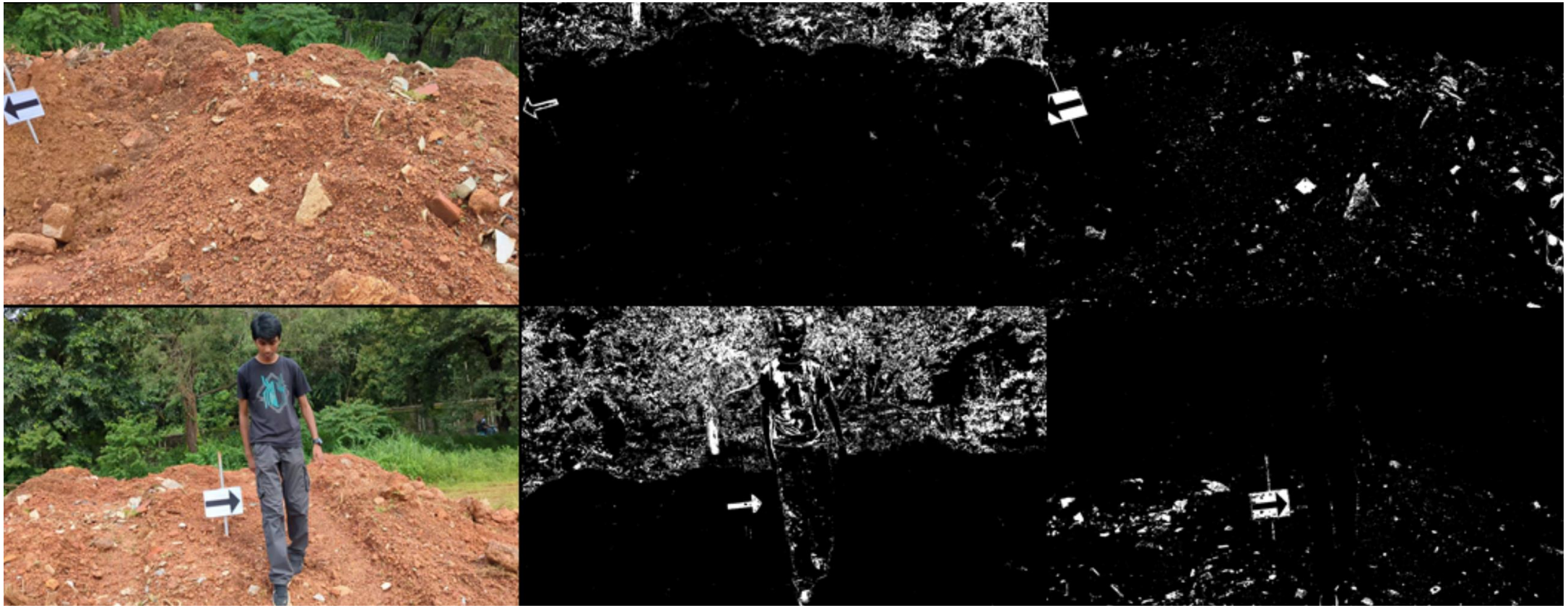
CSE 3172: Foundations of Computer Vision



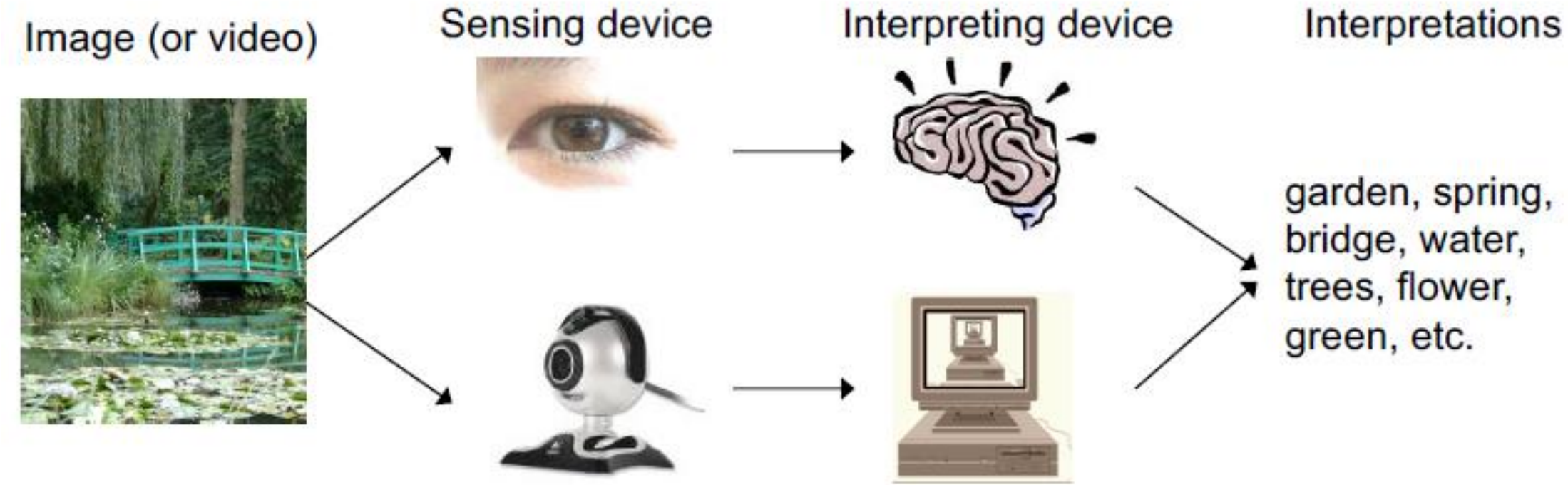
Slide Credits

- http://www.cs.washington.edu/education/courses/cse455/17wi/
- Center for Research in Computer Vision (CRCV), UCF
- Stanford AI Lab

Agenda

- *Introduction to computer vision*
- *References*
- *Mode of Assignments*

What is (computer) vision?



What is computer vision?

Extract descriptions of the world from images

Descriptions of what kind? qualitative vs. quantitative

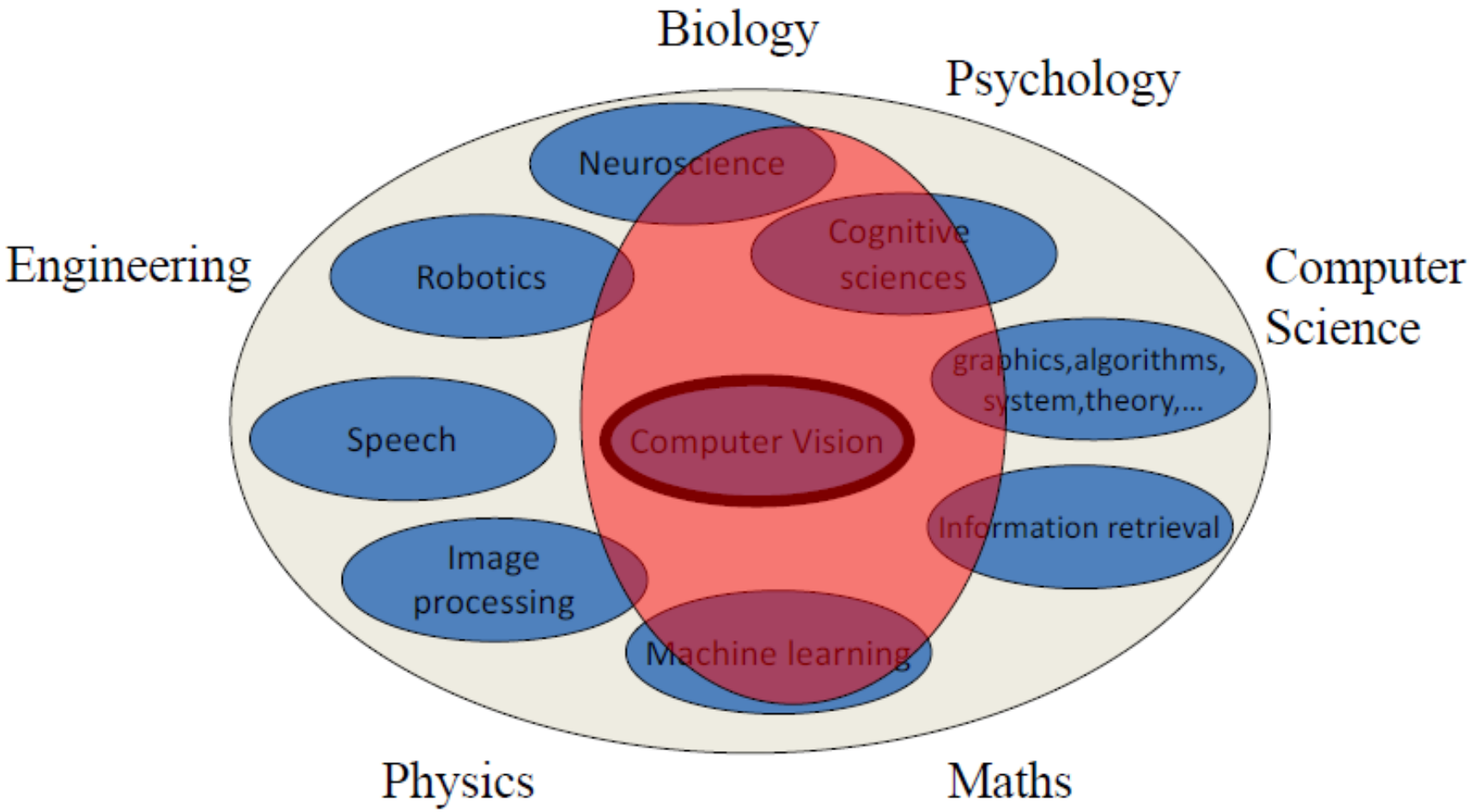
Geometric: shape and position of object or distances in 3D world

Semantic: what objects do I see?

Dynamic: scene changes, object velocities, human actions, ...

Credit: Francesca Odone, University of Genova

What is it related to?



The goal of computer vision

- To extract “meaning” from pixels

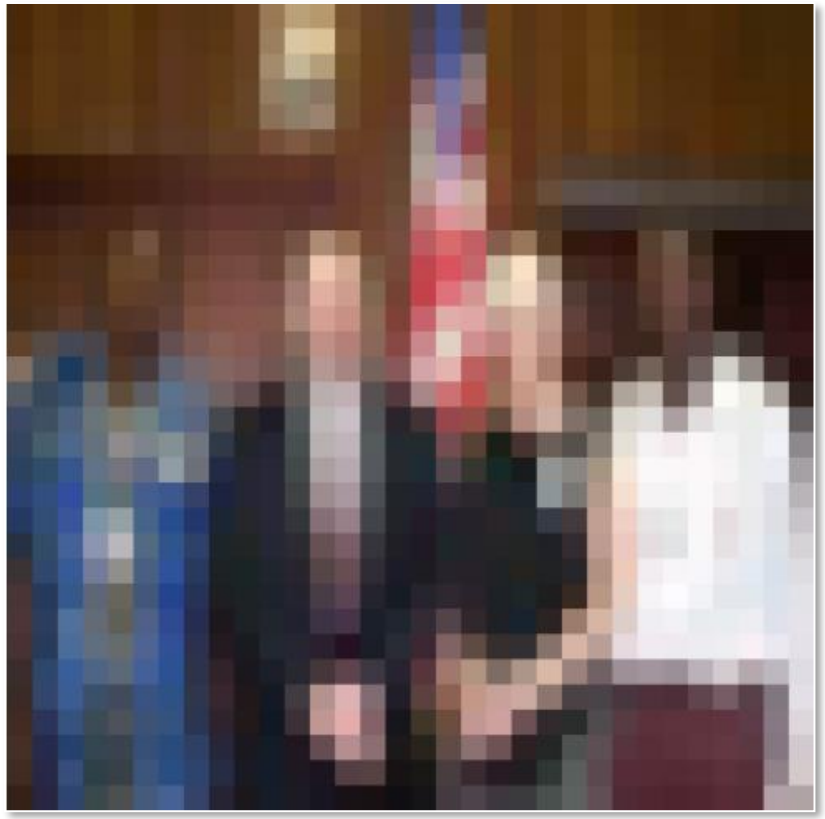
What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

