ECSE-626

Statistical Computer Vision

Introductory Lecture

Background Requirements

- Designed for graduate students pursuing interests in:
  - Computer vision
  - Robot vision
  - Medical imaging
  - Artificial intelligence (e.g. machine learning, decision-making)
- Not meant as a general introduction to computer vision

Background Requirements

- Describe and address wide range of problems in computer vision
- Focus on solutions based on probability and information theory

Background Requirements

- Assumes mathematical background (basic) in:
  - Probability and statistics (basic – course will provide most of the background required)
  - Linear algebra
  - Calculus
  - Image processing (basic)
  - Optimization (basic)
- Assumes student comfortable programming in Matlab

Course Motivation

- Computer Vision has always concerned itself with solving hard, ill-posed problems:

  Given measurements of a scene (e.g. from a camera, from a laser), what can we infer about the underlying process that resulted in these measurements?

Vision as a source of semantic information

slide credit: Fei-Fei, Fergus &

Object categorization

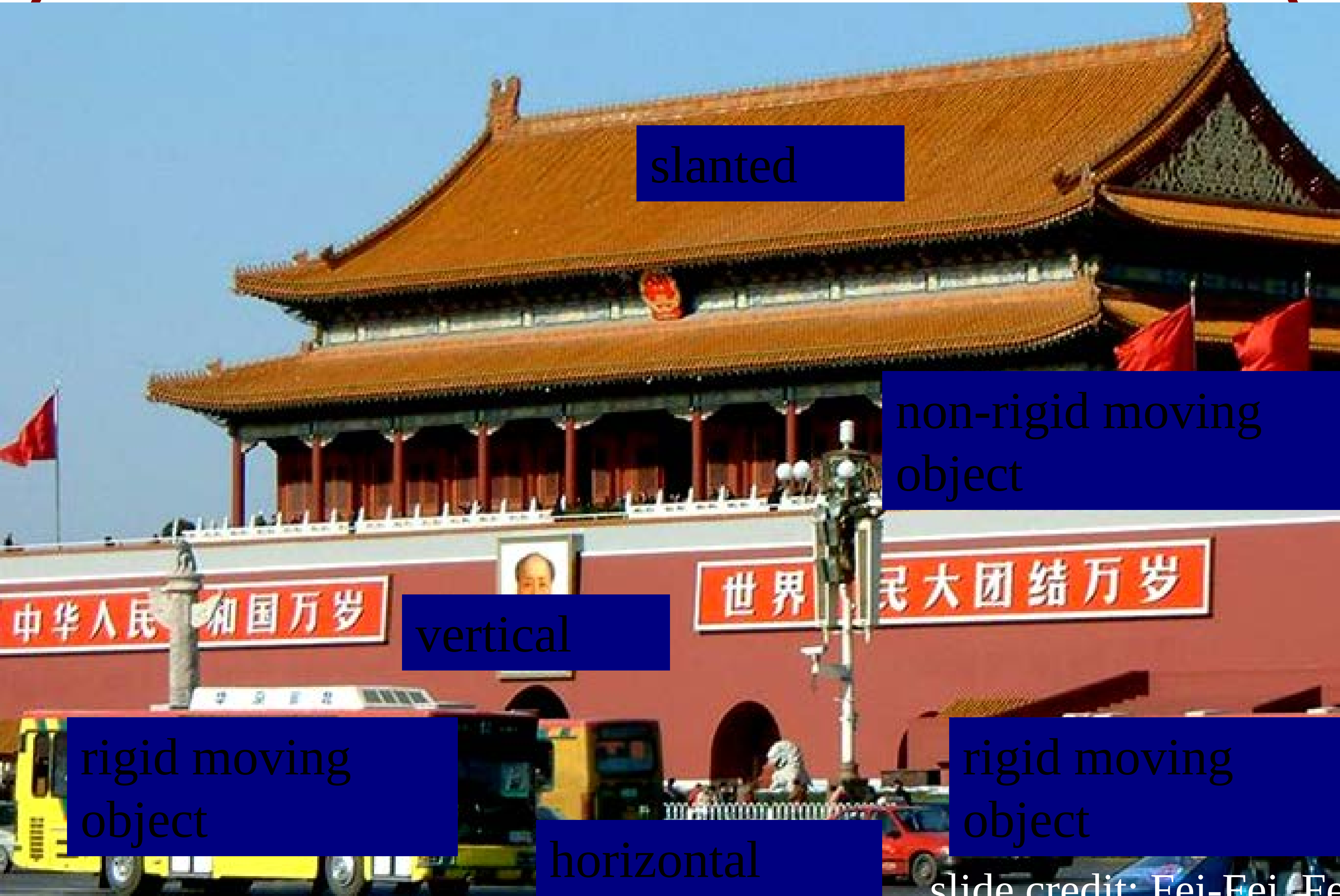
slide credit: Fei-Fei, Fergus &

Scene and context categorization

- outdoor
- city
- traffic
- …

slide credit: Fei-Fei, Fergus &

Qualitative spatial information



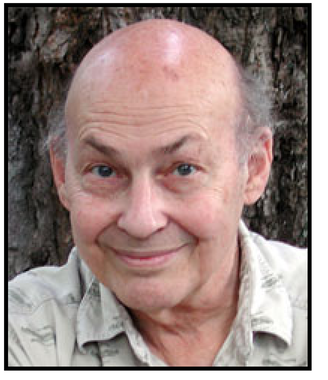
slide credit: Fei-Fei, Fe

Computer Vision

- These don't seem like hard problem. After all, people do it all the time!
- Human visual system evolved to be good at this task.

The Summer Project

- In 1966, Minsky hired a first-year undergraduate student and assigned him a problem to solve over the summer: connect a television camera to a computer and get the machine to describe what it sees

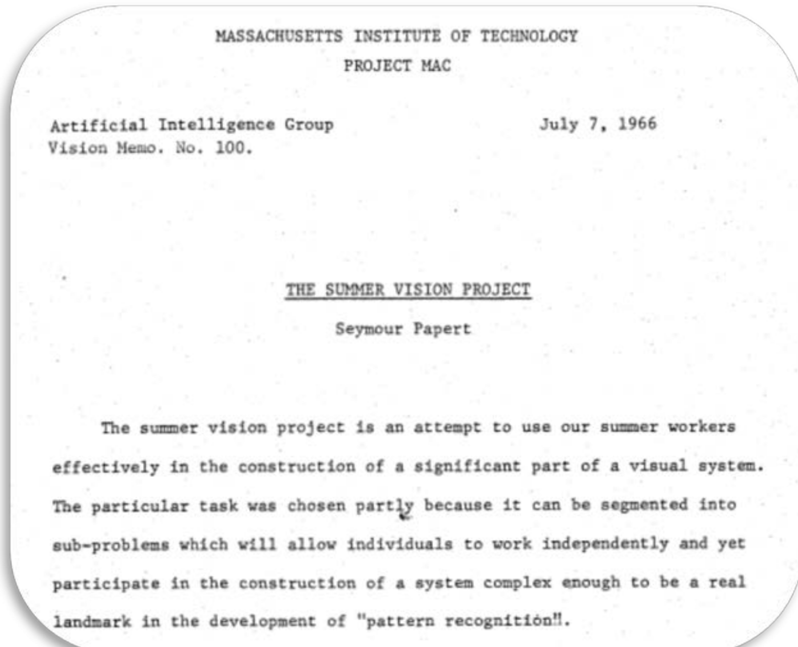


Marvin Minsky, MIT
Turing award,1969

The Summer Project

- “You’ll notice that Sussman never worked in vision again!” – Berthold Horn

Gerald Sussman, MIT



MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

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

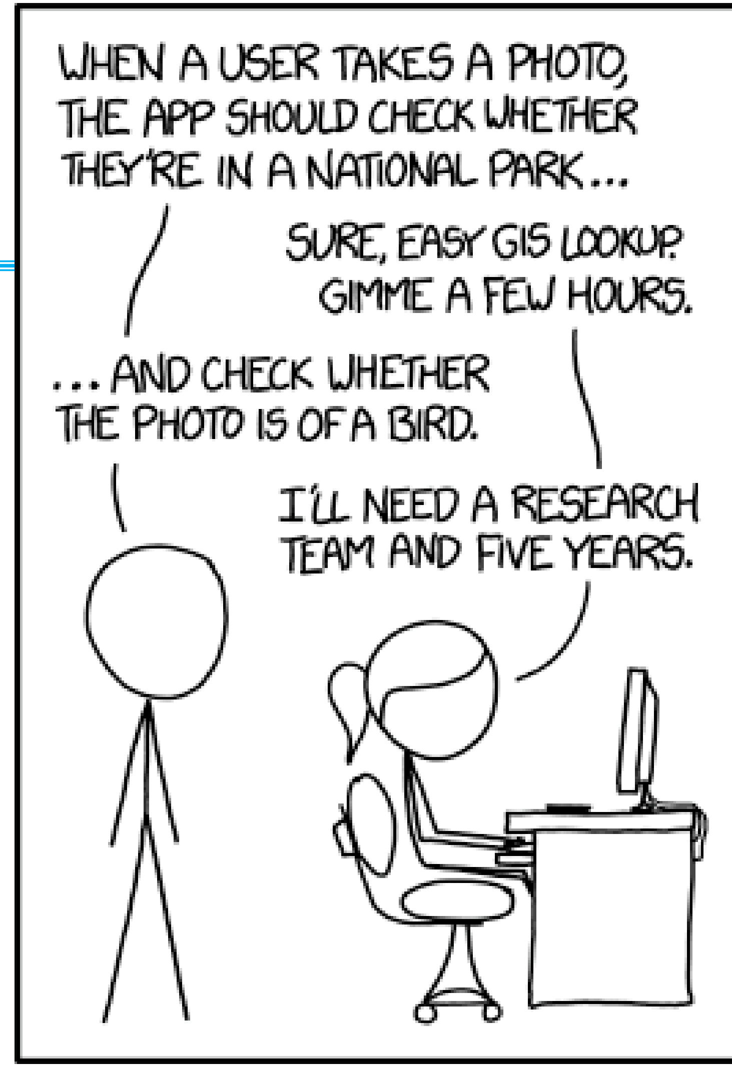
THE SUMMER VISION PROJECT

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".

