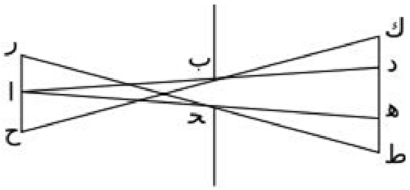
Lecture 1:

Brief history of computer vision

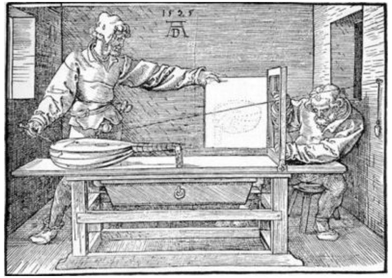
CSE 455 - Computer Vision

Science stands on the shoulder of giants

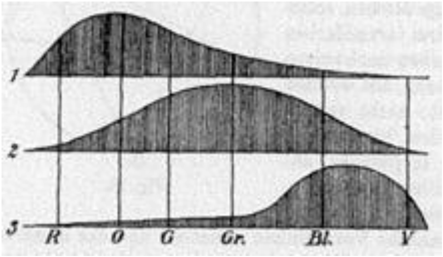
Computer vision draws origins from math & physics



Pinhole projection, optics



Projective geometry



Models of color vision (trichromacy)

Early theories of visual perception: Helmholtz, Two big technologies changed how computer vision was studied and how we understand them today.

Q. Can anyone here guess what those two events were?

First technology: Camera

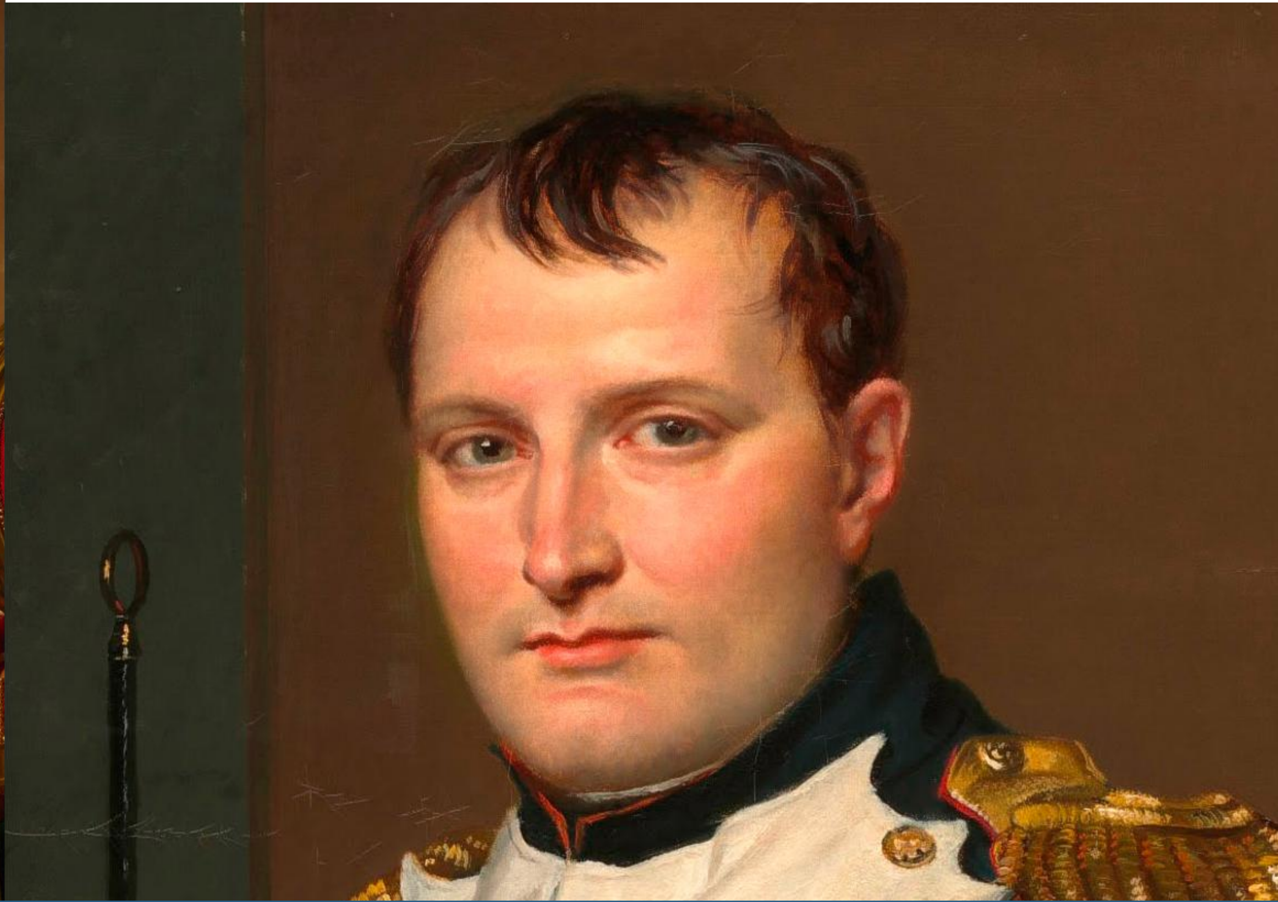
Aside from physics and math, computer vision also has connections to art