The goal of computer vision

- To extract “meaning” from pixels



Humans are remarkably good at this…

Origins of computer vision: an MIT undergraduate summer project

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

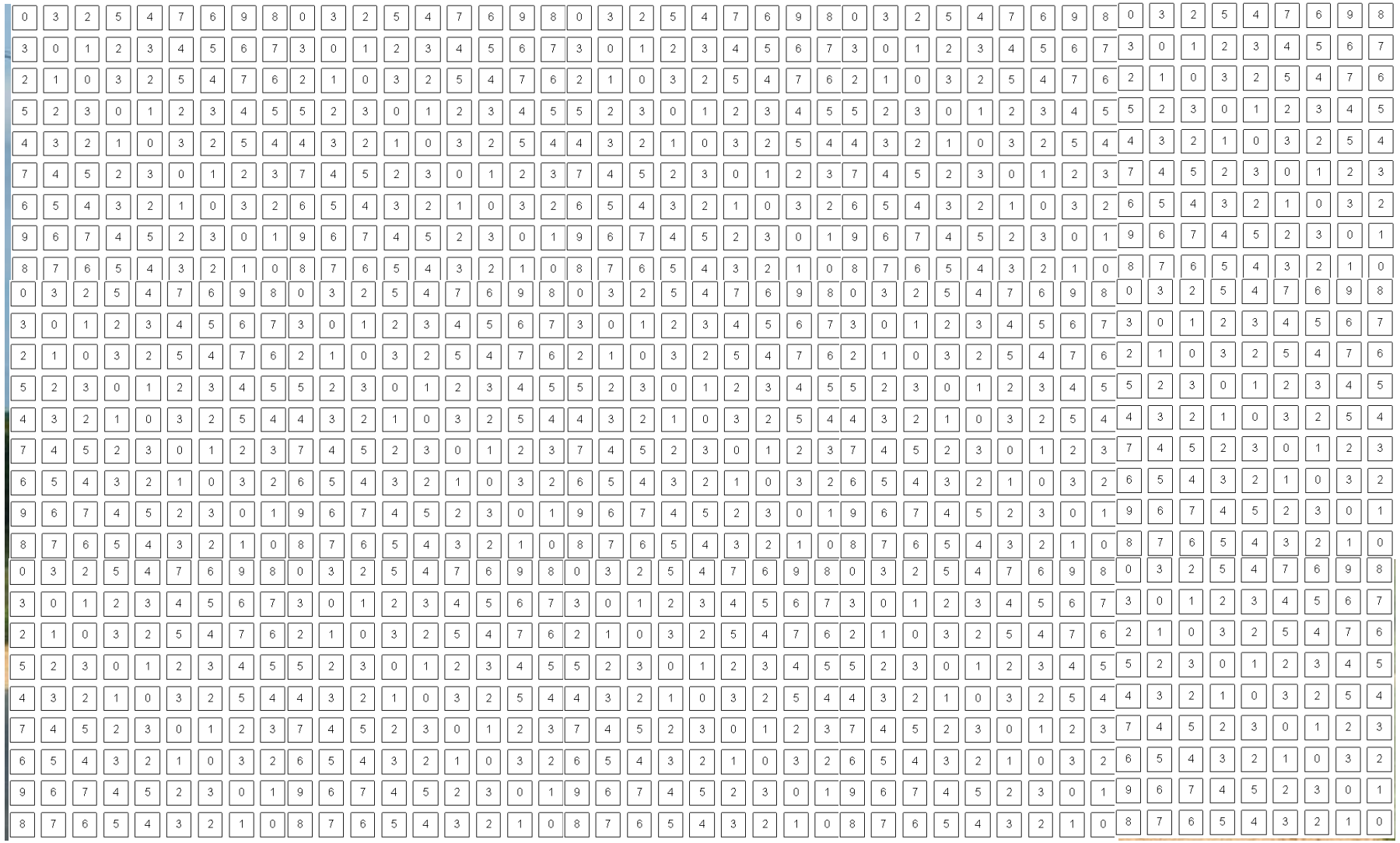
THE SUMMER VISION PROJECT

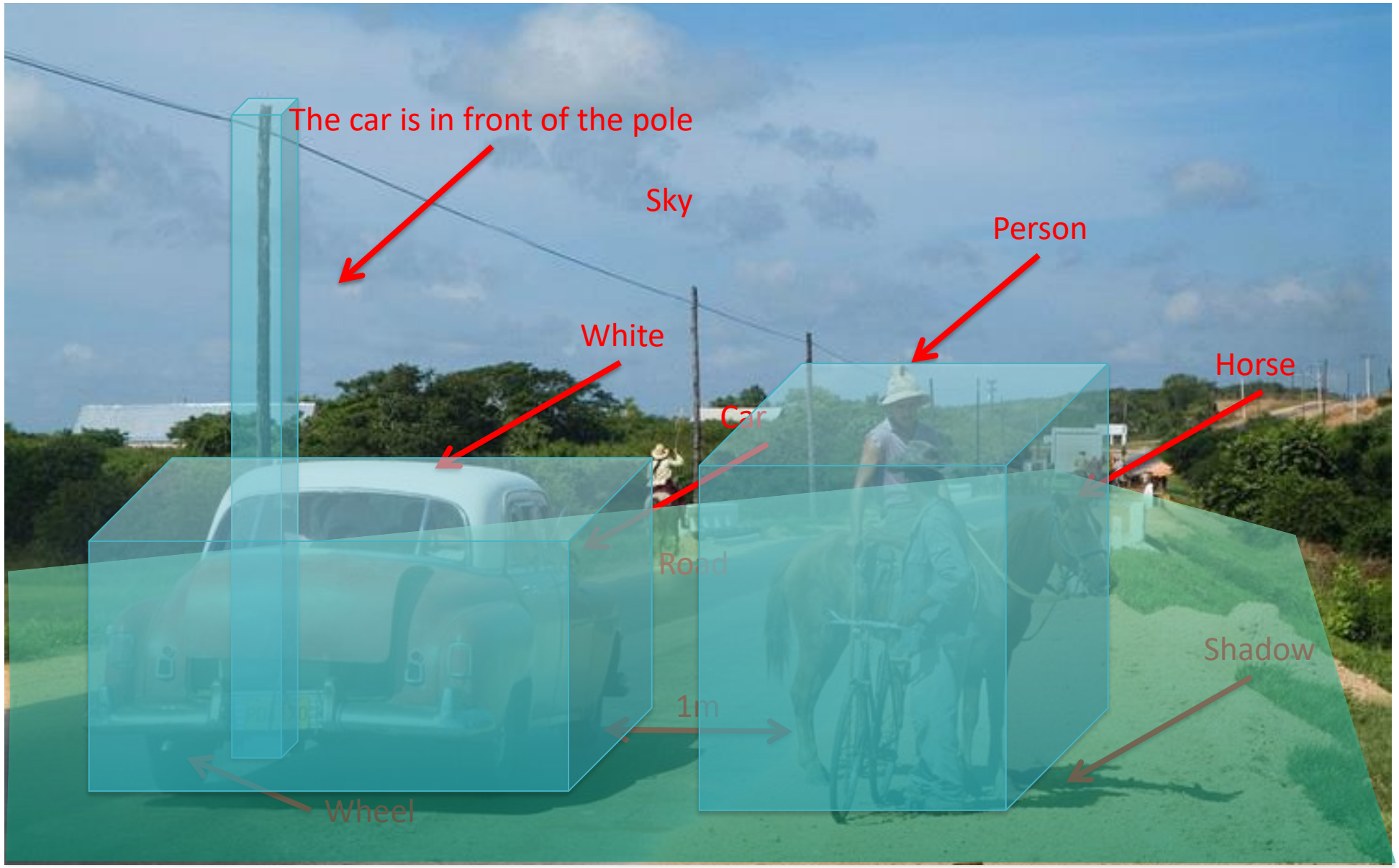
Seymour Papert

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 sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

What kind of information can be extracted from an image?

- Metric 3D information
-


0 3 2 5 4 7 6 9 8 0 3 2 5 4 7 6 9 8 0 3 2 5 4 7 6 9 8 0 3 2 5 4 7 6 9 8 0 3 2 5 4 7 6 9 8
3 0 1 2 3 4 5 6 7 3 0 1 2 3 4 5 6 7 3 0 1 2 3 4 5 6 7 3 0 1 2 3 4 5 6 7 3 0 1 2 3 4 5 6 7
2 1 0 3 2 5 4 7 6 2 1 0 3 2 5 4 7 6 2 1 0 3 2 5 4 7 6 2 1 0 3 2 5 4 7 6 2 1 0 3 2 5 4 7 6
5 2 3 0 1 2 3 4 5 5 2 3 0 1 2 3 4 5 5 2 3 0 1 2 3 4 5 5 2 3 0 1 2 3 4 5 5 2 3 0 1 2 3 4 5
4 3 2 1 0 3 2 5 4 4 3 2 1 0 3 2 5 4 4 3 2 1 0 3 2 5 4 4 3 2 1 0 3 2 5 4 4 3 2 1 0 3 2 5 4
7 4 5 2 3 0 1 2 3 7 4 5 2 3 0 1 2 3 7 4 5 2 3 0 1 2 3 7 4 5 2 3 0 1 2 3 7 4 5 2 3 0 1 2 3
6 5 4 3 2 1 0 3 2 6 5 4 3 2 1 0 3 2 6 5 4 3 2 1 0 3 2 6 5 4 3 2 1 0 3 2 6 5 4 3 2 1 0 3 2
9 6 7 4 5 2 3 0 1 9 6 7 4 5 2 3 0 1 9 6 7 4 5 2 3 0 1 9 6 7 4 5 2 3 0 1 9 6 7 4 5 2 3 0 1
8 7 6 5 4 3 2 1 0 8 7 6 5 4 3 2 1 0 8 7 6 5 4 3 2 1 0 8 7 6 5 4 3 2 1 0 8 7 6 5 4 3 2 1 0
0 3 2 5 4 7 6 9 8 0 3 2 5 4 7 6 9 8 0 3 2 5 4 7 6 9 8 0 3 2 5 4 7 6 9 8 0 3 2 5 4 7 6 9 8
3 0 1 2 3 4 5 6 7 3 0 1 2 3 4 5 6 7 3 0 1 2 3 4 5 6 7 3 0 1 2 3 4 5 6 7 3 0 1 2 3 4 5 6 7
2 1 0 3 2 5 4 7 6 2 1 0 3 2 5 4 7 6 2 1 0 3 2 5 4 7 6 2 1 0 3 2 5 4 7 6 2 1 0 3 2 5 4 7 6
5 2 3 0 1 2 3 4 5 5 2 3 0 1 2 3 4 5 5 2 3 0 1 2 3 4 5 5 2 3 0 1 2 3 4 5 5 2 3 0 1 2 3 4 5
4 3 2 1 0 3 2 5 4 4 3 2 1 0 3 2 5 4 4 3 2 1 0 3 2 5 4 4 3 2 1 0 3 2 5 4 4 3 2 1 0 3 2 5 4
7 4 5 2 3 0 1 2 3 7 4 5 2 3 0 1 2 3 7 4 5 2 3 0 1 2 3 7 4 5 2 3 0 1 2 3 7 4 5 2 3 0 1 2 3
6 5 4 3 2 1 0 3 2 6 5 4 3 2 1 0 3 2 6 5 4 3 2 1 0 3 2 6 5 4 3 2 1 0 3 2 6 5 4 3 2 1 0 3 2
9 6 7 4 5 2 3 0 1 9 6 7 4 5 2 3 0 1 9 6 7 4 5 2 3 0 1 9 6 7 4 5 2 3 0 1 9 6 7 4 5 2 3 0 1
8 7 6 5 4 3 2 1 0 8 7 6 5 4 3 2 1 0 8 7 6 5 4 3 2 1 0 8 7 6 5 4 3 2 1 0 8 7 6 5 4 3 2 1 0