The Summer Project

- 50 years and thousands of researchers later, the problem is still unsolved, but the computer vision field is making progress!



IN CS, IT CAN BE HARD TO EXPLAIN THE DIFFERENCE BETWEEN THE EASY AND THE VIRTUALLY IMPOSSIBLE.

Computer Vision

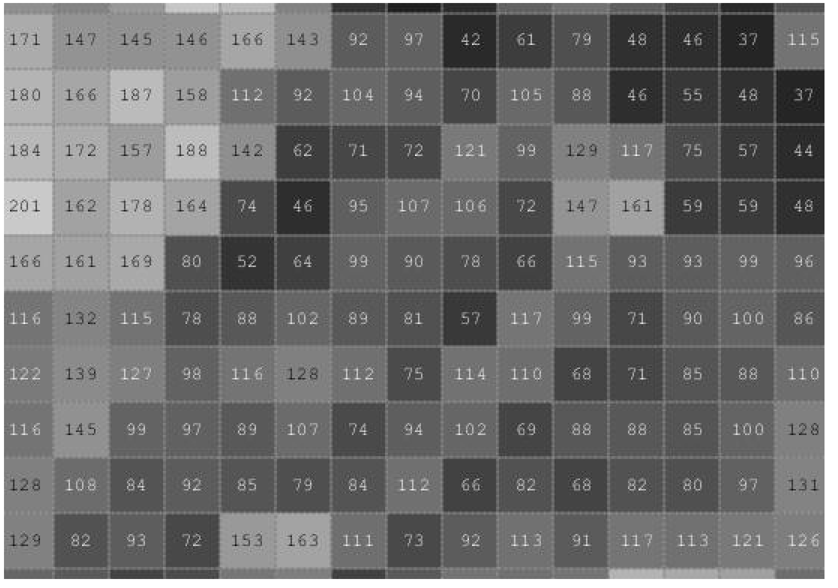
- Computers still find most of these tasks to be really hard.
- To understand why it is hard for computers, let's ask the following questions:
  - What information does the computer get?
  - What is an image?

Every picture tells a story

- Humans are able to infer an entire story from a single monochrome picture.
- We are able to identify the objects in the scene, activities, intentions, relationships between the components.

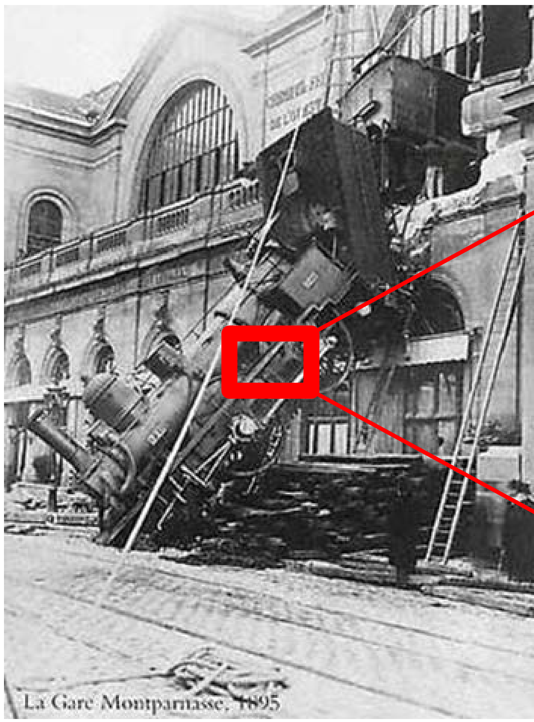
What do computers “see”?

An image is just a set of numbers.



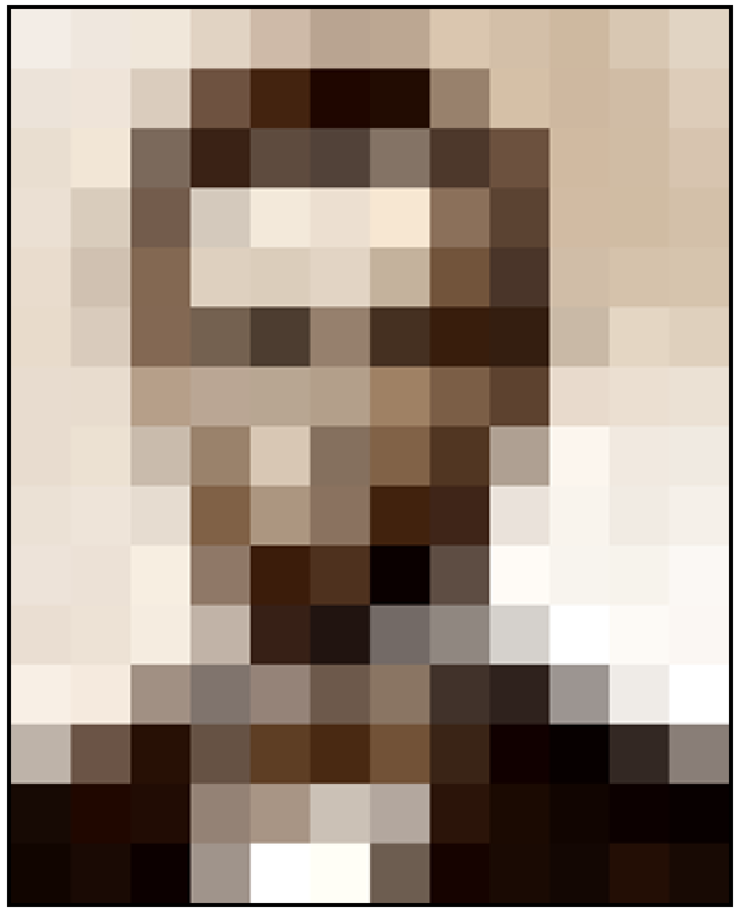
Can you tell what scene is being viewed just by looking at these numbers?

What do computers “see”?

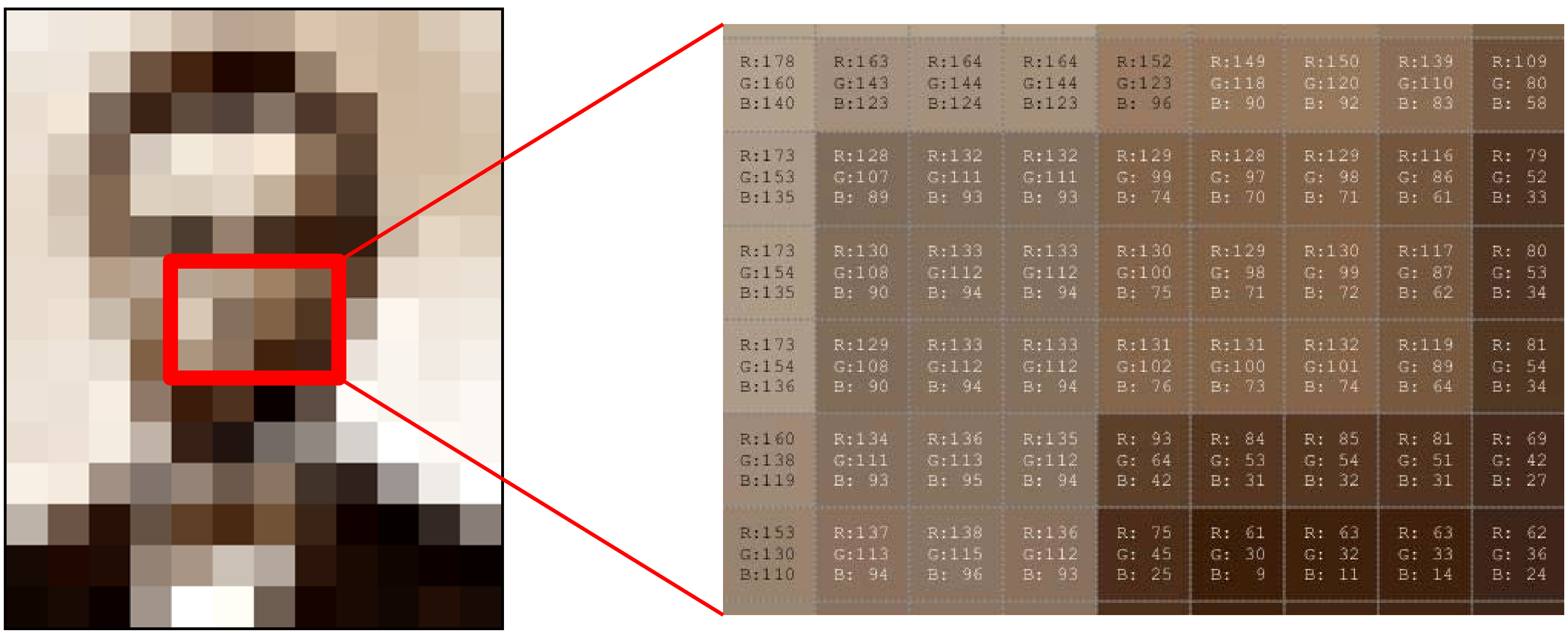


171	147	145	146	166	143	92	97	42	61	79	48	46	37	115
180	166	187	158	112	92	104	94	70	105	88	46	55	48	37
184	172	157	188	142	62	71	72	121	99	129	117	75	57	44
201	162	178	164	74	46	95	107	106	72	147	161	59	59	48
166	161	169	80	52	64	99	90	78	66	115	93	93	99	96
116	132	115	78	88	102	89	81	57	117	99	71	90	100	86
122	139	127	98	116	128	112	75	114	110	68	71	85	88	110
116	145	99	97	89	107	74	94	102	69	88	88	85	100	128
128	108	84	92	85	79	84	112	66	82	68	82	80	97	131
129	82	93	72	153	163	111	73	92	113	91	117	113	121	126