Pictures before 1838

Portraiture - artists would spend hours/days drawing their subjects who stood still in front of them

1812: Jacques-Louis-David

The Emperor Napoleon at his Study at the Tuileries

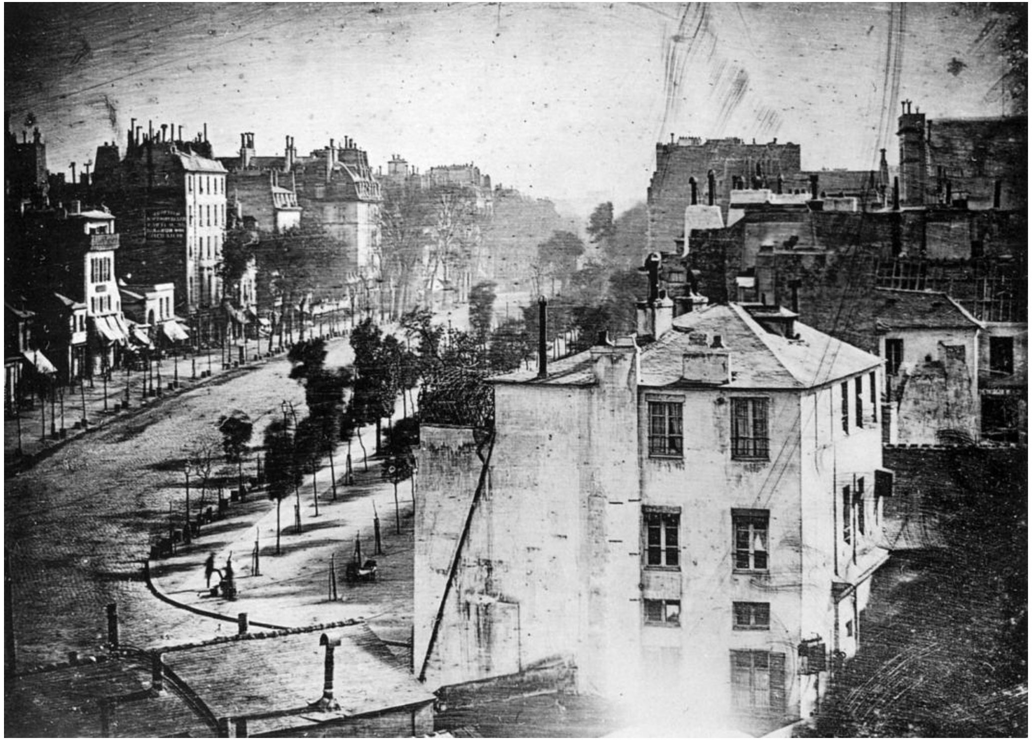


1775-1817: Jane Austin, famous author.