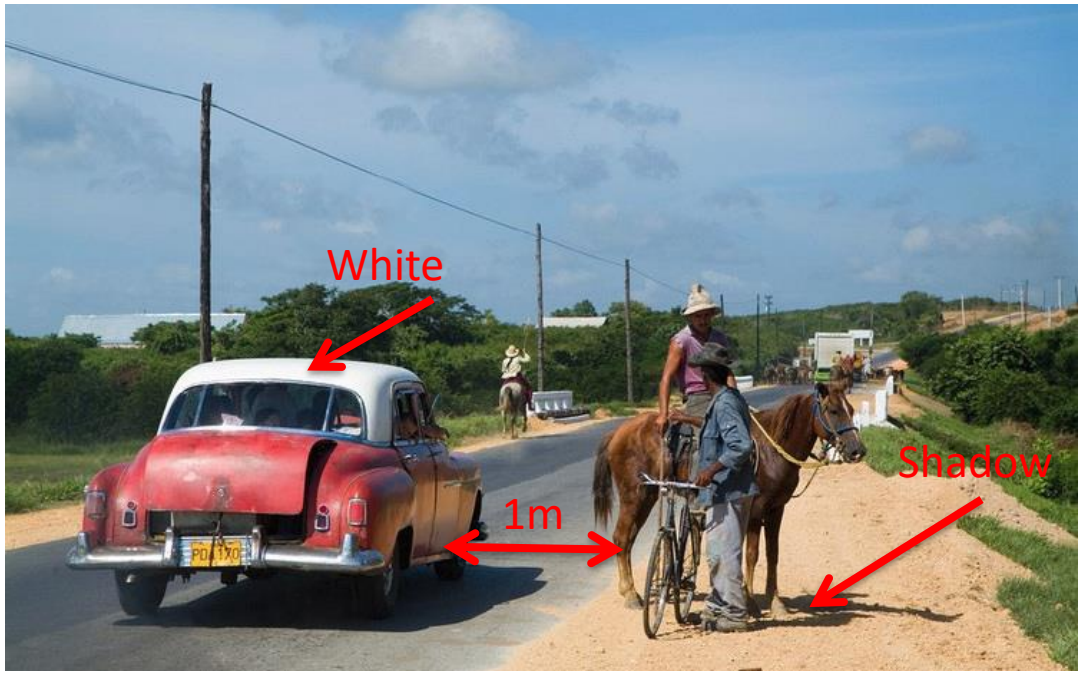

The car is in front of the pole
Sky
Person
White
Horse
Car
Road
Shadow
1m
Wheel

Computer Vision

- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings

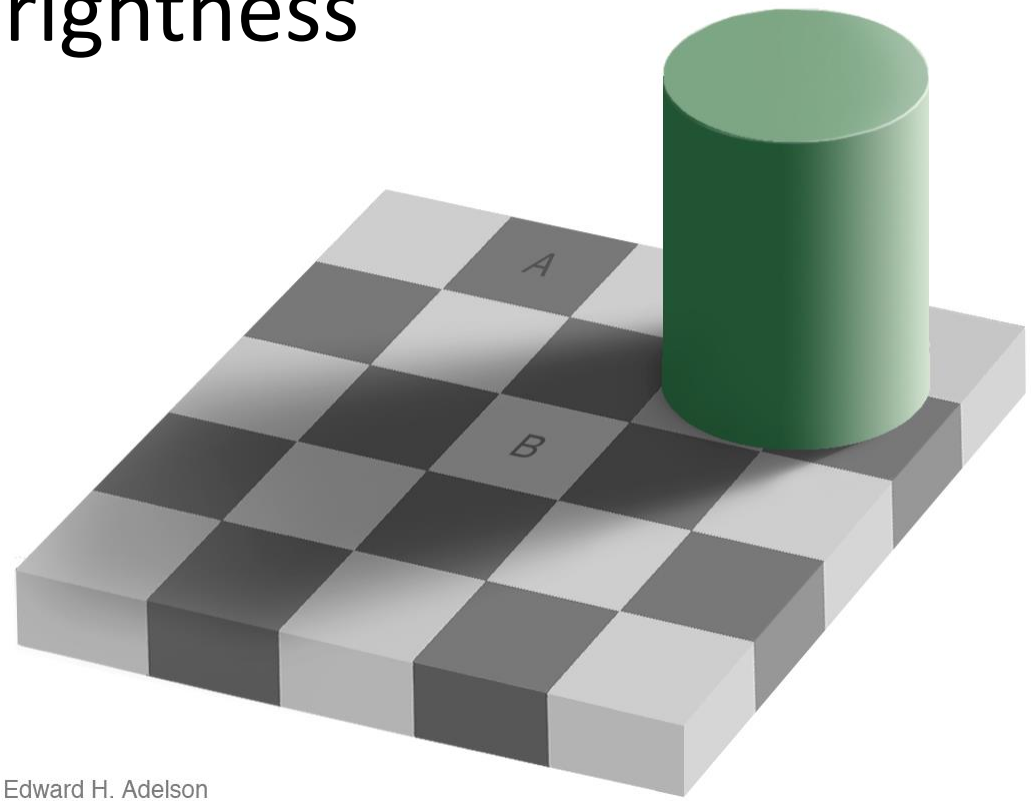
Computer Vision

- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings



Measurement

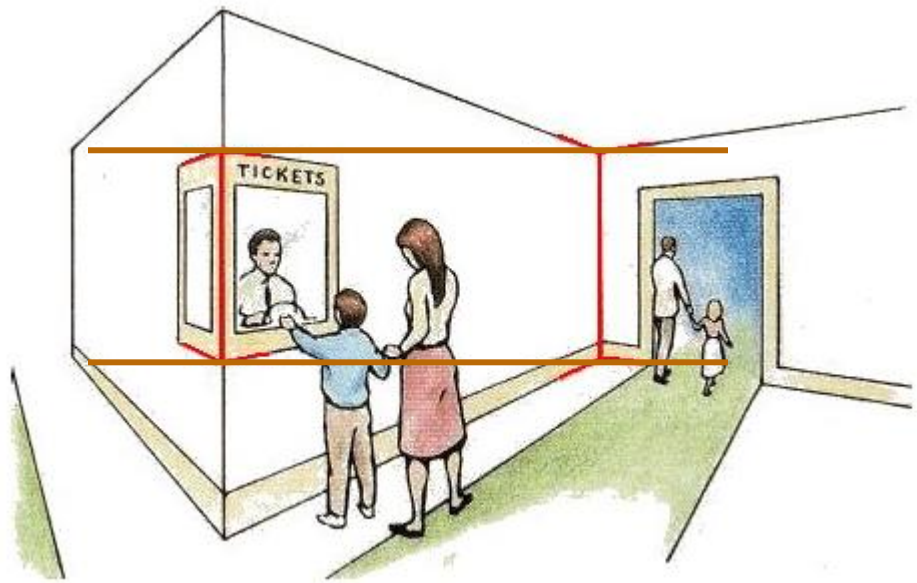
Brightness



Edward H. Adelson

Measurement

Length



http://www.michaelbach.de/ot/sze_muelue/index.html

Image Enhancement

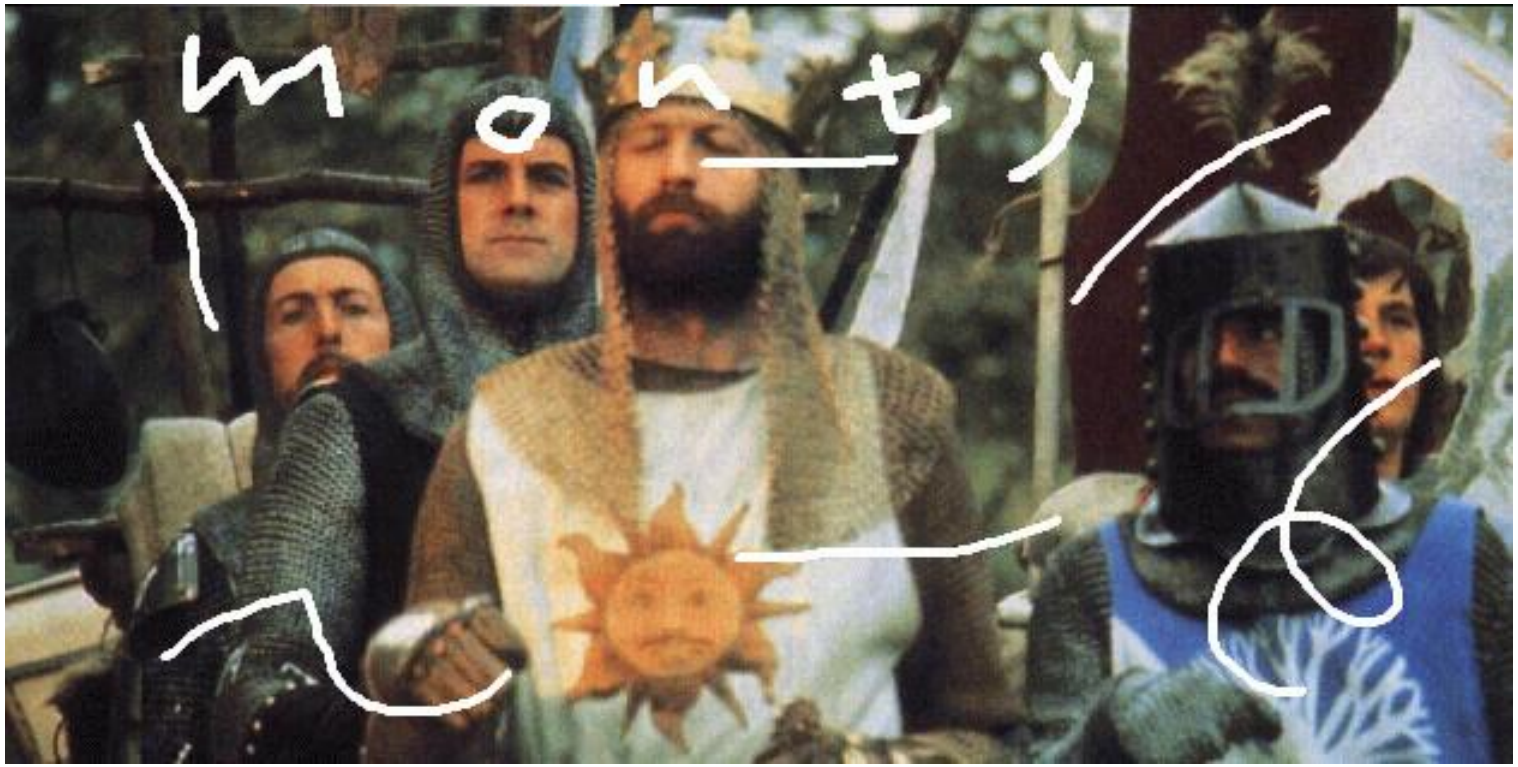
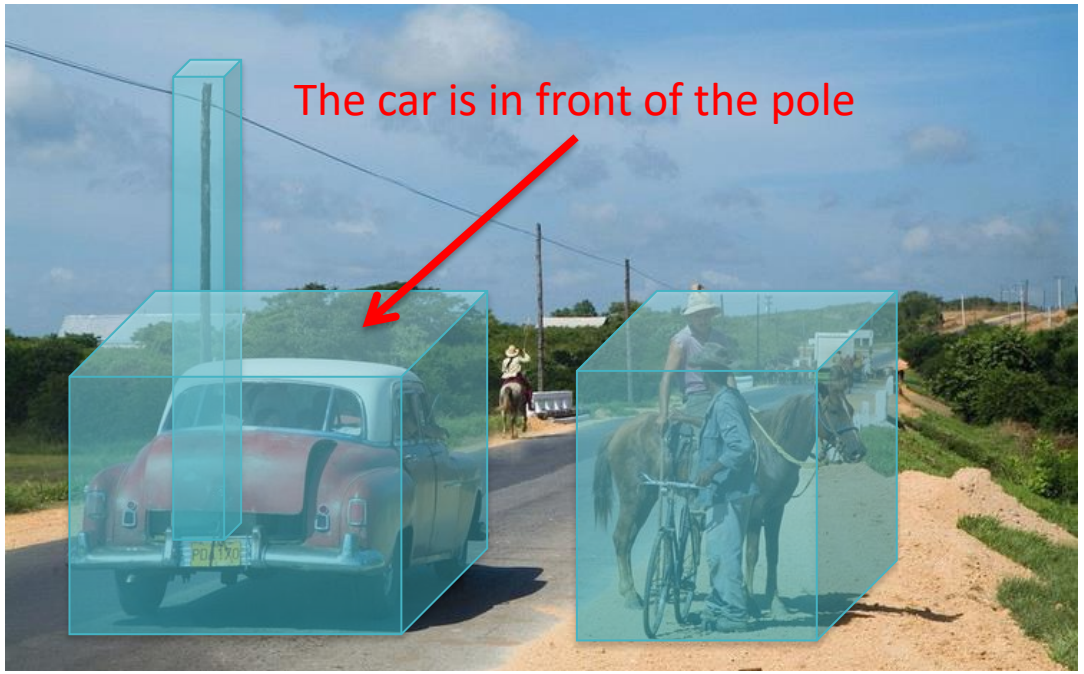


Image Enhancement

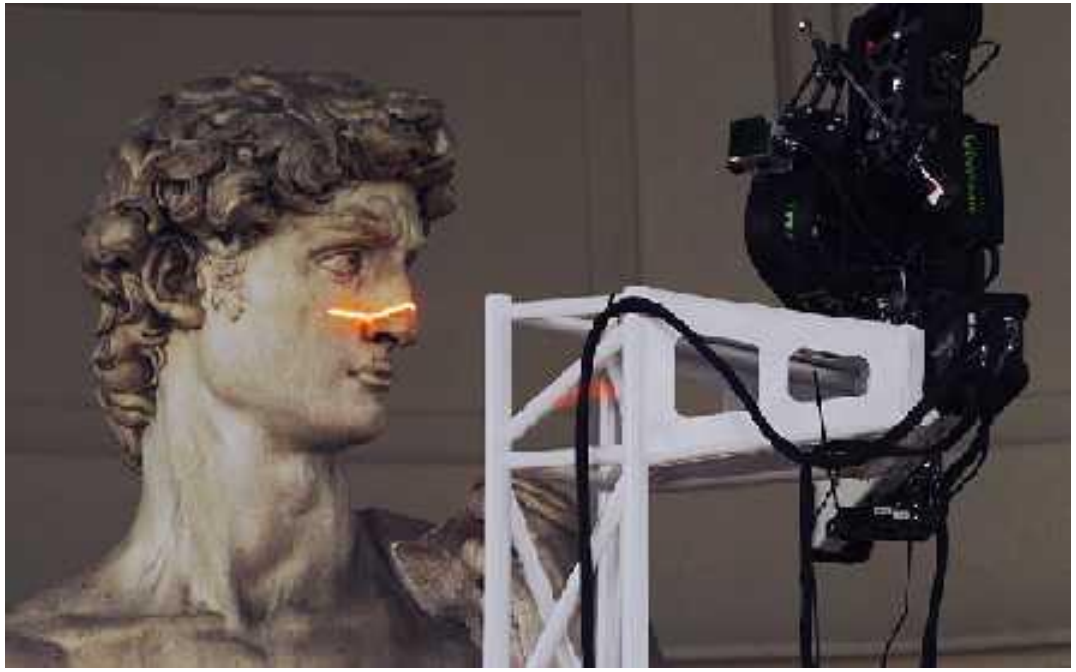
Image Inpainting, M. Bertalmío et al.
http://www.iua.upf.es/~mbertalmio//restoration.html

Computer Vision

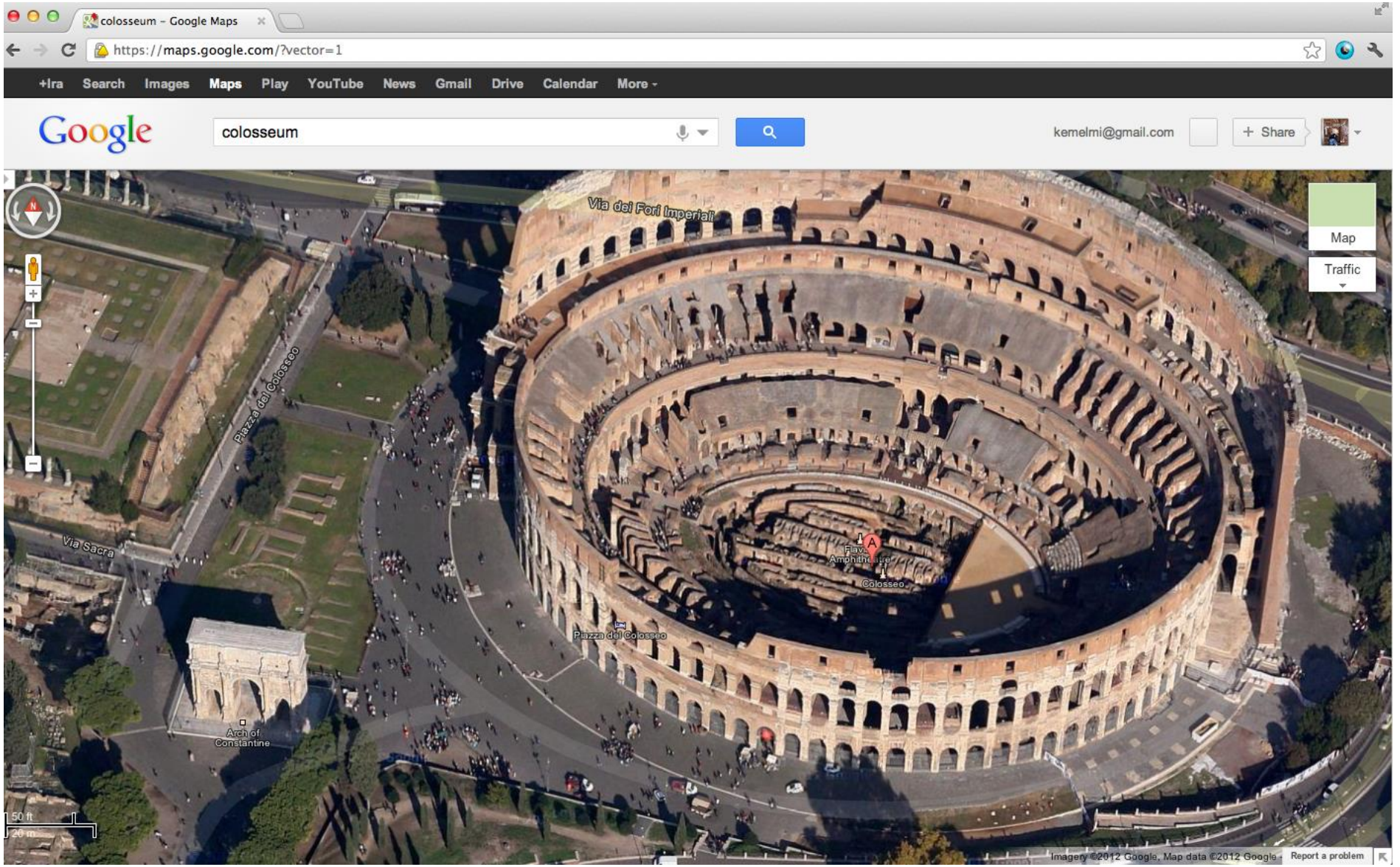
- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings



