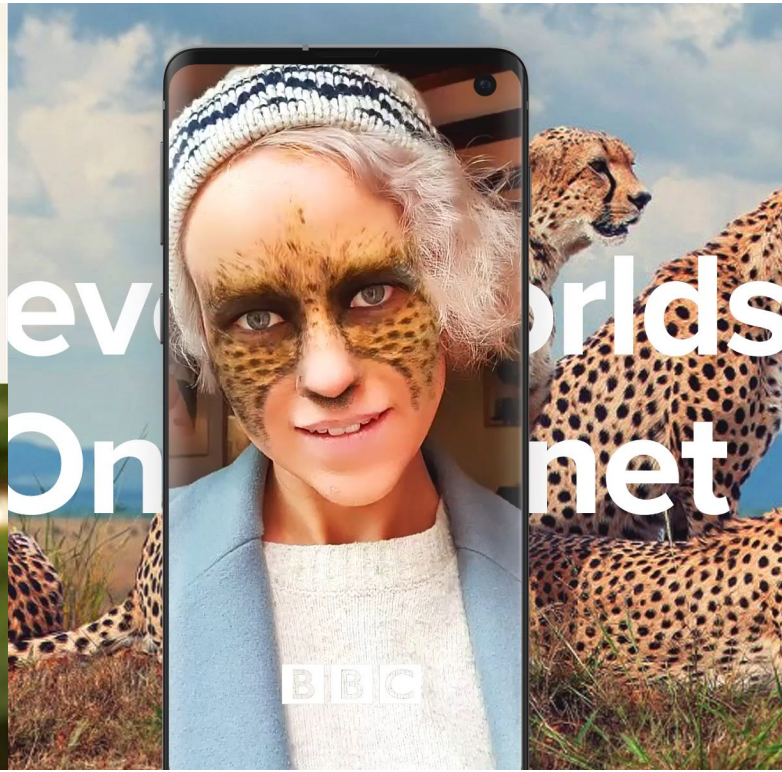
Reconstruct room.

iPad uses computer vision to track body movements in its Fitness app.



AI Face Generator.

AI Face Filters



Thank You For Listening!