From biography.com

1837: Niépce, First photo of one’s meal

1838: Boulevard du Temple, Daguerre



1838: First selfie, Robert Cornelius

Technology often begets fear

"From today, painting is dead"
— painter Paul Delaroche
at a demonstration of the Daguerreotype, 1839

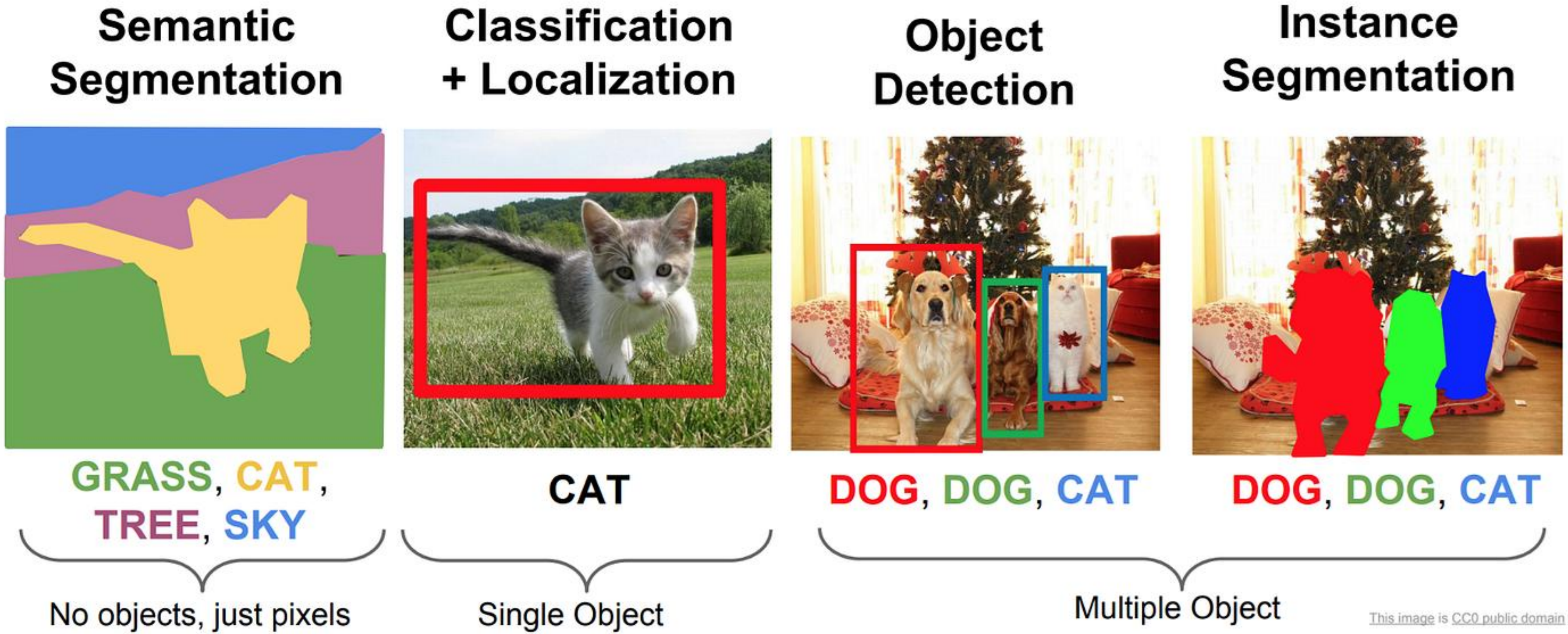
Second technology

- 1957: Digital scanner invented at NIST

With smaller cameras and larger storage,

We began curating large scale databases of images online

With those images, we now train models to understand what is in an image



We can also train models to generate new images

Neural Style Transfer [Gatys et al. 2015]

AI Art

In February 2016, Gray Area and Google presented the world's first exhibition of generative AI art, titled "DeepDream: The art of neural networks".

New technology begets fear

Can Computers Create Art?

Aaron Hertzmann
Adobe Research*
Working draft†

January 16, 2018

Abstract

This paper discusses whether computers, using Artifical Intelligence (AI), could create art. The first part concerns AI-based tools for assisting with art making. The history of technologies that automated aspects of art is covered, including photography and animation. In each case, we see initial fears and denial of the technology, followed by acceptance, and a blossoming of new creative and professional opportunities for artists. The hype and reality of Artificial Intelligence (AI) tools for art making is discussed, together with predictions about how AI tools will be used. The second part concerns AI systems that could conceive of artwork, and be credited with authorship of an artwork.