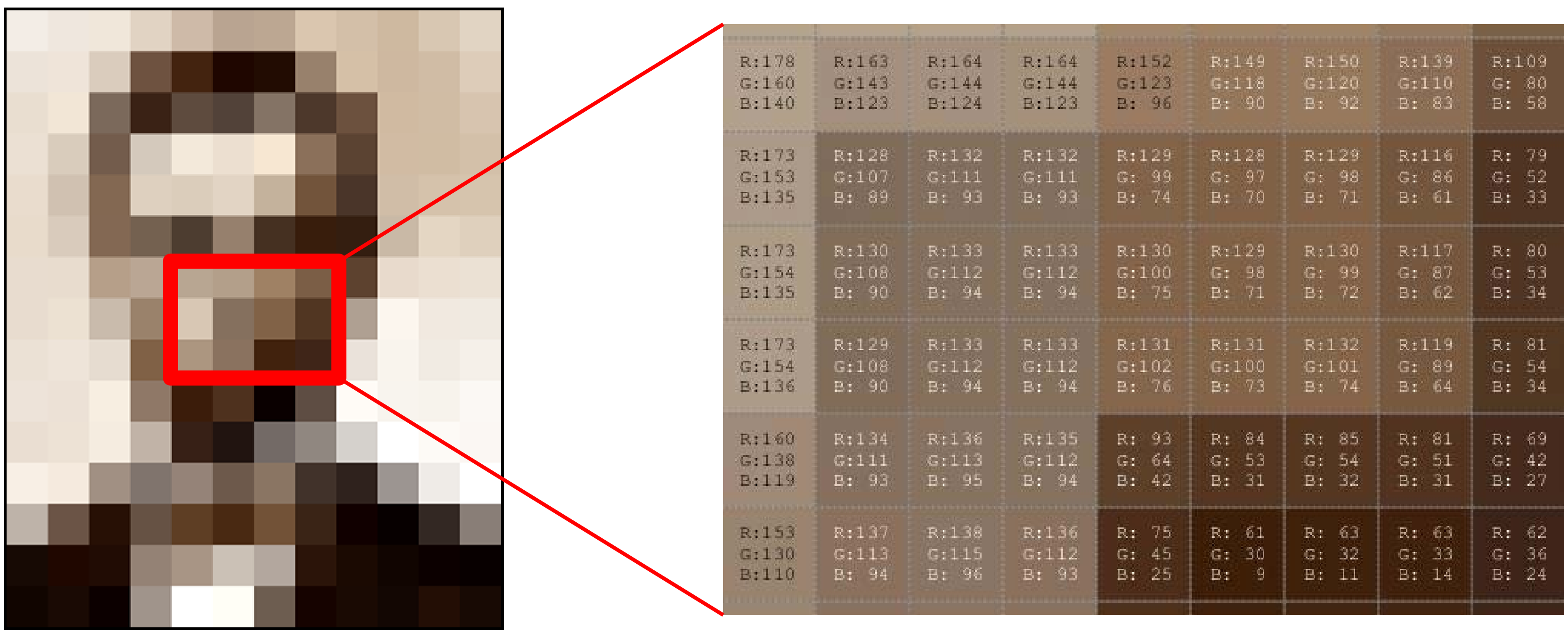
What about this one?



What do computers "see"?



What do computers "see"?



These "lifeless" numbers do not tell us what is going on in the scene.

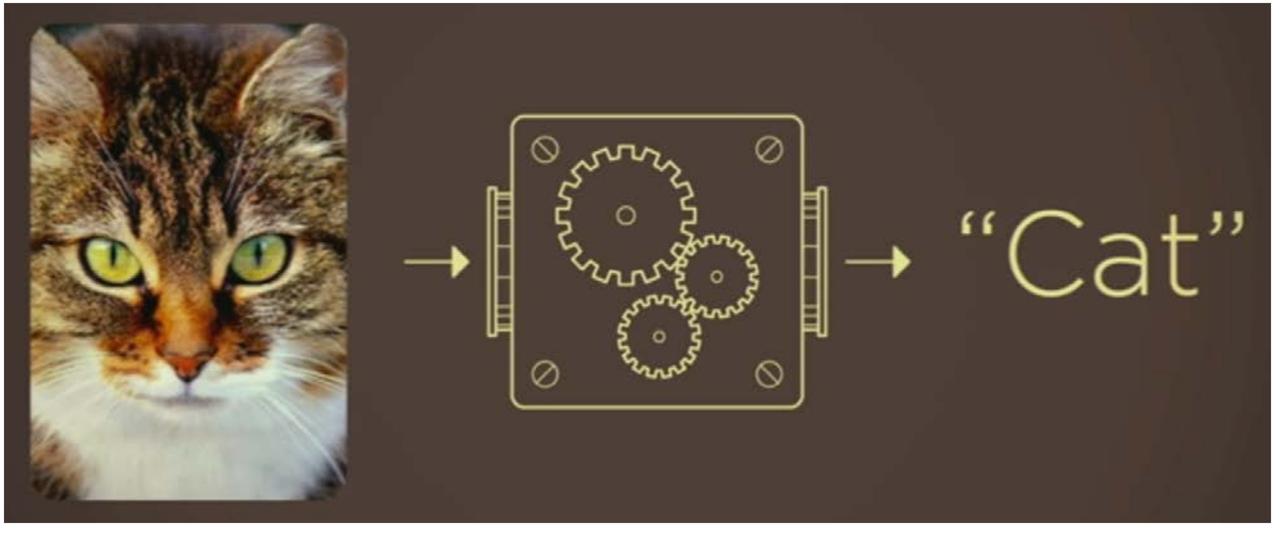
We are not there yet…

- These are a series of lifeless numbers.
- We are not able to yet interpret them as humans do.
- Taking pictures (or “looking”) is not the same as “seeing”, just as hearing is not the same as listening.
- By seeing, we really mean understanding…

Fei-Fei Li: How we're teaching computers to understand pictures, TED2015

Let’s consider the problem of teaching computers to recognize one object, let’s say cats

- We show the computer some training images of cats, and design a model that learn from these training images

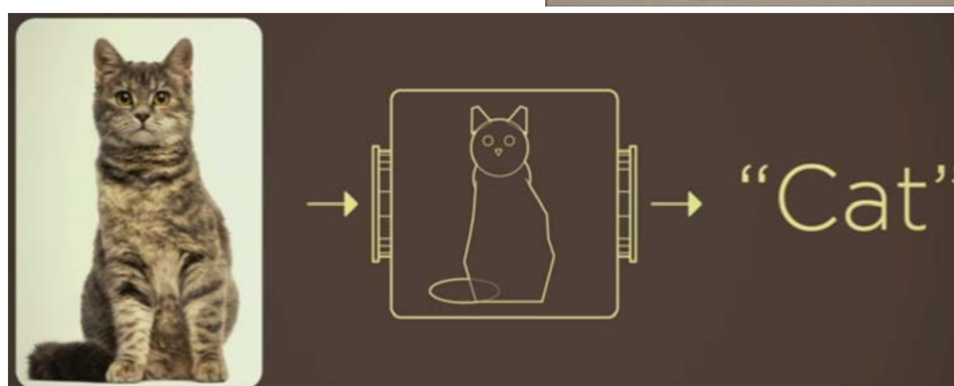


Fei-Fei Li: How we're teaching computers to understand pictures, TED2015

Let's consider the problem of teaching computers to recognize cats

- We can mathematically describe a cat as a collection of shapes and colors
  - Round face
  - Chubby body
  - Two pointy ears
  - Long tail
  - …

- This simplifies our model



Fei-Fei Li: How we're teaching computers to understand pictures, TED2015

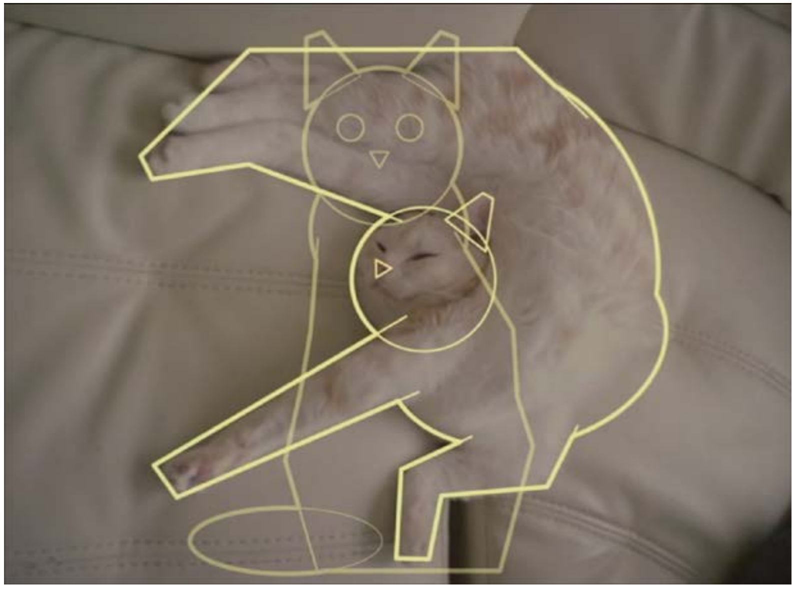
Let's consider the problem of teaching computers to recognize cats

- But what about this cat?