Applications: 3D Scanning

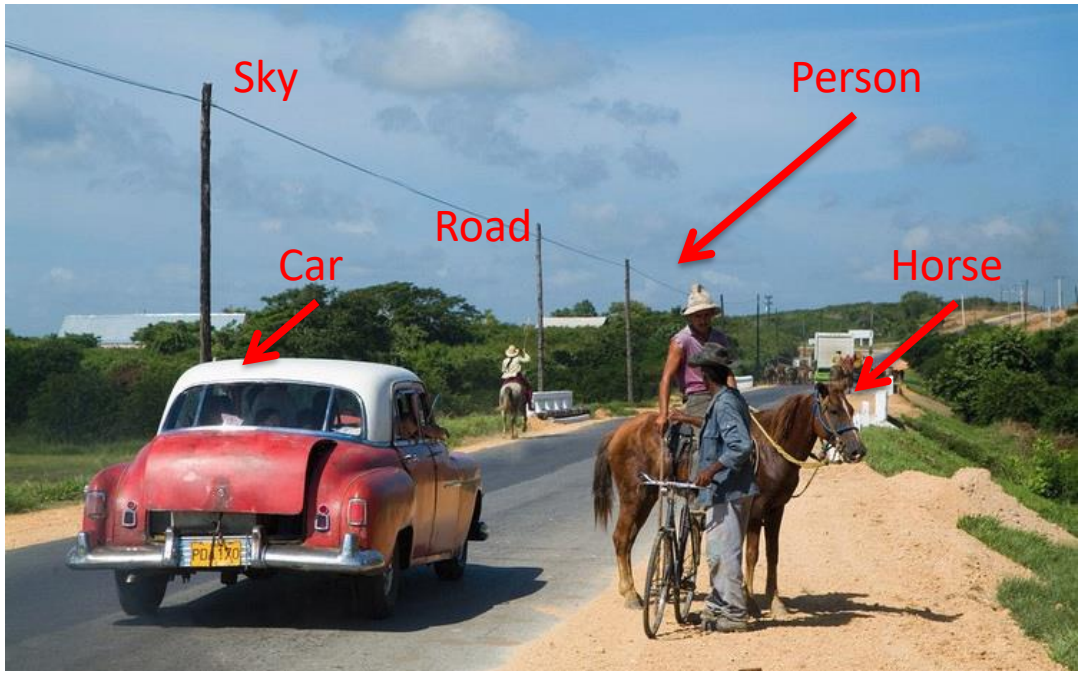


Google's 3D Maps
Structure estimation from tourist photos

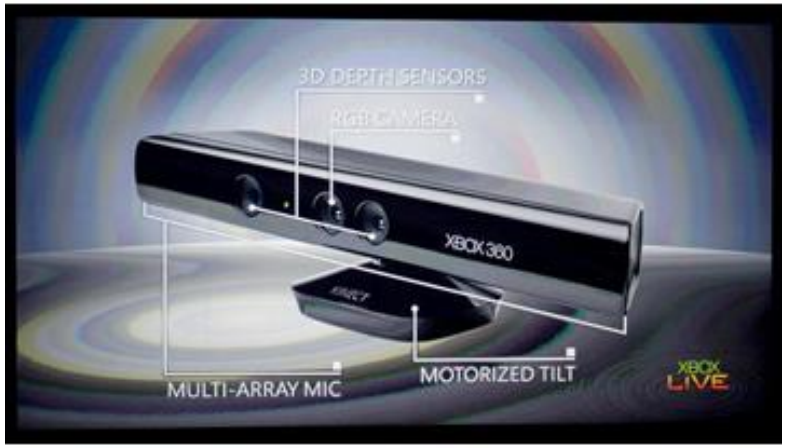


Computer Vision

- Low Level Vision
  - Measurements
  - Enhancements
  - Region segmentation
  - Features
- Mid Level Vision
  - Reconstruction
  - Depth
  - Motion Estimation
- High Level Vision
  - Category detection
  - Activity recognition
  - Deep understandings
  - Pose estimation



Vision-based interaction: Xbox Kinect

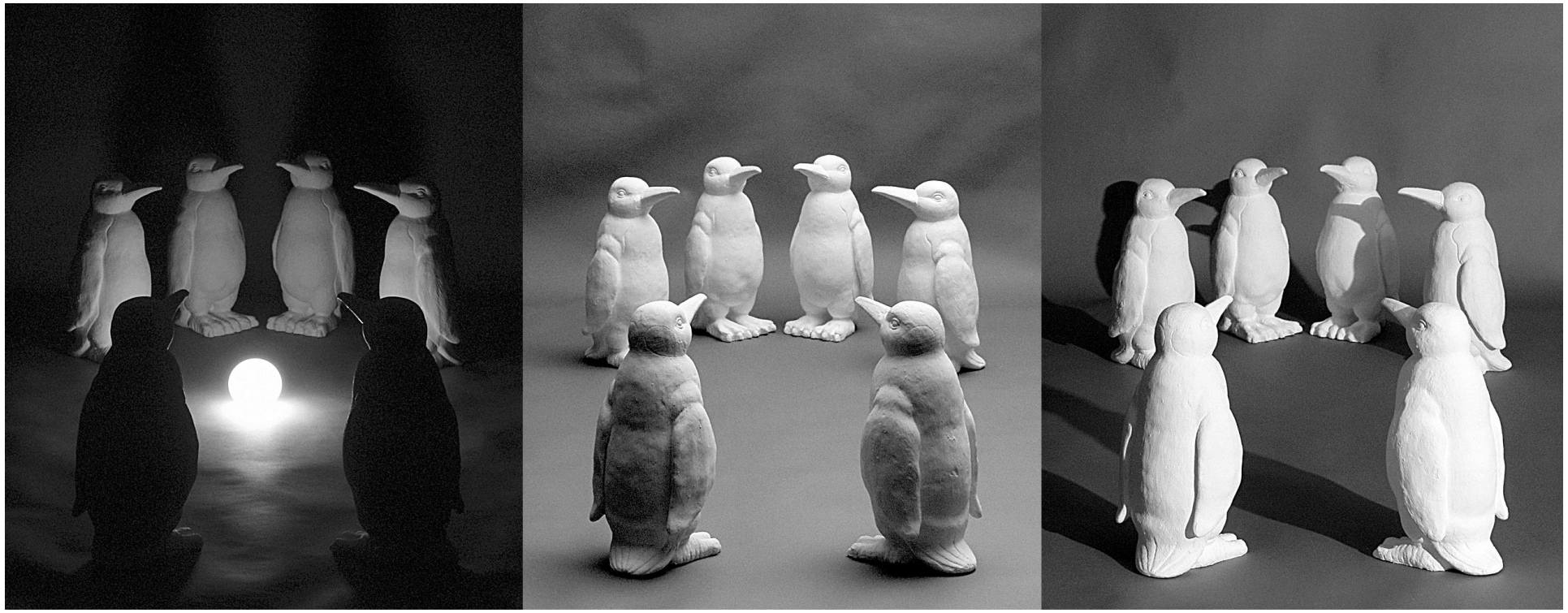


Why is computer vision difficult?