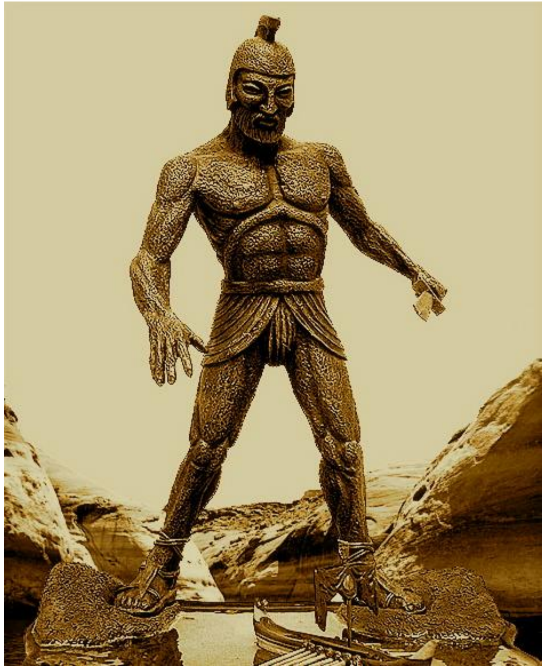
At the end of the day, vision is for doing

One single model controls multiple robot embodiments [Link to paper]

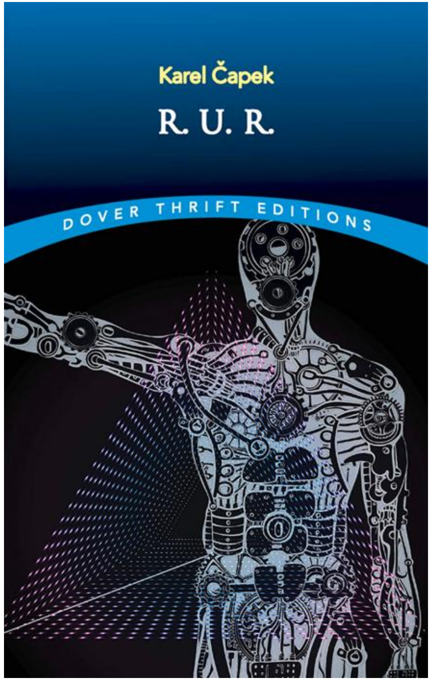
Robots and agents powered by vision

have often been depicted by popular media

Depictions of AI: Myths and Stories



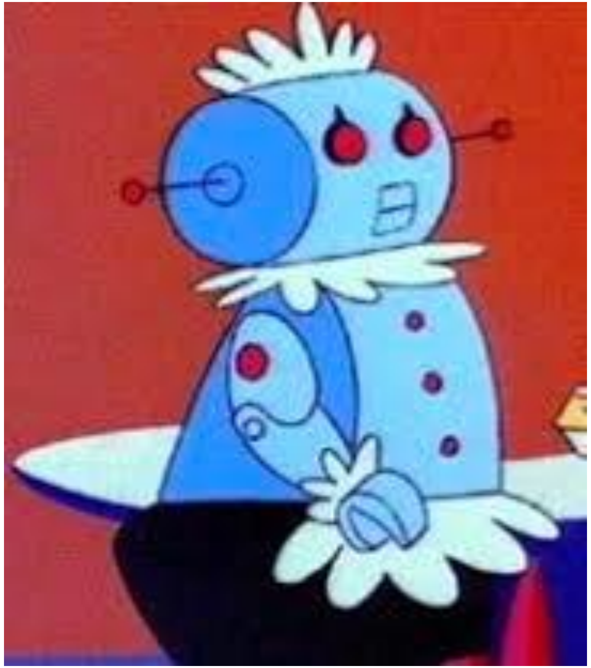
Legend of Talos
Adrienne Mayor, *Gods and Robots*



R. U. R. (1920)

Data in *Star Trek* (1987)