Fei-Fei Li: How we're teaching computers to understand pictures, TED2015

Let's consider the problem of teaching computers to recognize cats

- But what about this cat?

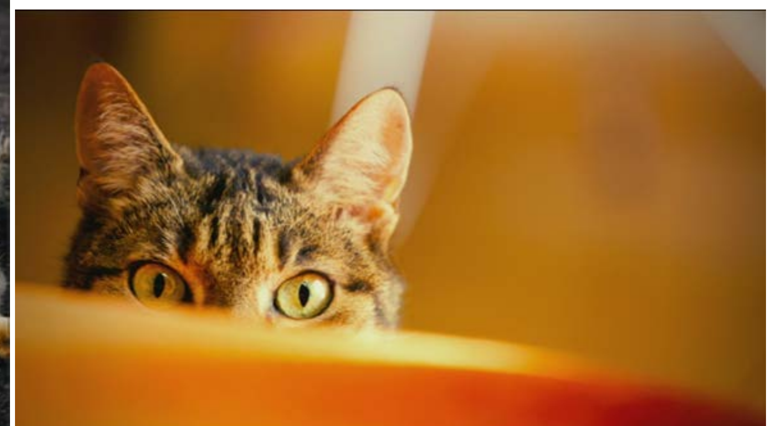


- Now we have to add another shape to our model

Fei-Fei Li: How we're teaching computers to understand pictures, TED2015

Let's consider the problem of teaching computers to recognize cats

- But what about these?

- Infinite number of variations for the cat model
- All these just for one object!

Fei-Fei Li: How we're teaching computers to understand pictures, TED2015

Course Motivation

Why is this a hard problem?

Course Motivation

- Difficult because information is lost when projecting 3D world onto 2D images.
- Result: Different scenes give rise to same image.
- More than one possible solution exists.
- Example: Object recognition problem…

Object Recognition

Object Recognition Problem

Generate algorithm to get a computer recognize an object placed in front of it *in any orientation* from a single 2D image.

What is this?

Database of objects

Which 3D object in the database gave rise to this image?

Object Recognition Problem

Why is Problem Hard?

(1) Several objects can give rise to same measurement.

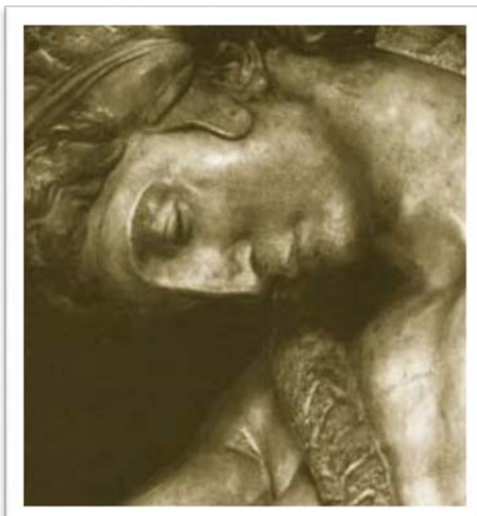
(2) Experimental uncertainties give rise to uncertain measurements.

What is this?

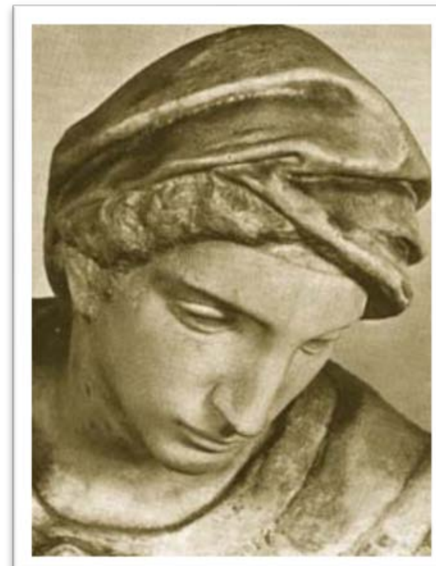
Viewpoints where impossible to pick one solution!

Why vision is so hard?

- View point variation



Michelangelo 1475-1564



Fei Fei, Fergus & Torralba

Why vision is so hard?

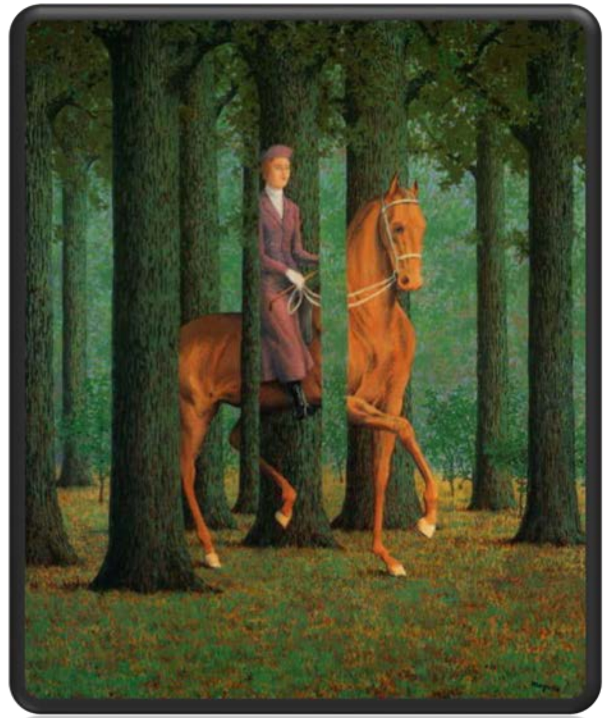
- Illumination variations

Yale Database of faces

Why vision is so hard?

- Occlusion

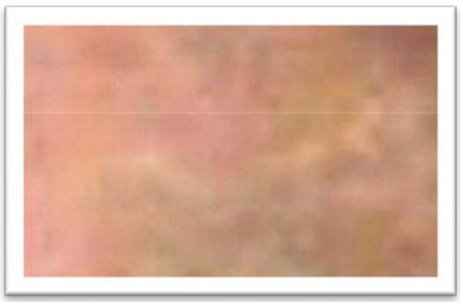
Magritte, 1957

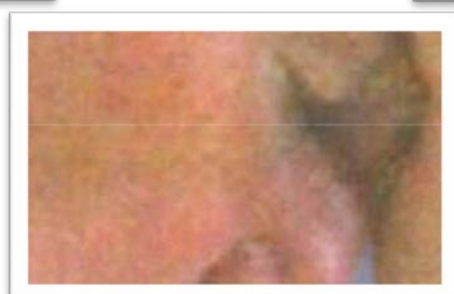


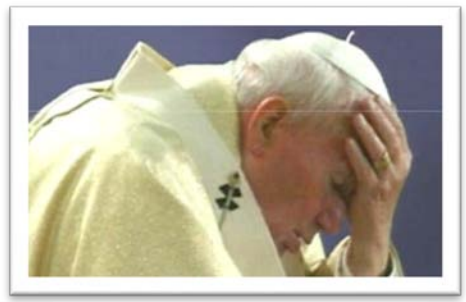
Fei Fei, Fergus & Torralba

Why vision is so hard?

- Scale

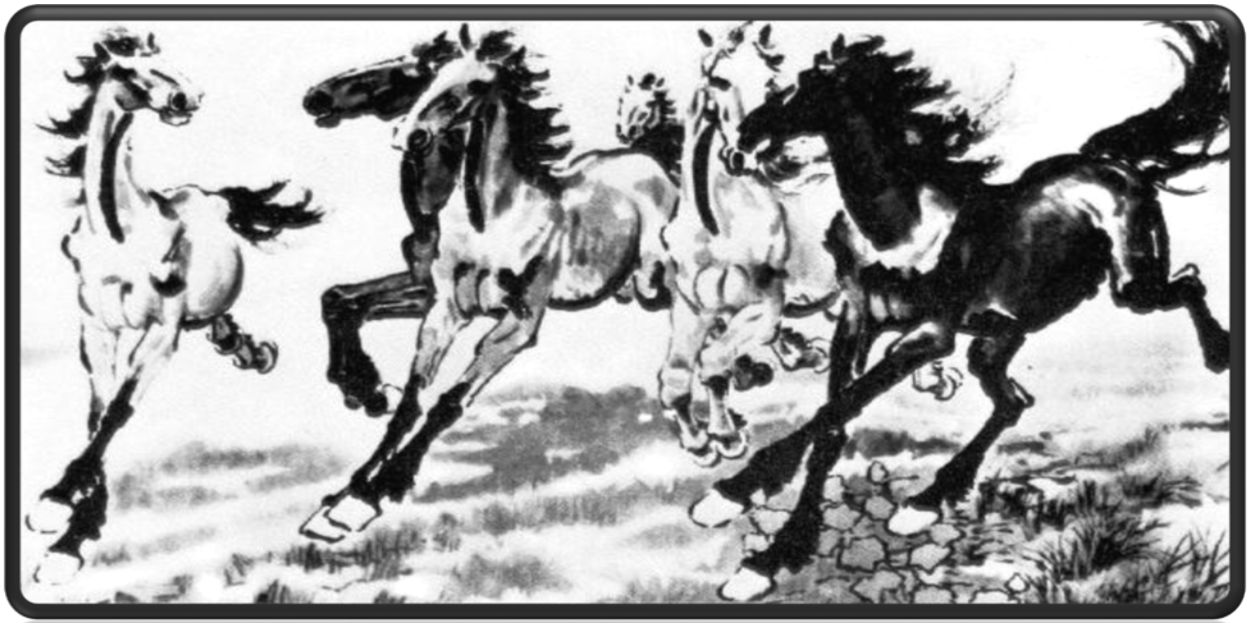






Why vision is so hard?