Challenges: viewpoint variation

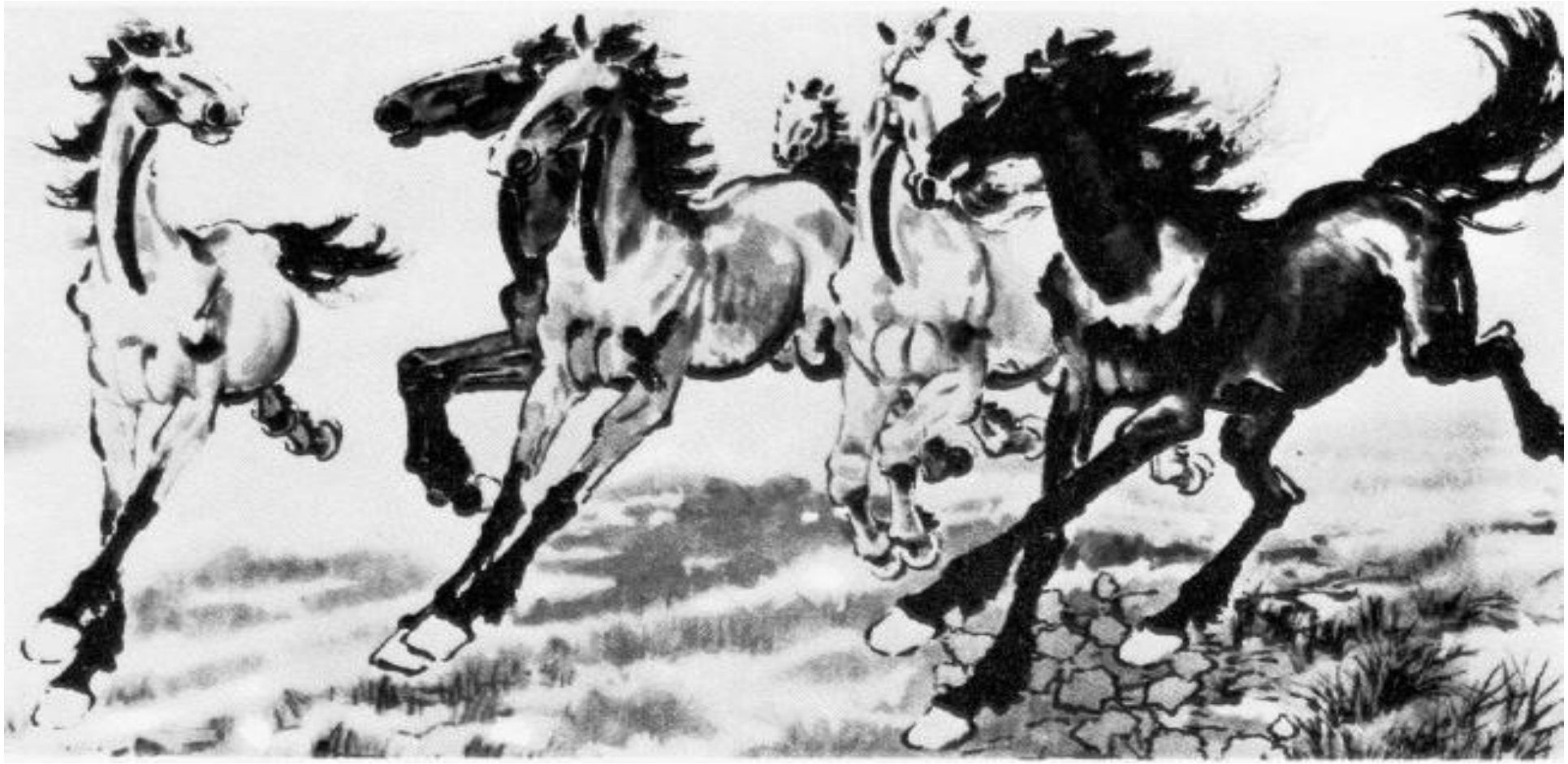
Michelangelo 1475-1564

Challenges: illumination

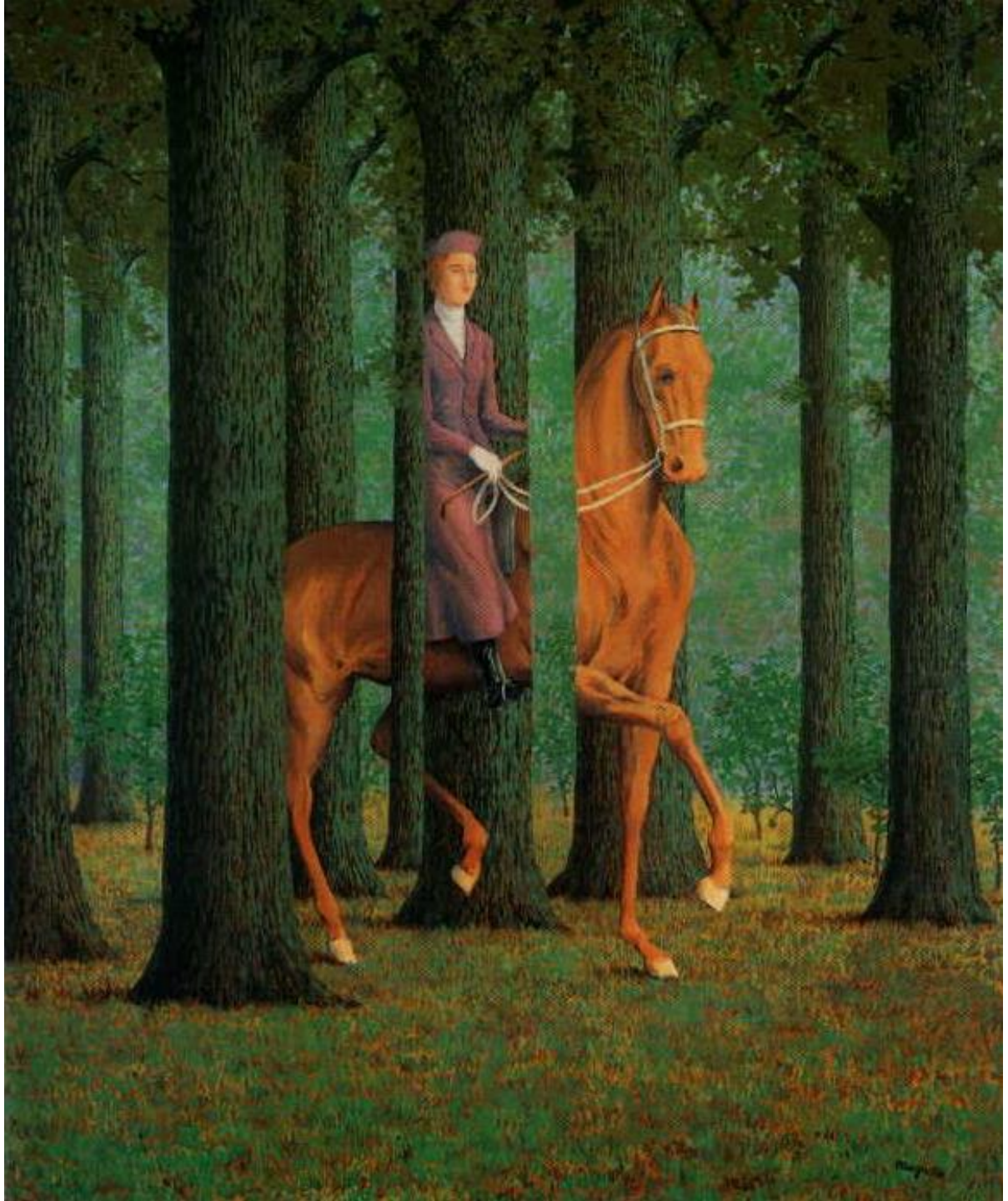


Challenges: scale

Challenges: deformation



Challenges: occlusion

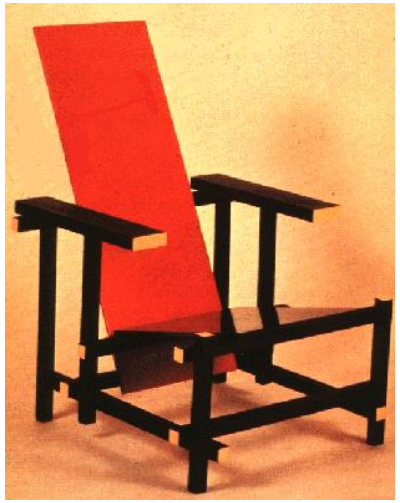


Magritte, 1957

Challenges: background clutter

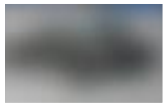
Challenges: Motion

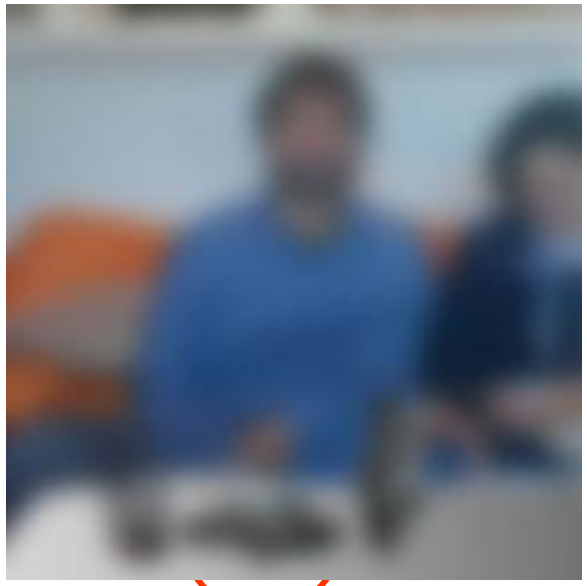
Challenges: object intra-class variation



slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity





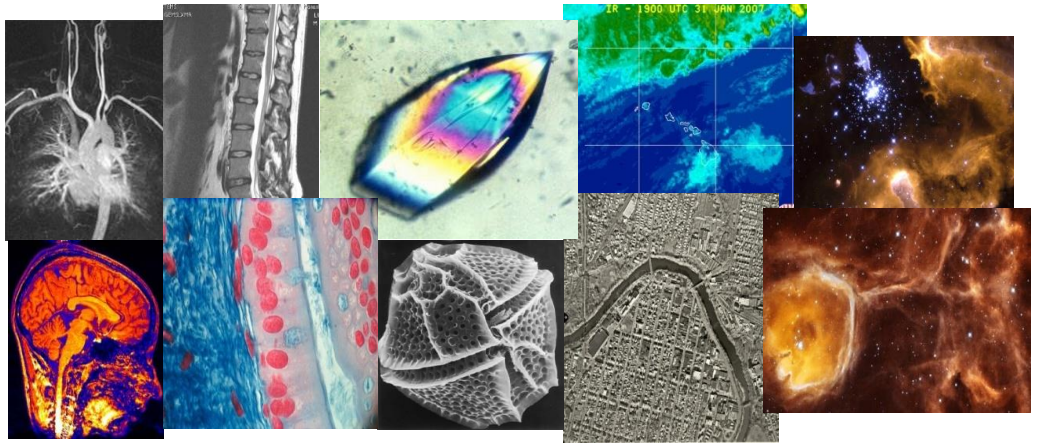
Challenges or opportunities?

- Images are confusing, but they also reveal the structure of the world through numerous cues
- Our job is to interpret the cues!

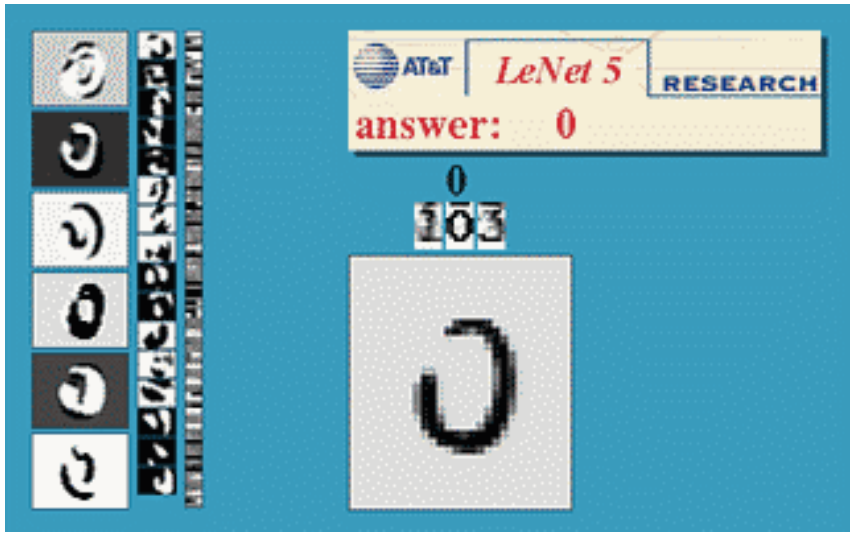
Why study computer vision?

Vision is useful: Images and video are everywhere!

Google Image Search™ Picasa™ flickr™ facebook YouTube Broadcast Yourself™



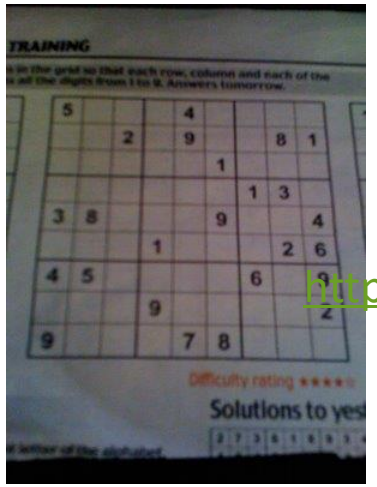
Optical character recognition (OCR)



Digit recognition
yann.lecun.com

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Automatic check processing



Sudoku grabber
http://sudokugrab.blogspot.com/

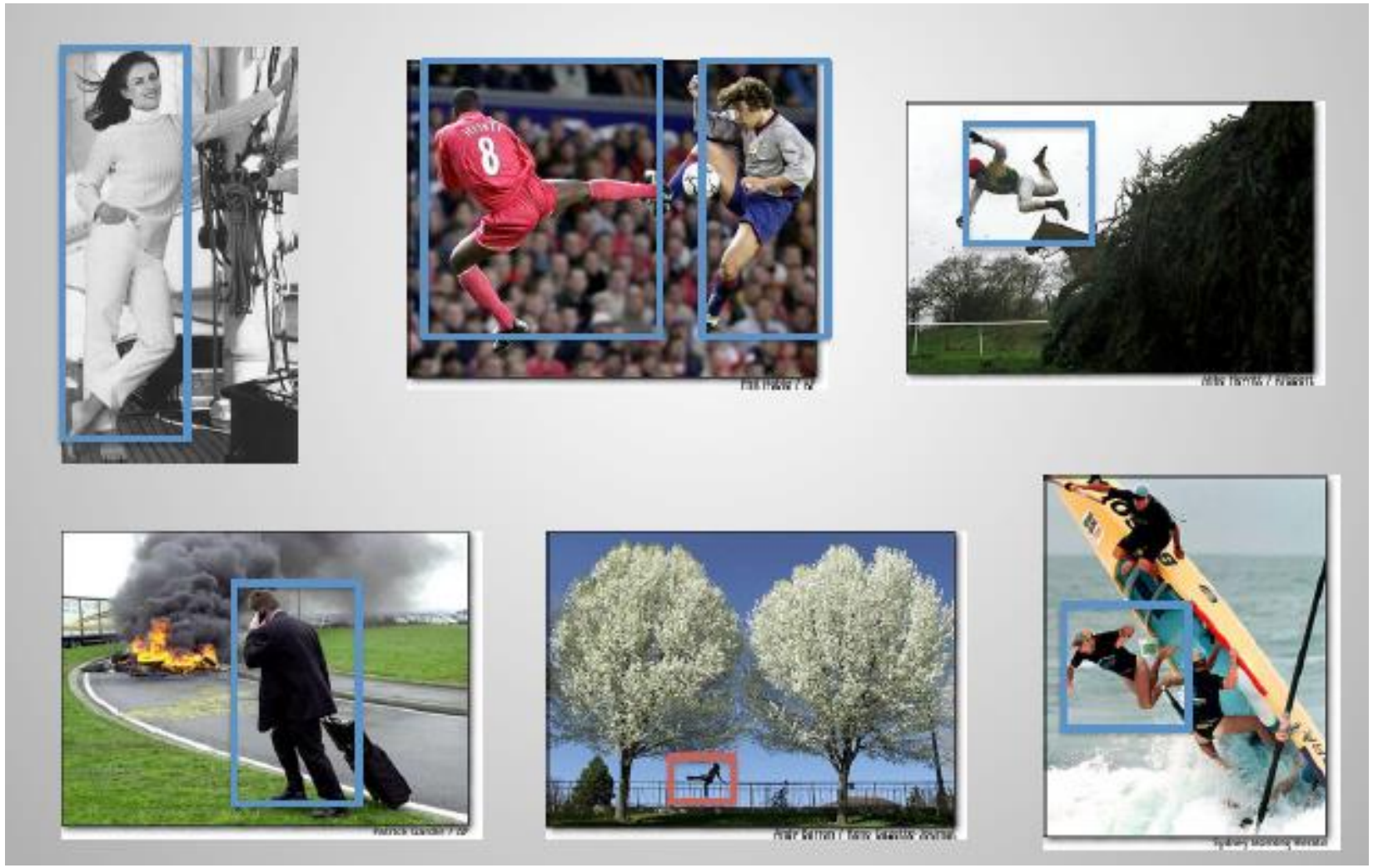
Ex. Object Recognition

- Problem: Given an image A, does A contain an image of a person?

Ex. Object Recognition

- Problem: Given an image A, does A contain an image of a person?

Ex. Object localization



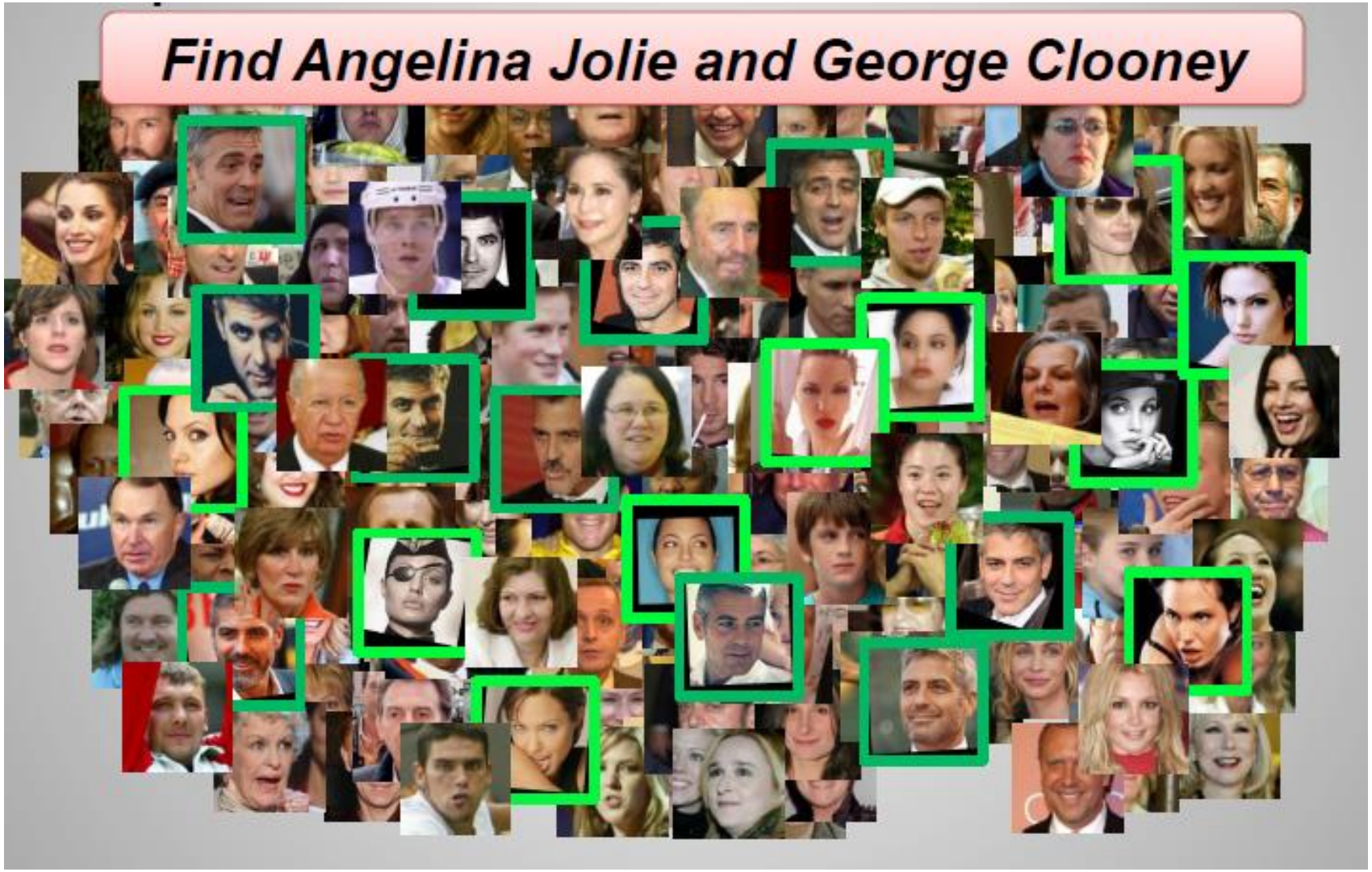
Ex. Human Detection

Ex: Face Detection: Apple iPhoto, Facebook, Google..