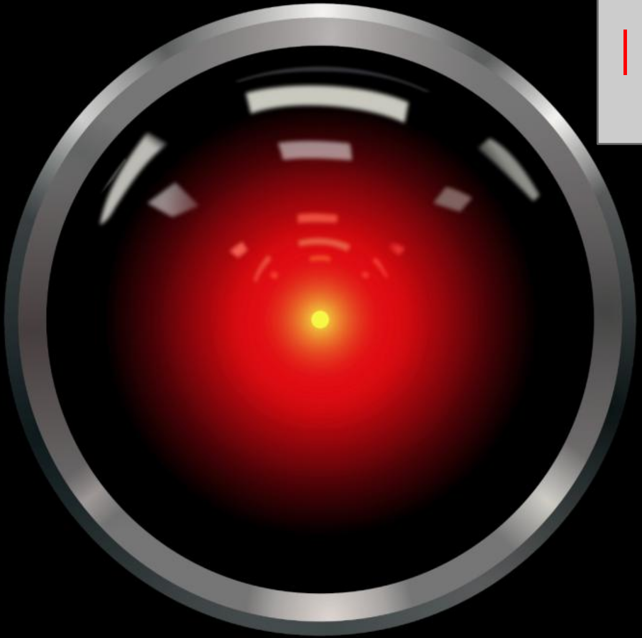
Jetson's robot



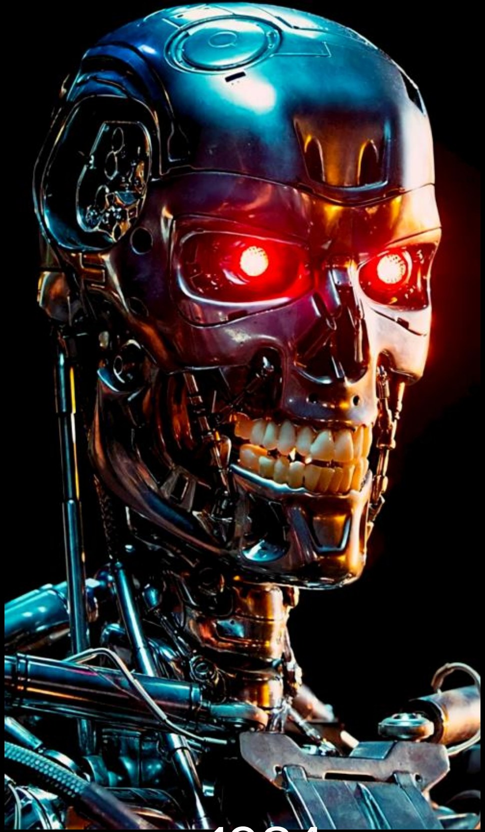
1956

1962

1965



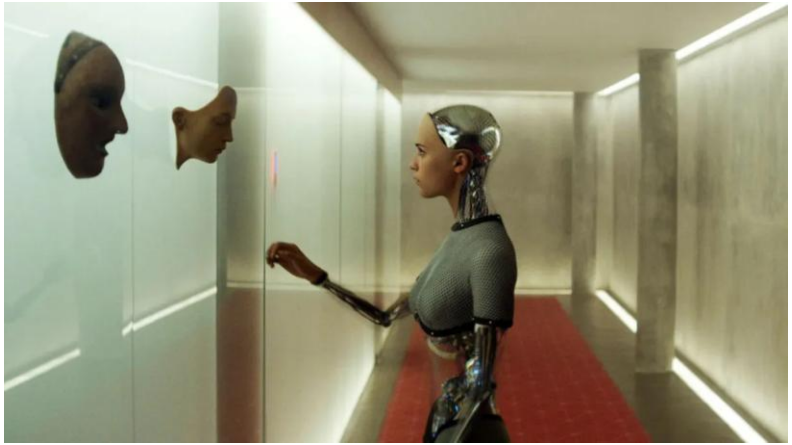
I’m sorry, Dave.
I’m afraid
I can’t do that.



1968 (2001: A Space Odyssey)