- Deformation



Xu, Beihong 1943

Why vision is so hard?

- Background Clutter

Fei Fei, Fergus & Torralba

Why vision is so hard?

- Intra-class variation






Fei-Fei, Fergus & Torralba

Why vision is so hard?

- Ambiguity

Fei Fei, Fergus & Torralba

Course Motivation

- Computer vision literature focuses on techniques that constrain the solution set in order to address these difficulties.
- In this course, we focus on probabilistic techniques for addressing various problems in computer vision.

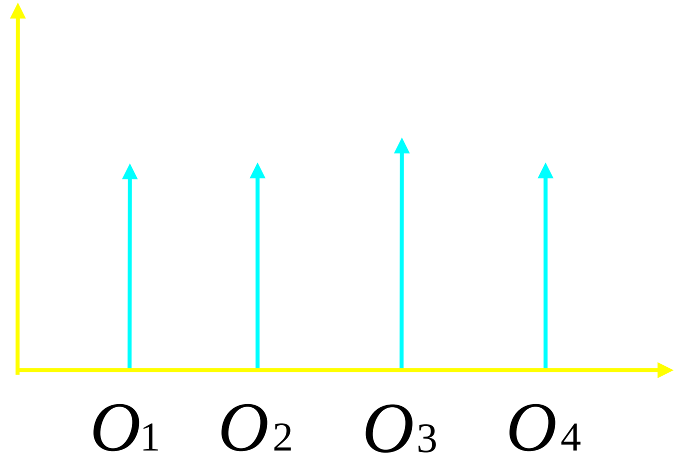
Course Philosophy

- We explore the conjecture that the world is uncertain, and should therefore should be described through the language of probabilities.
- We explore representing sources of information in the form of probability density function.
- We will examine solutions provided by probability and information theory.

Course Philosophy

- Example: object recognition.

$$p(O_i \mid \mathbf{d})$$



- Rather than force a single solution (which would be misleading in this case), represent result in the form of a posterior probability density function, $p(O_i|d)$.
- Permits explicit representation of all sources of uncertainty.
- Permits higher level processes to assess the result.

Course Objectives

- We will describe several problems in computer vision.
- We will briefly explore some standard techniques for their solution.
- We will mainly focus on probabilistic inference methods for their solution.
- **Goal:** To provide the student with the necessary tools to be able to apply these techniques to other problems in computer vision and other research areas.

Statistical Computer Vision

- All course work will be found on mycourses.

- The course assignments (e.g. literature review, final project) will be assigned on mycourses and are to be submitted electronically through mycourses.

- There will be one in-class exam at the end of term.

- Students will each give a seminar presentation of their project at the end of the term.

Statistical Computer Vision

- The final course grade will be determined as follows:
    - Assignments: 20%
    - Midterm literature review: 10%
    - Final project:
        - document: 30% + presentation: 20% = 50%
    - Final test: 20%

Statistical Computer Vision

Topics to be covered (subject to change):

- Regularization approach to solving ill-posed problems
- Applications of regularization to computer vision problems
- Bayesian formulation of regularization
- Bayesian inference
- Markov Random Field (MRF) models
- Information Theory / Pattern Recognition
- Linear models - PCA, ICA
- Sequential Bayesian methods
- Deep learning
- Applications such as : Image registration, medical image segmentation, face classification

Course Motivation

Are there jobs in this field?

What's the current state?

Update on computer vision today

- Computer vision is currently in a phase of high growth worldwide.
- Companies are leveraging what we can do already and pushing research into new directions. There is a lot of money to be made.
- What companies are investing in this area?

Update on computer vision today

- Companies:
  - Hardware companies (e.g. Intel, Samsung, Qualcomm, NVDIA),
  - Software companies (e.g. Google, Facebook, Dropbox, Snapchat),
  - Car and entertainment companies (e.g. Uber, Disney, Microsoft: Xbox)

are all currently heavily investing in these domains, and aggressively recruiting in the area.

NVIDIA's focus on pushing its graphics cards for deep learning and artificial intelligence acceleration paid off in a big way, and the stock exploded higher as a result.

Read more: http://www.fool.com/investing/2017/01/03/why-nvidia-stock-tripled-in-2016.aspx#ixzz4UjEwob5u

Update on computer vision today

- Many companies with products in the telecommunication area such as Apple, Blackberry, Google, Bell, Apple, and Nokia all have significant activities in developing vision and image manipulation apps for use in cellphones.
- Furthermore, a large number of startups in this area have recently enjoyed tremendous success due to the maturity and availability of the algorithms in the field as well as the ubiquitous nature of cameras today.

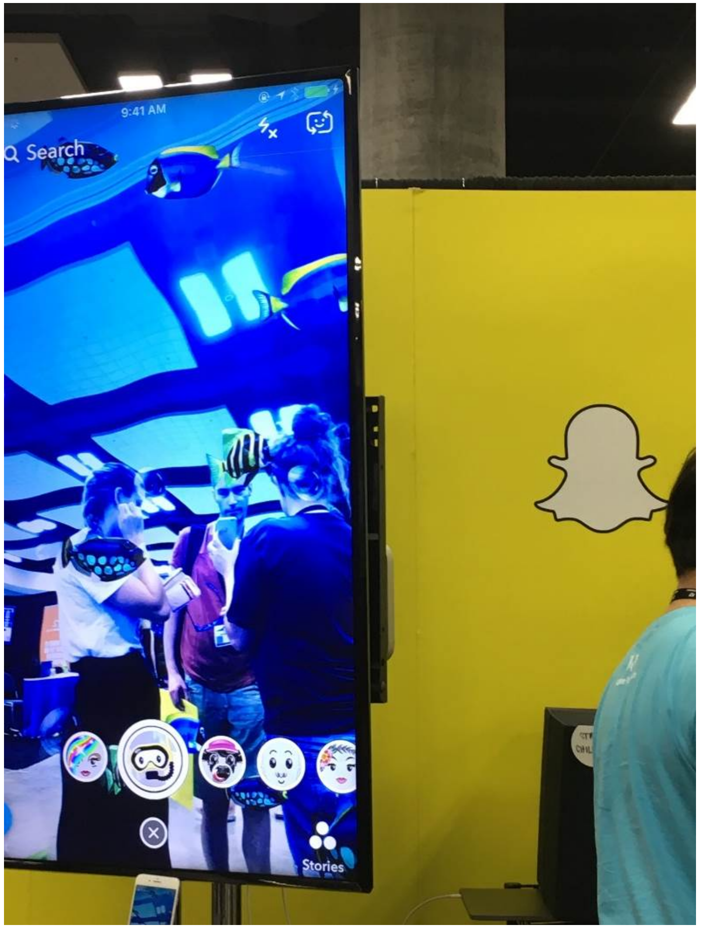
International Conference on Computer Vision and Pattern Recognition (CVPR) 2017

International Conference on Computer Vision and Pattern Recognition (CVPR) 2017

International Conference on Computer Vision and Pattern Recognition (CVPR) 2017

International Conference on Computer Vision and Pattern Recognition (CVPR) 2017

International Conference on Computer Vision and Pattern Recognition (CVPR) 2017

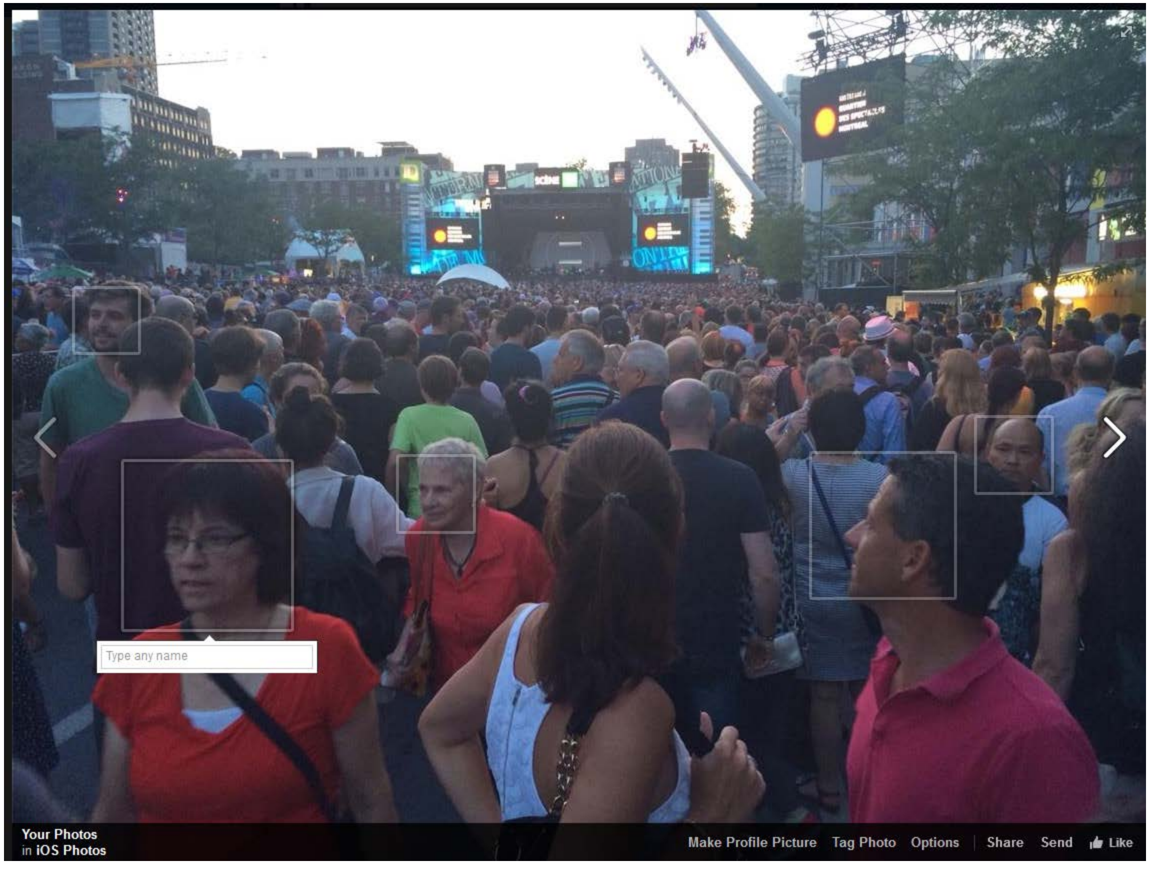


International Conference on Computer Vision and Pattern Recognition (CVPR) 2017

What works today?