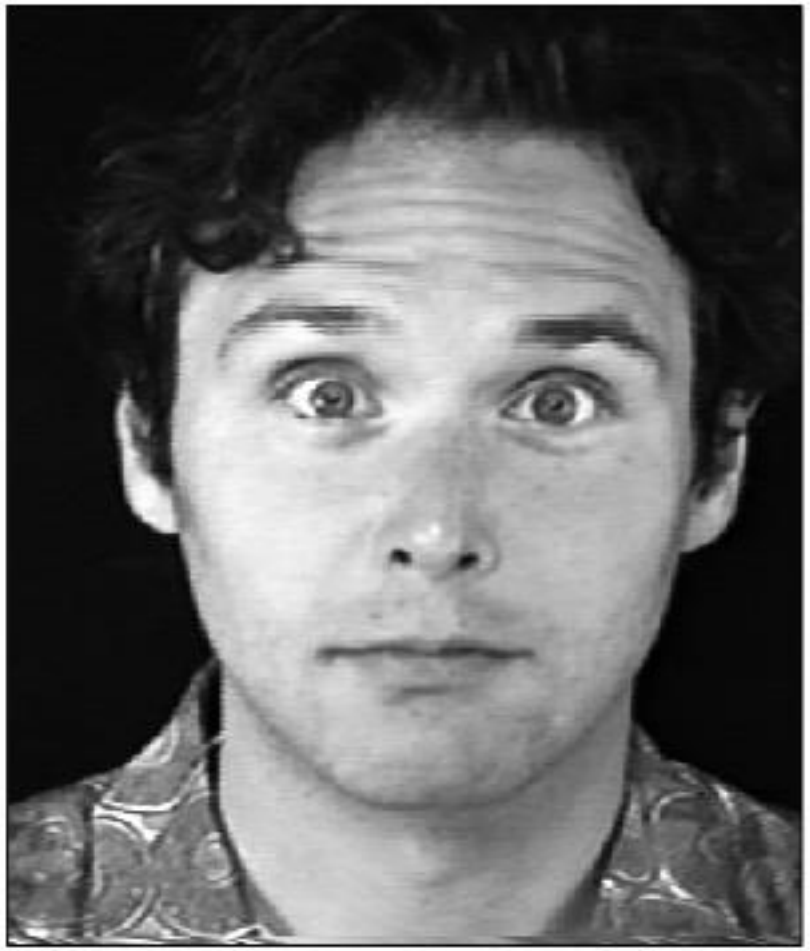
Ex. Face Recognition

Open-Universe Face Identification

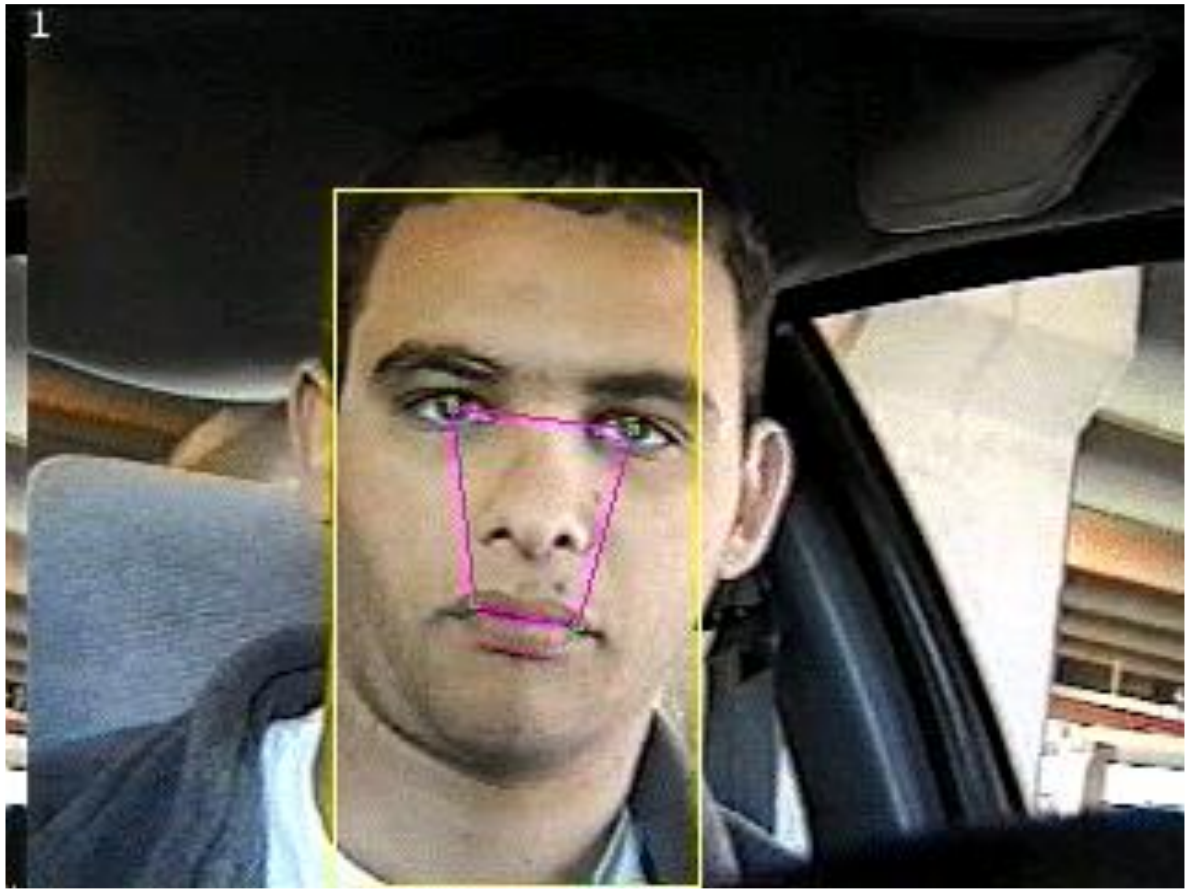
Open-Universe Face Identification



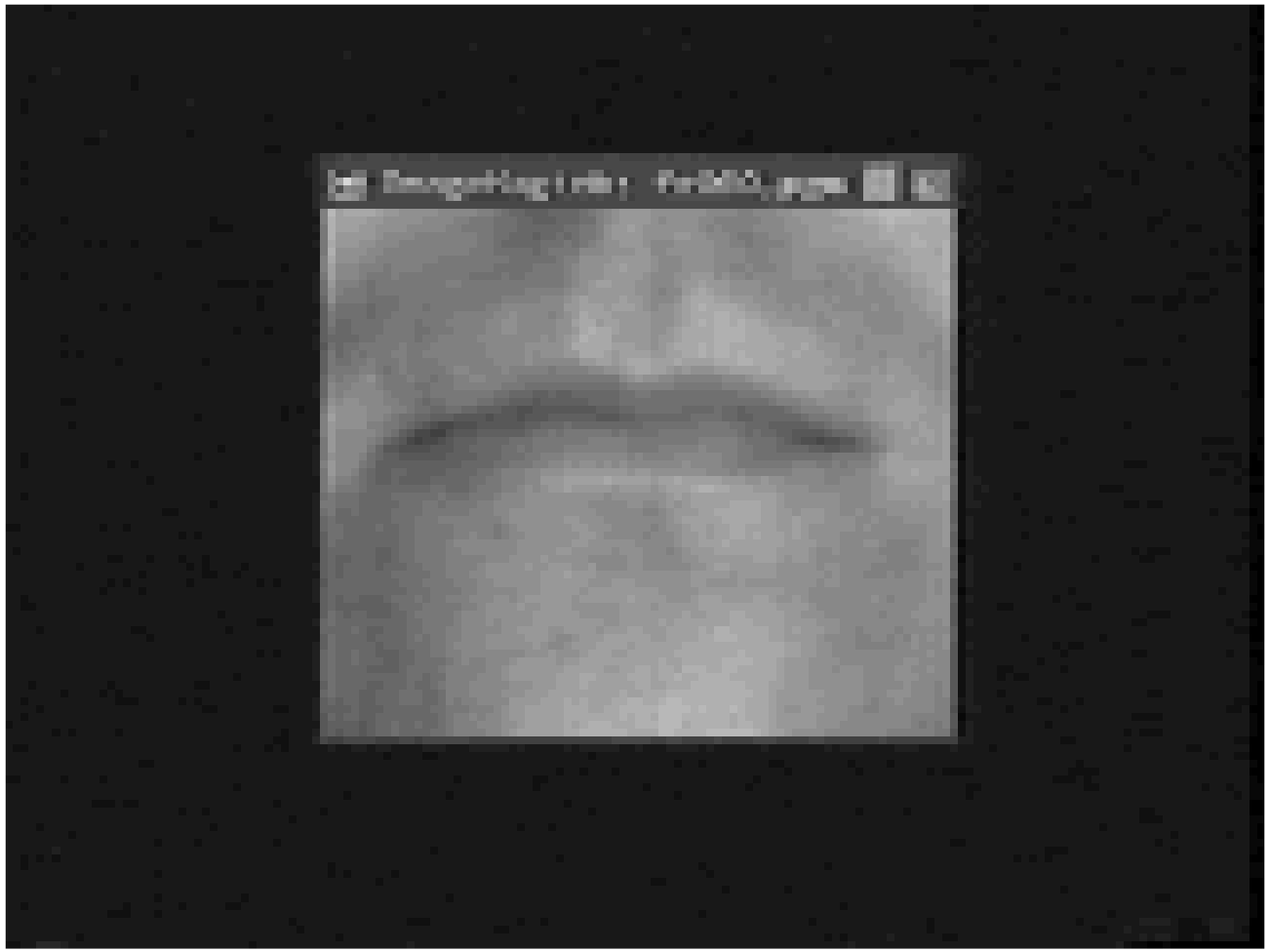
Ex. Facial expression



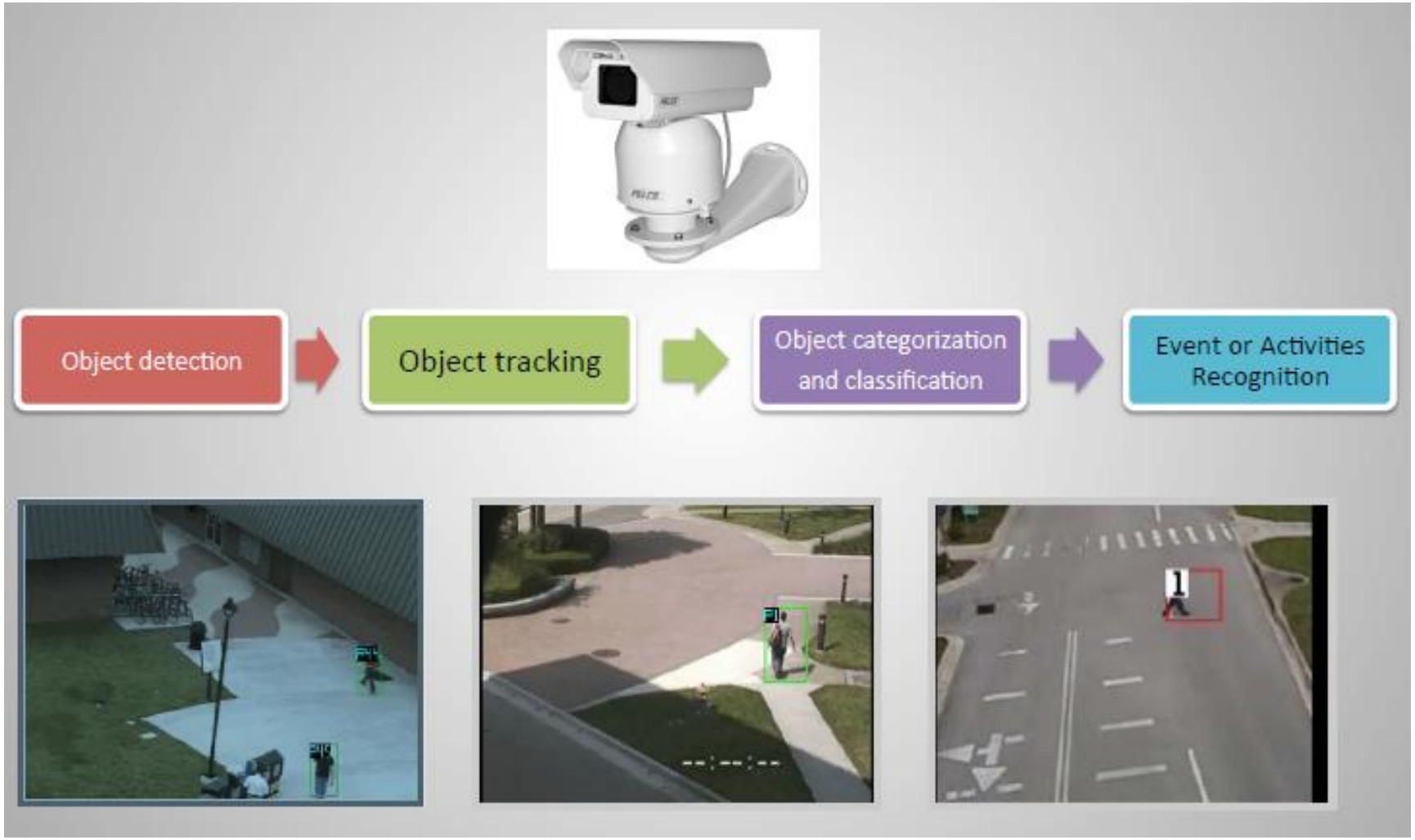
Ex. Fatigue detection



Ex. Lip-reading

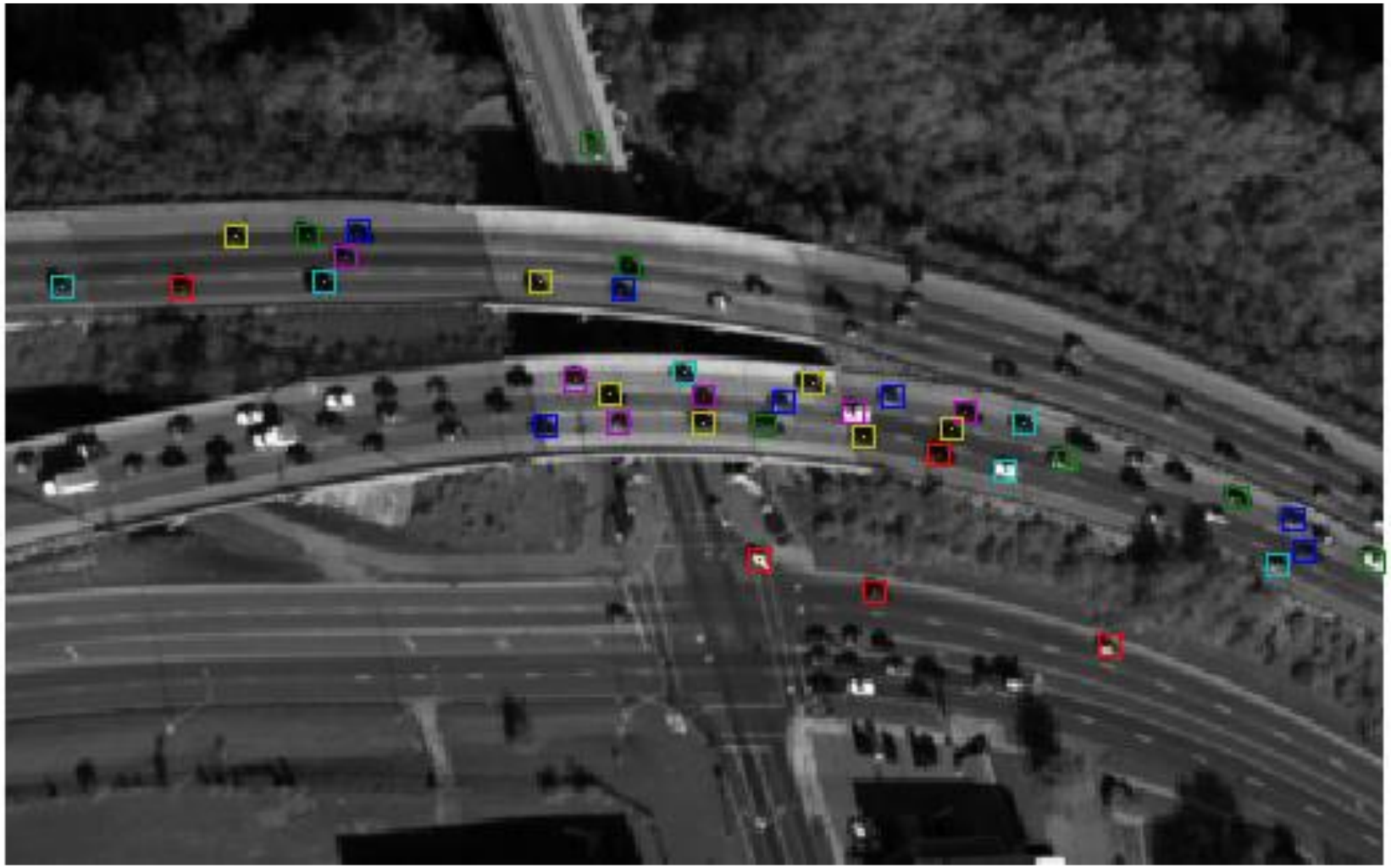


Video Surveillance and Monitoring