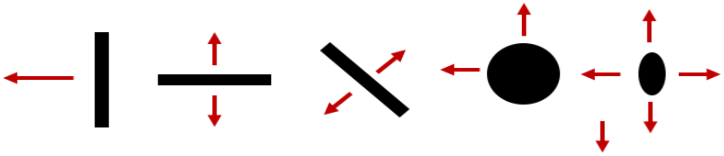
1984



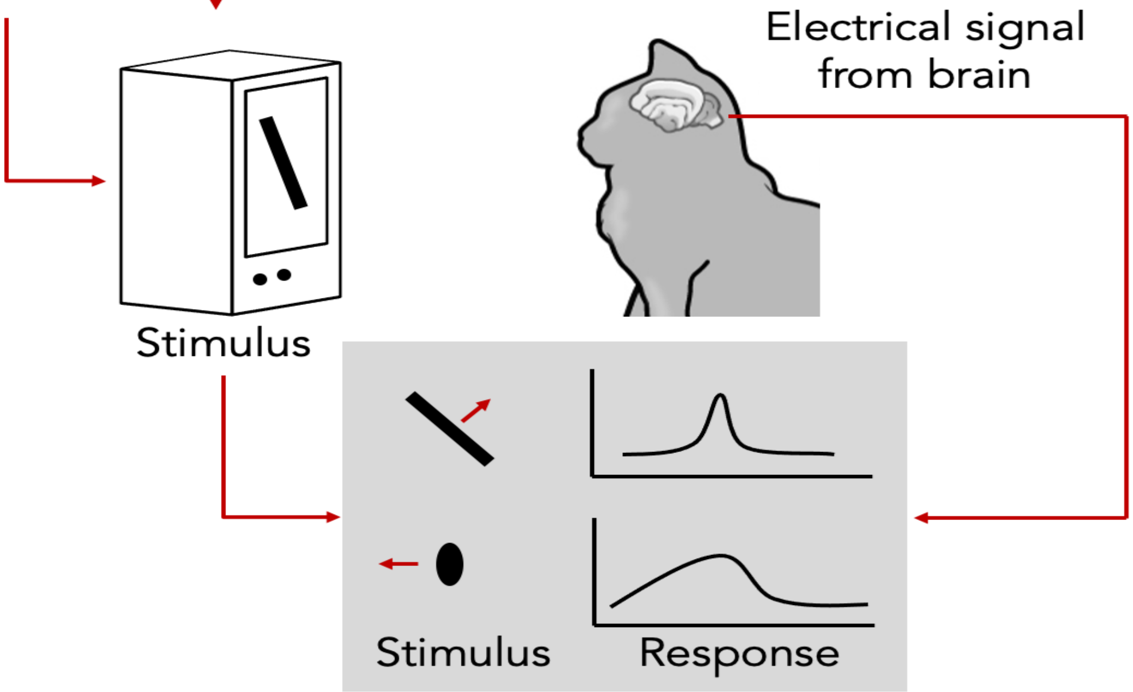
Aside from physics, math, art, popular media,

Computer Vision also draws on fundamental findings in neuroscience

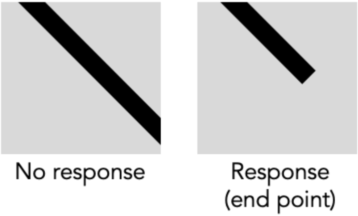


Hubel & Wiesel, 1959

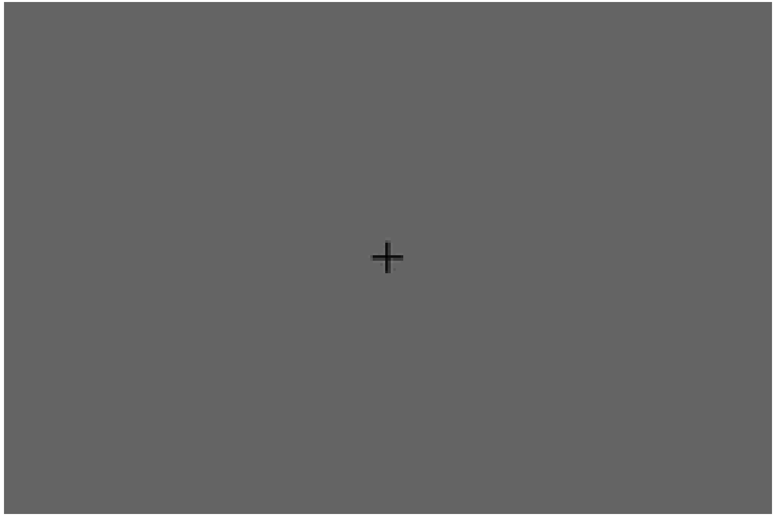
How does animal vision work?



Won Nobel Prize in 1981
Visual processing is hierarchical, involving recognizing simpler structures, edges, etc.