Face Detection

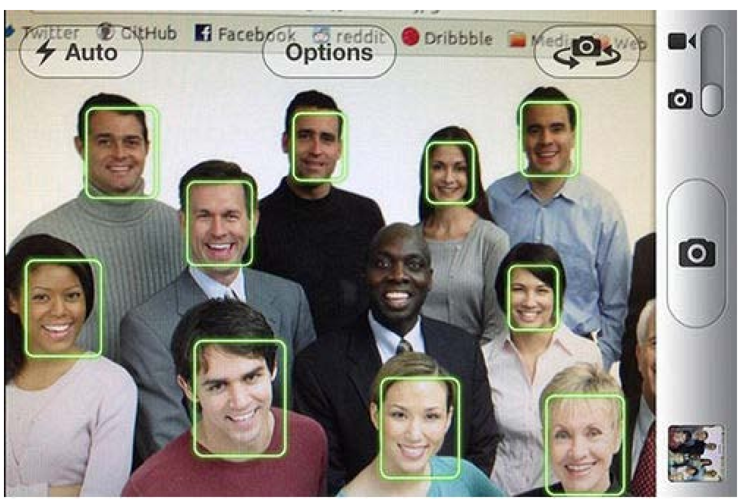
Face Detection



Facebook detects faces in uploaded images to aid in tagging

Face Detection

- Most newer digital cameras can detect faces. This is so that the picture taker can make sure the faces are in focus and properly exposed.
- It doesn't always work…

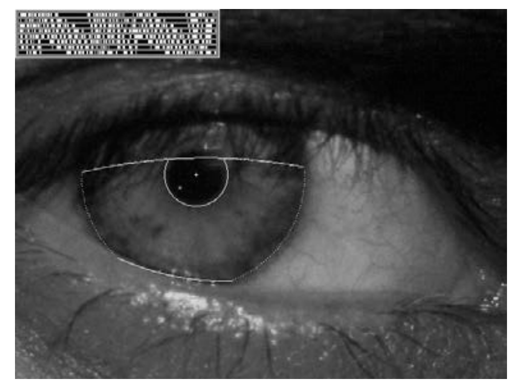


Face Recognition

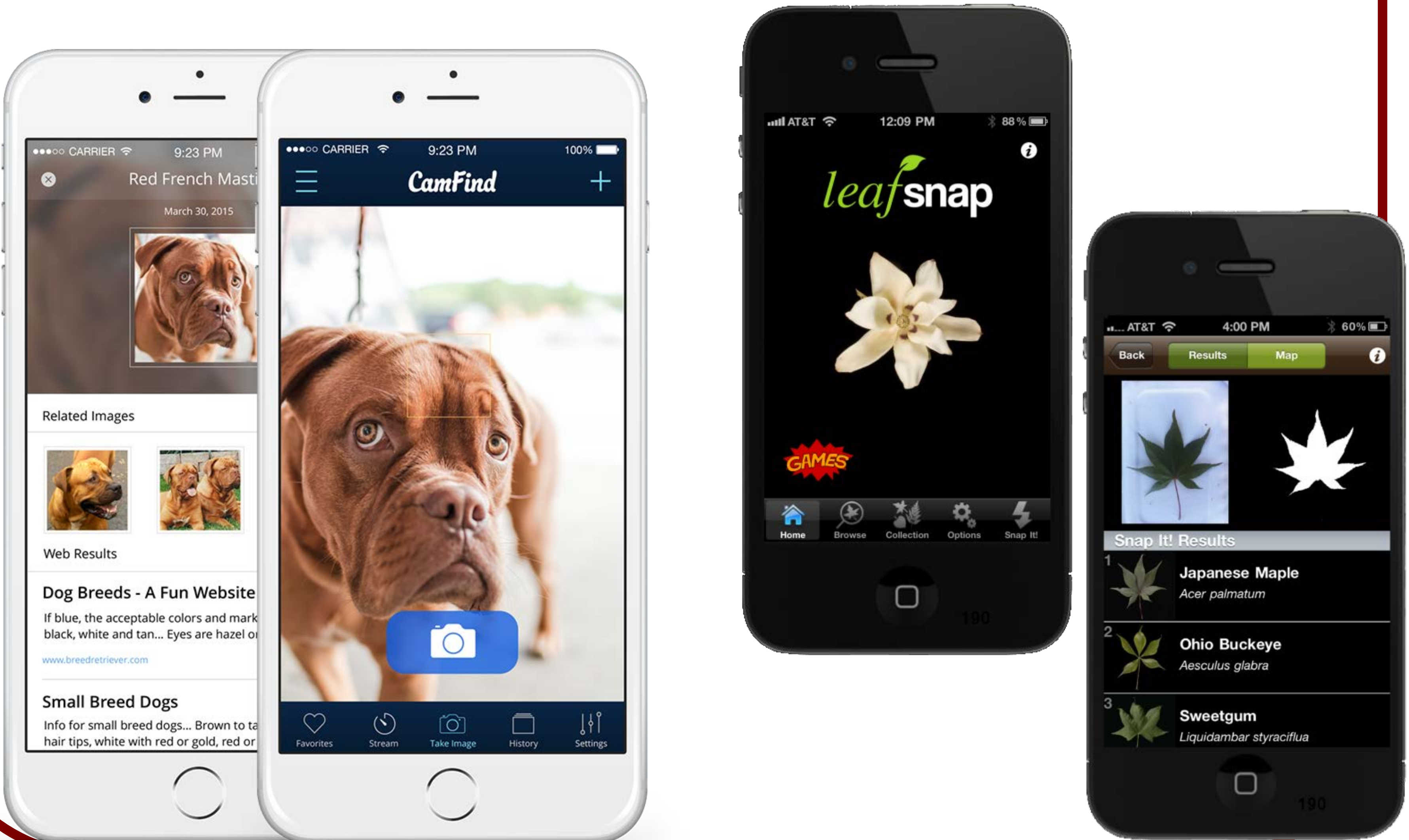
Face Recognition: the many faces of Madonna

Face Recognition: vision based biometrics

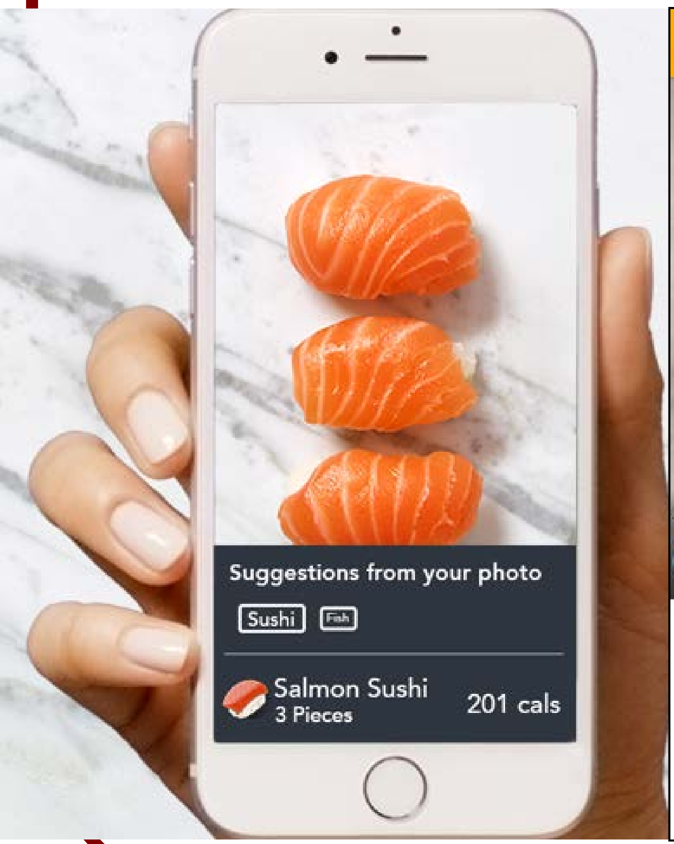
"*How the Afghan Girl was Identified by Her Iris Patterns, 18 years later*"
Read the story