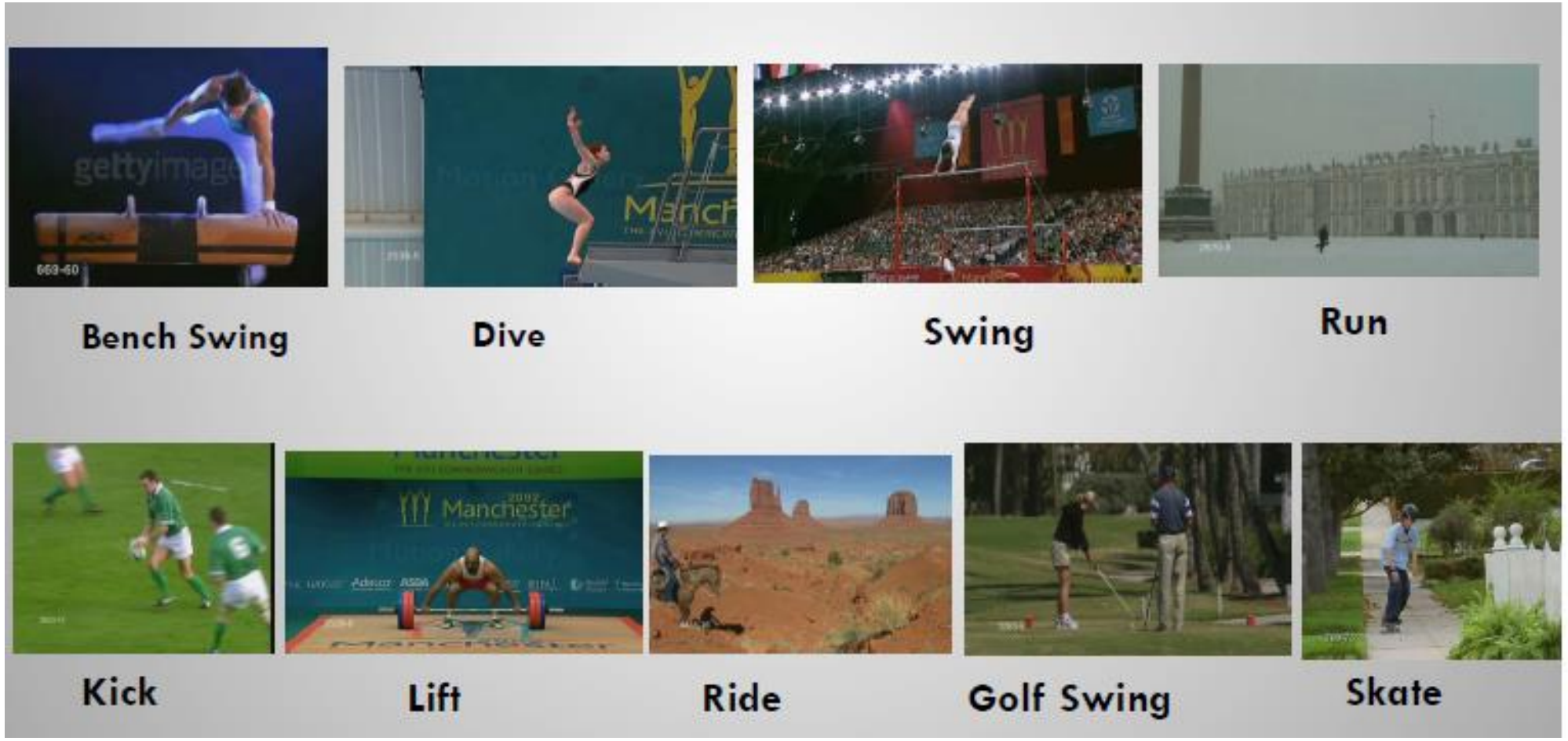


UAVs: Unmanned Aerial Vehicles (drones)

Ex. Tracking (multi-object)



Ex. Human Action Recognition



Ex. Counting in Extremely Dense Crowd Images

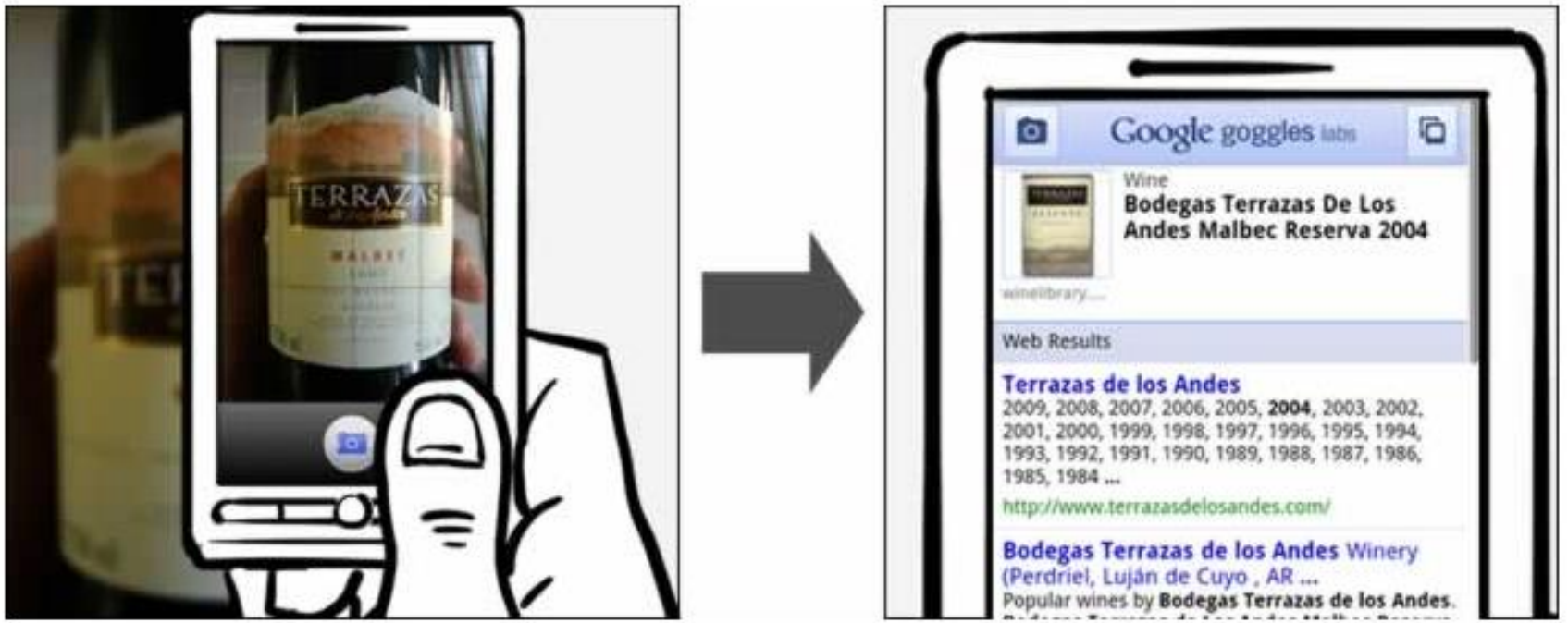
Ground truth=634 Proposed Method by Idrees and Shah=640

Mobile visual search: Google Goggles

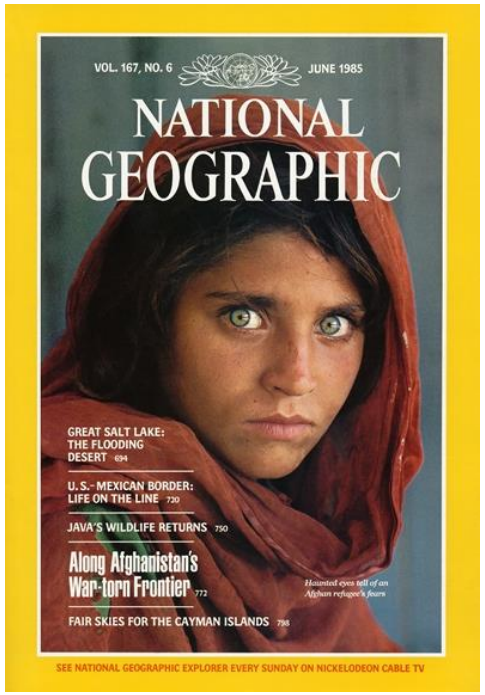
Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.