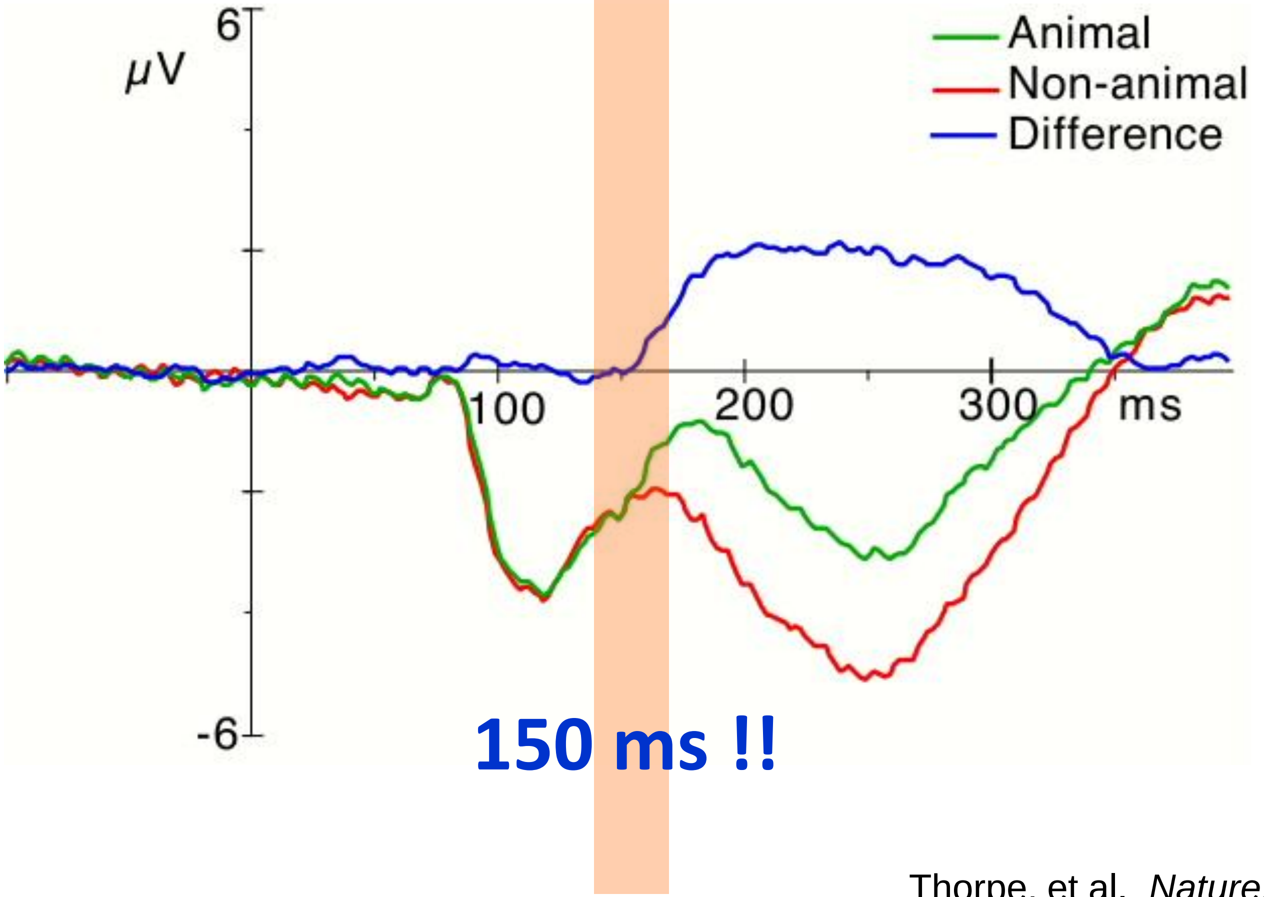


Human vision is superbly efficient



Potter, Biederman, etc. 1970s

Thorpe, et al. *Nature,*



150 ms !!