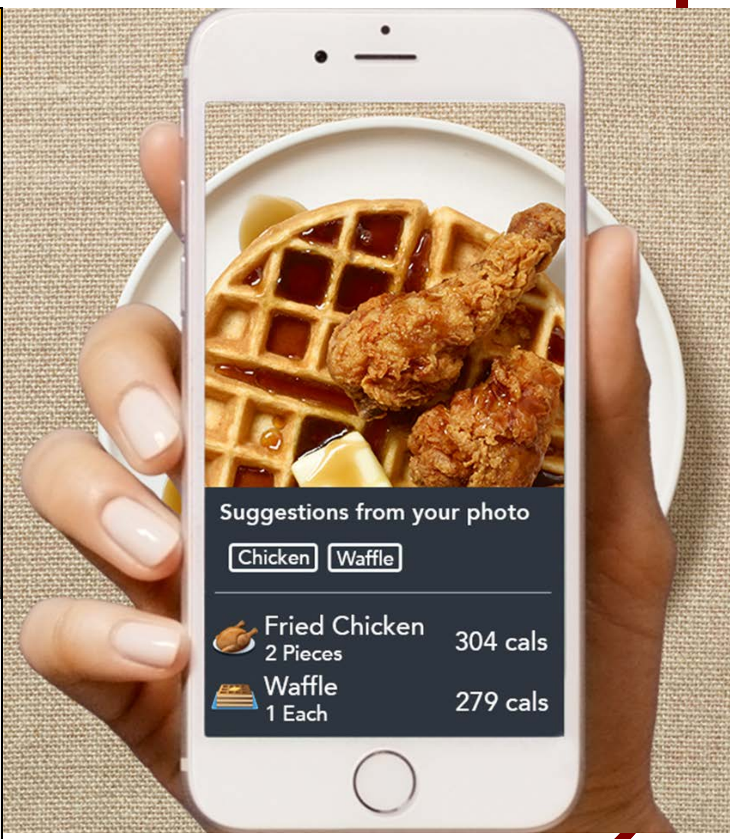


Object Recognition - Apps

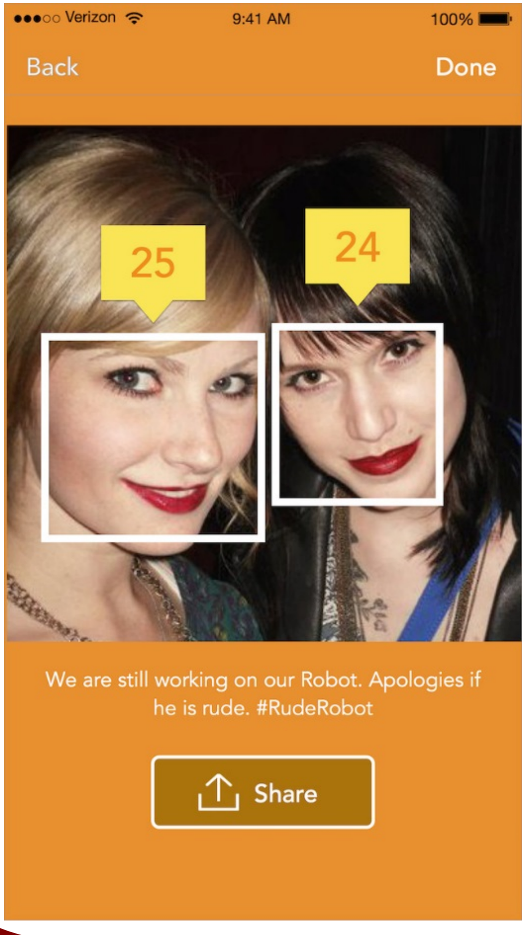


Personal Dietary Assistants





Age Estimation Apps



Sports

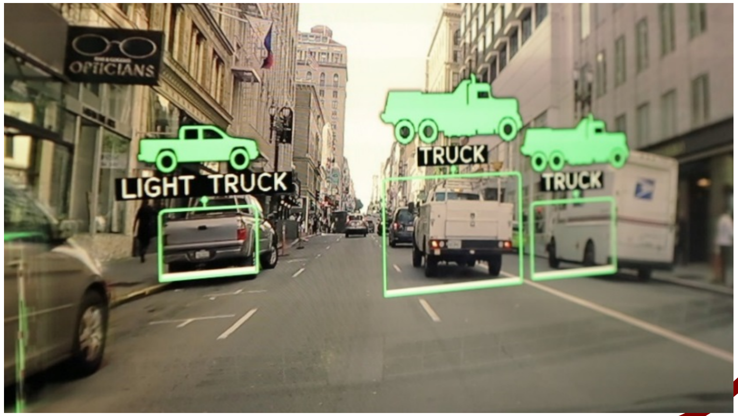
- Sportvision first down line
- Nice explanation on www.howstuffworks.com

Industrial Manufacturing and Quality Control

Smart Cars

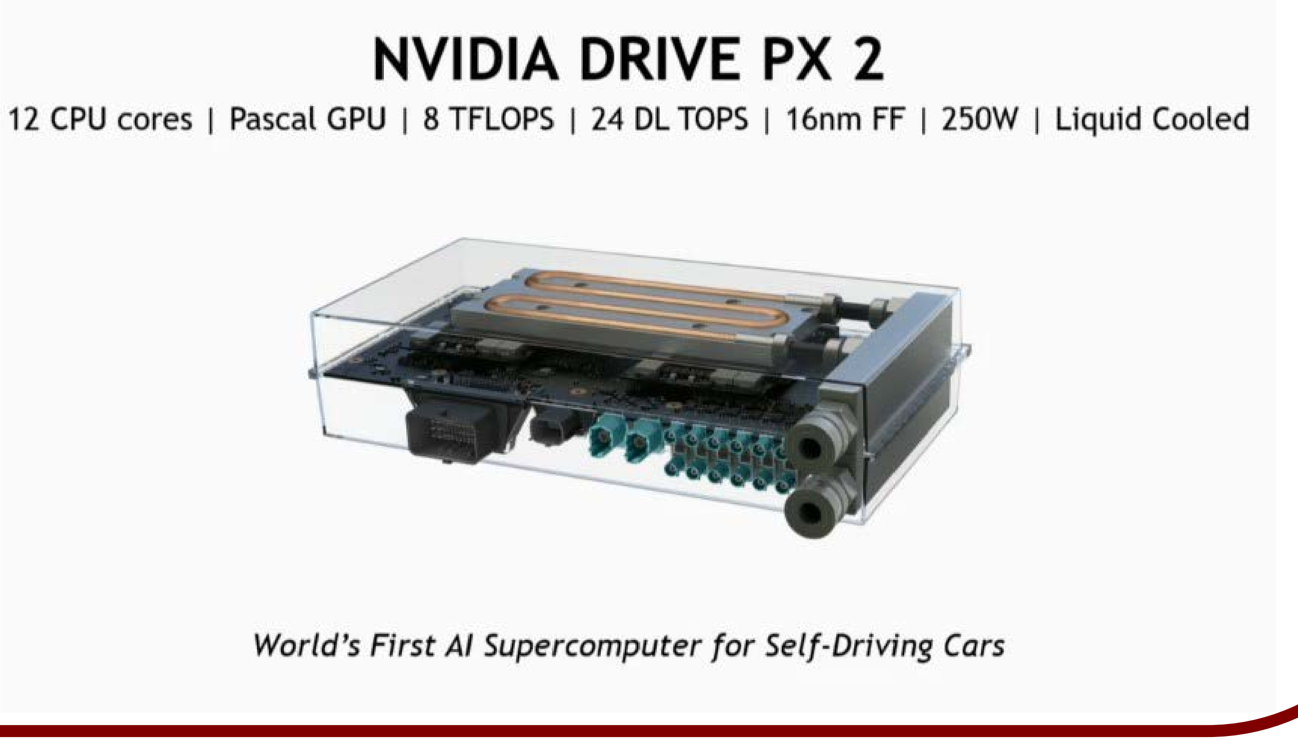
- Mobileye
  - Vision systems currently in high-end BMW, GM, Volvo models
  - Pedestrian collision warning
  - Forward collision warning
  - Lane departure warning
  - Headway monitoring and warning
  - https://www.youtube.com/watch?v=kp3ik5f3-2c

- Slide content courtesy of Amnon Shashua



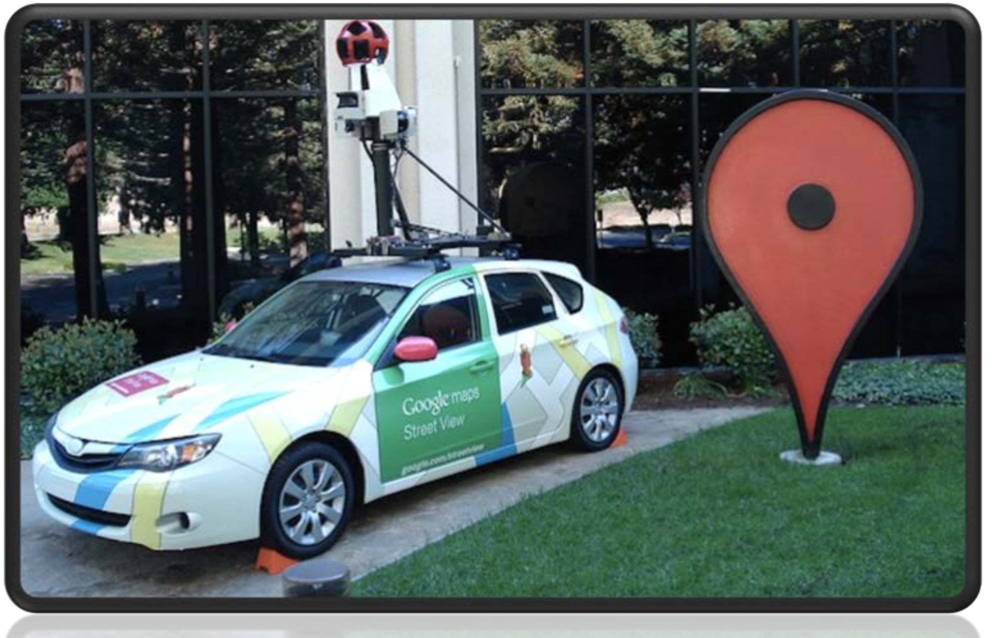
Smart Cars

- NVDIA
  - Have created special purpose hardware for computer vision in cars
  - https://www.youtube.com/watch?v=zsVsUvx8ieo