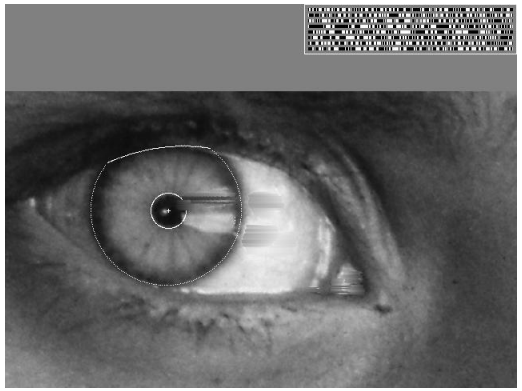
Google goggles labs

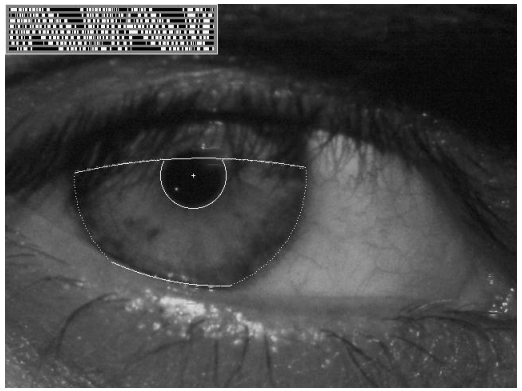


Biometrics



[How the Afghan Girl was Identified by Her Iris Patterns]



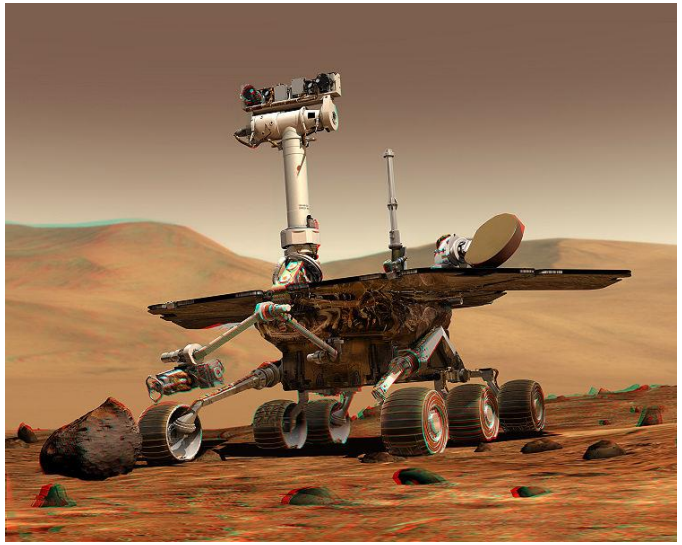


Automotive safety

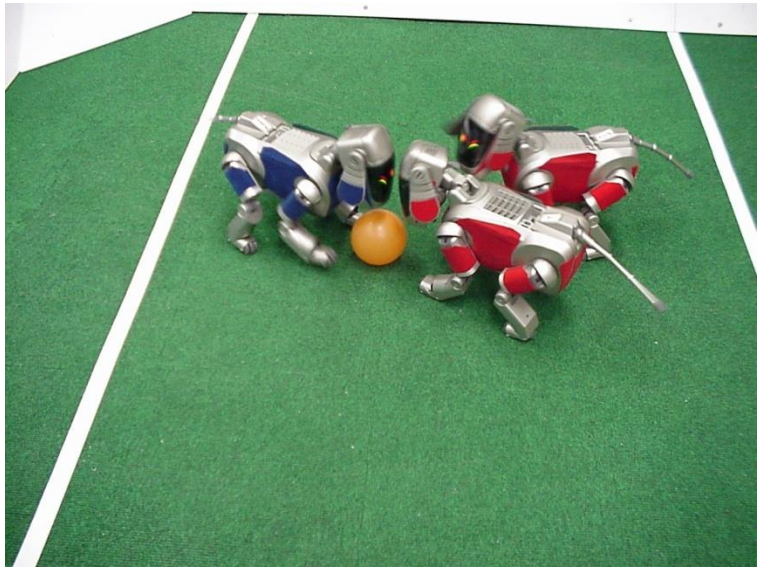
- Mobileye: Vision systems in high-end BMW, GM, Volvo models
  - Pedestrian collision warning
  - Forward collision warning
  - Lane departure warning
  - Headway monitoring and warning

AutoCars - Uber

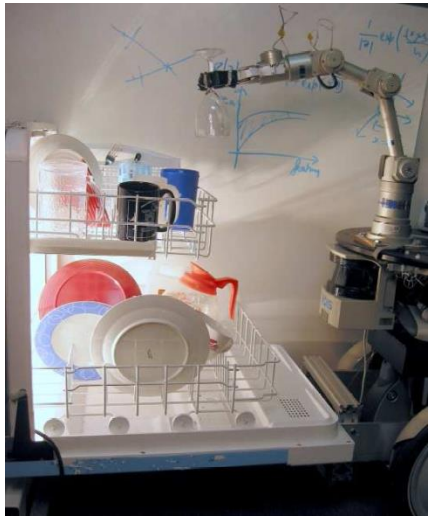
Mobile robots



NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



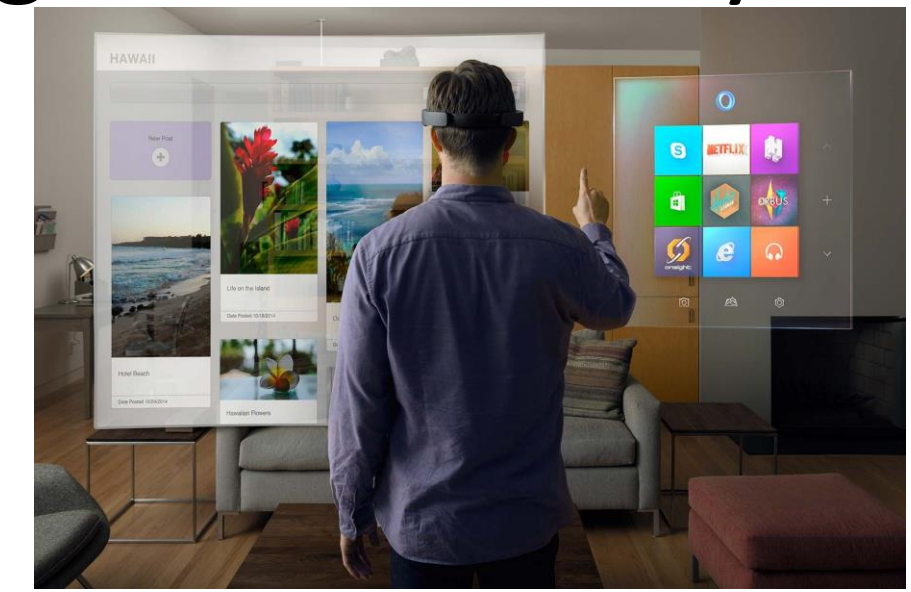
http://www.robocup.org/

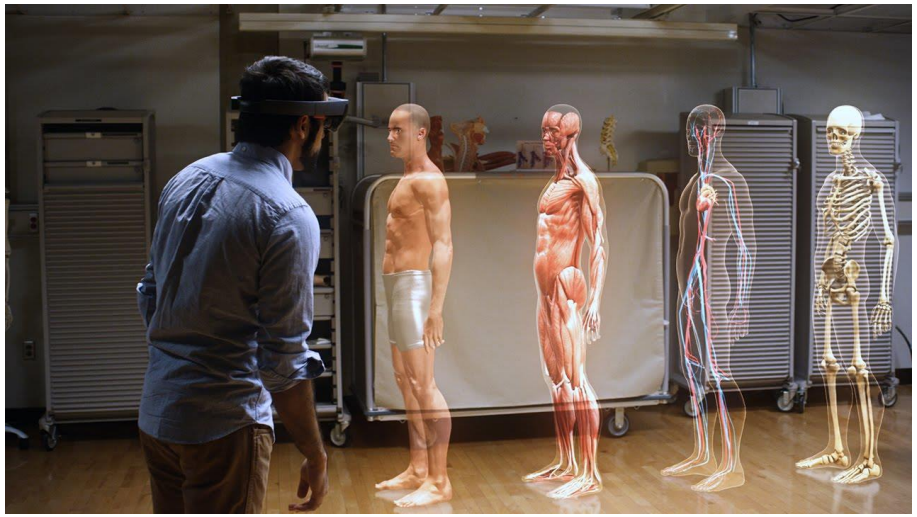


Saxena et al. 2008
STAIR at Stanford

amazon Prime Air

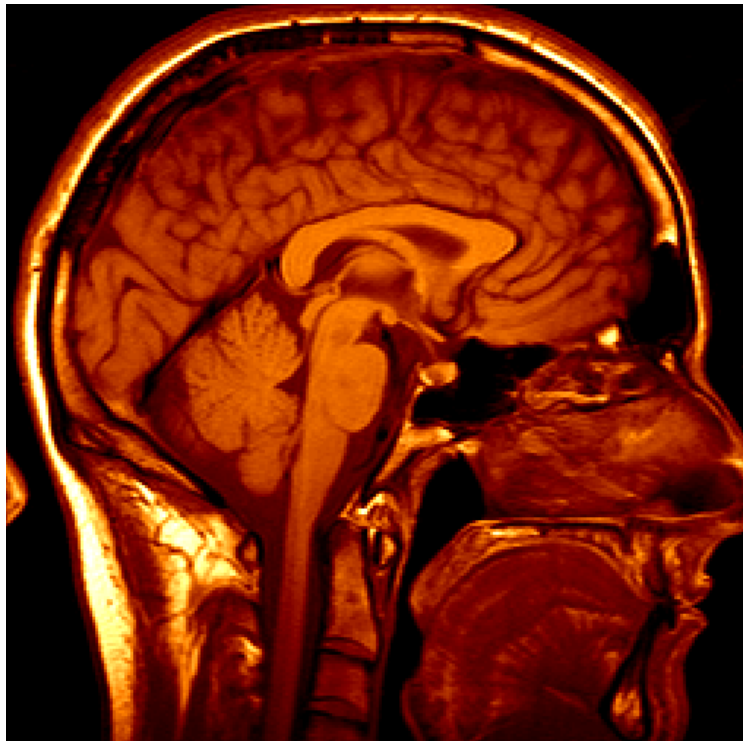
Augmented Reality and Virtual Reality





MS HoloLens, Oculus, Magic Leap,

Medical imaging



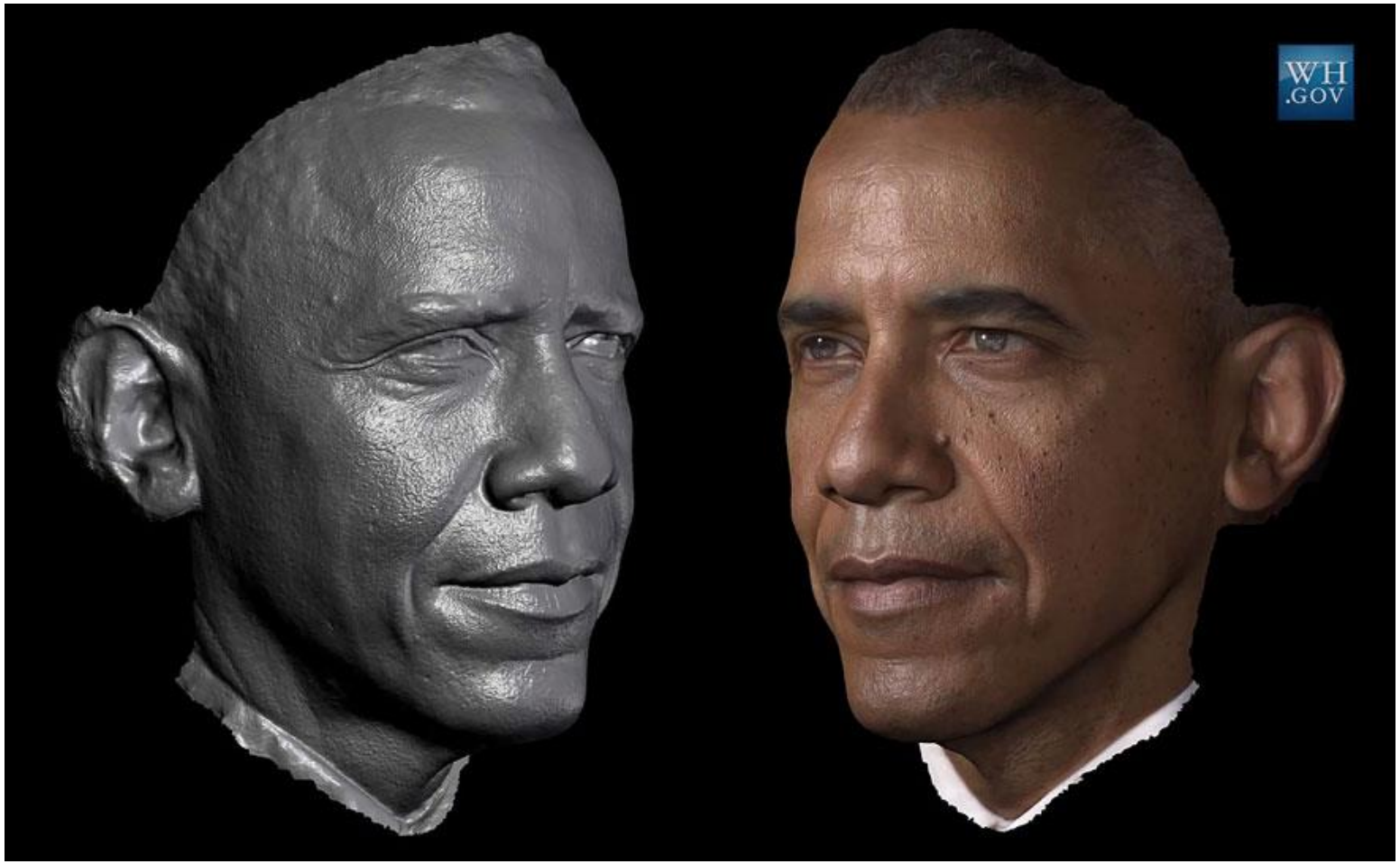
3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Human shape capture

Human shape capture

Human shape capture



Topics

- Filtering, Edge Finding, Transformations
- Color, Texture,
- Interest Points and Region Descriptors
- Segmentation- *Few advanced methods*
- Cameras, Stereo, Reconstruction
- Motion and Tracking
- Object Detection and Recognition
- Case study: CBIR/ Face detection & recognition

References

- Richard Szeliski, *Computer Vision: Algorithms and Applications* (available online) http://szeliski.org/Book/
- Forsyth & Ponce, *Computer Vision: A Modern Approach*
- *Gonzalez, "Digital Image Processing"*
- Relevant research papers