Smart Cars

- Google Cars



- Slide content courtesy of Amnon Shashua

Vision for Unmanned Aerial Vehicles

- Slide content courtesy of Amazon

Prototype of Amazon Prime Air drone delivery vehicle

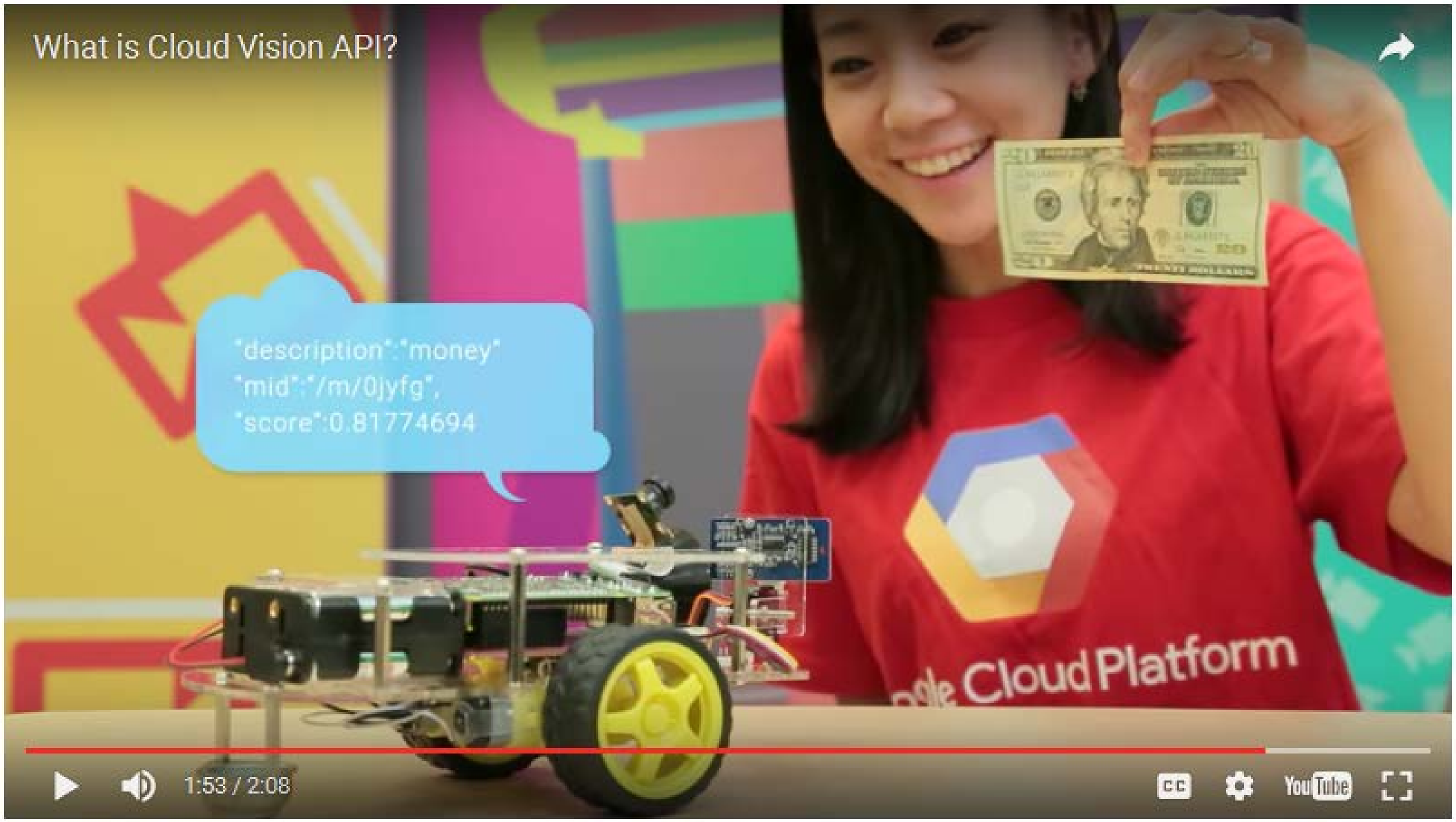
Vision for Unmanned Aerial Vehicles

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

- Vision systems (JPL) used for several tasks
  - Panorama stitching
  - 3D terrain modeling
  - Obstacle detection, position tracking

Robotics

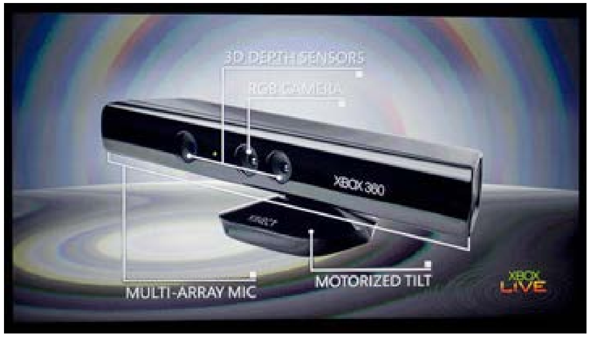
- https://www.youtube.com/watch?v=eve8DkkVdhI



Simple mobile robot using computer vision with the Google CloudVision API

Motion Capture

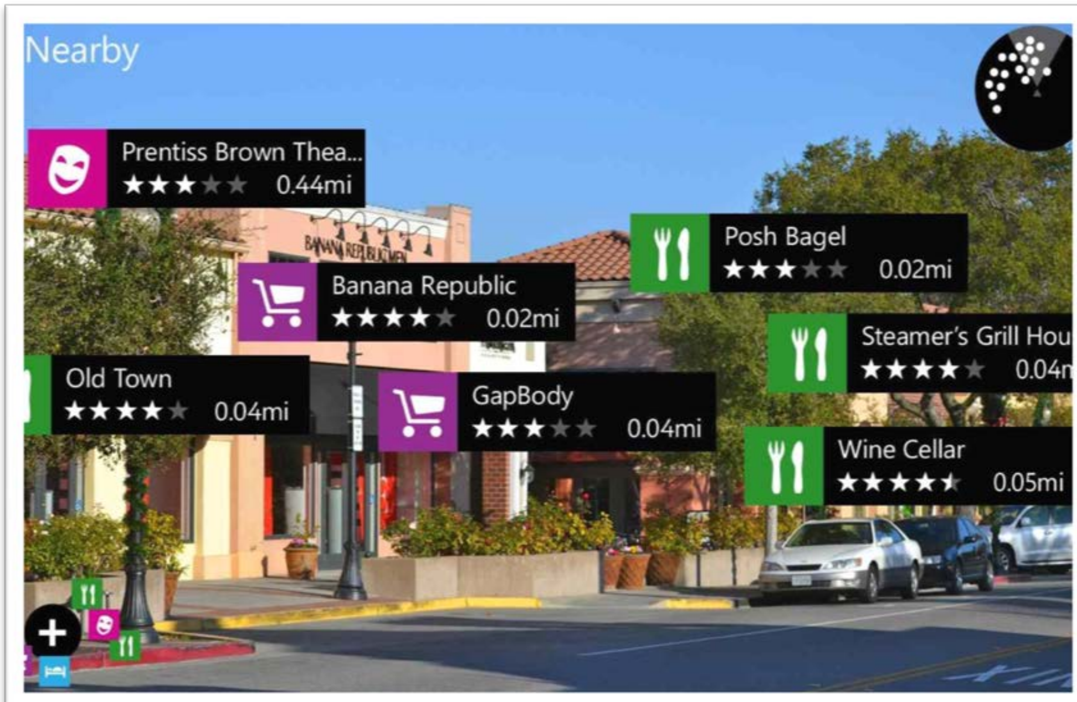
- Vision based interaction
  - Microsoft Kinect



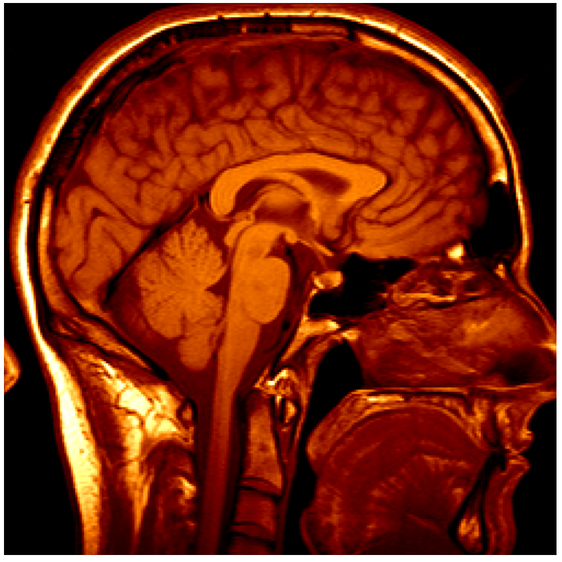
Special Effects: Shape Capture

- Andy Serkis, Gollum, Lord of the Rings

Augmented Reality



Medical Imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Current state of the art

- You just saw examples of current systems.
  - Many of these are less than 5 years old
- This is a very active research area, and rapidly changing
  - Many new apps in the next 5 years
- To learn more about vision applications and companies
  - David Lowe maintains an excellent overview of vision companies
    - http://www.cs.ubc.ca/spider/lowe/vision.html

Have fun!

Subway
Detected: Sedan
just now
Object detection Results
Sedan class