Thorpe, et al. *Nature,* 1996

Aside from physics, math, art, popular media, neuroscience

Computer Vision is also influenced by cognitive science explorations

Change Blindness

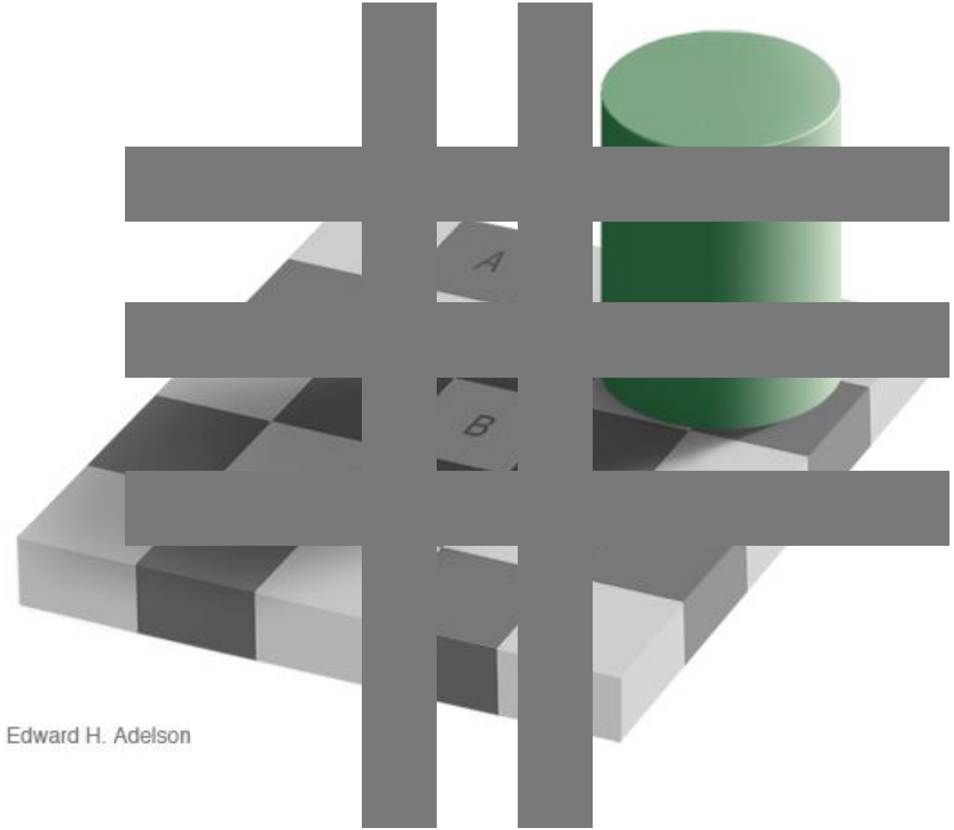
Rensink, O’regan, Simon, etc.

Change Blindness

Rensink, O'regan, Simon, etc.

camouflage

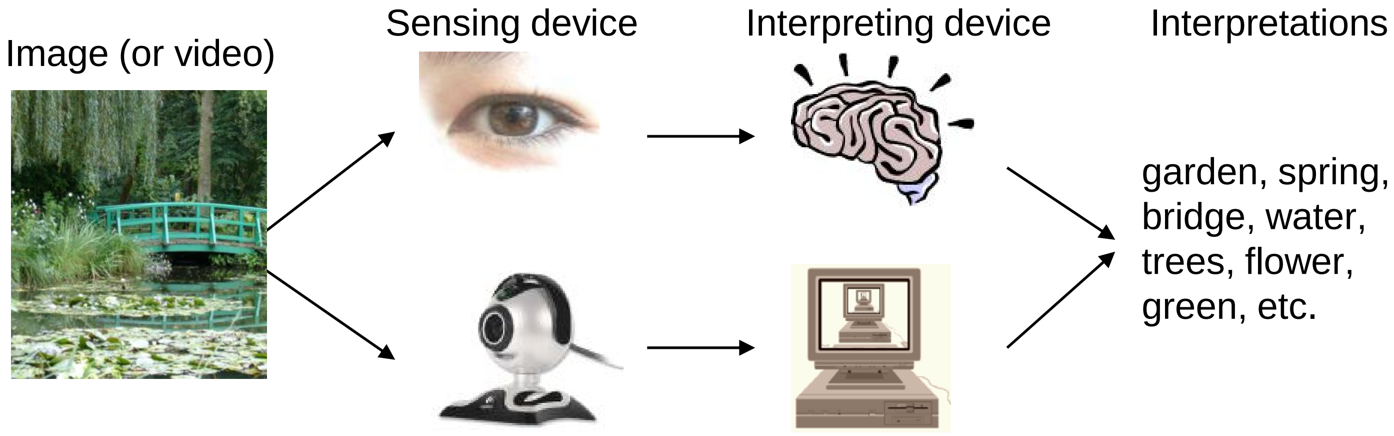
Who are these two people?


A
B
Edward H. Adelson

Motion without movement

Common theme in computer vision: which parts of human vision are necessary for intelligent systems?

So, what is computer vision?



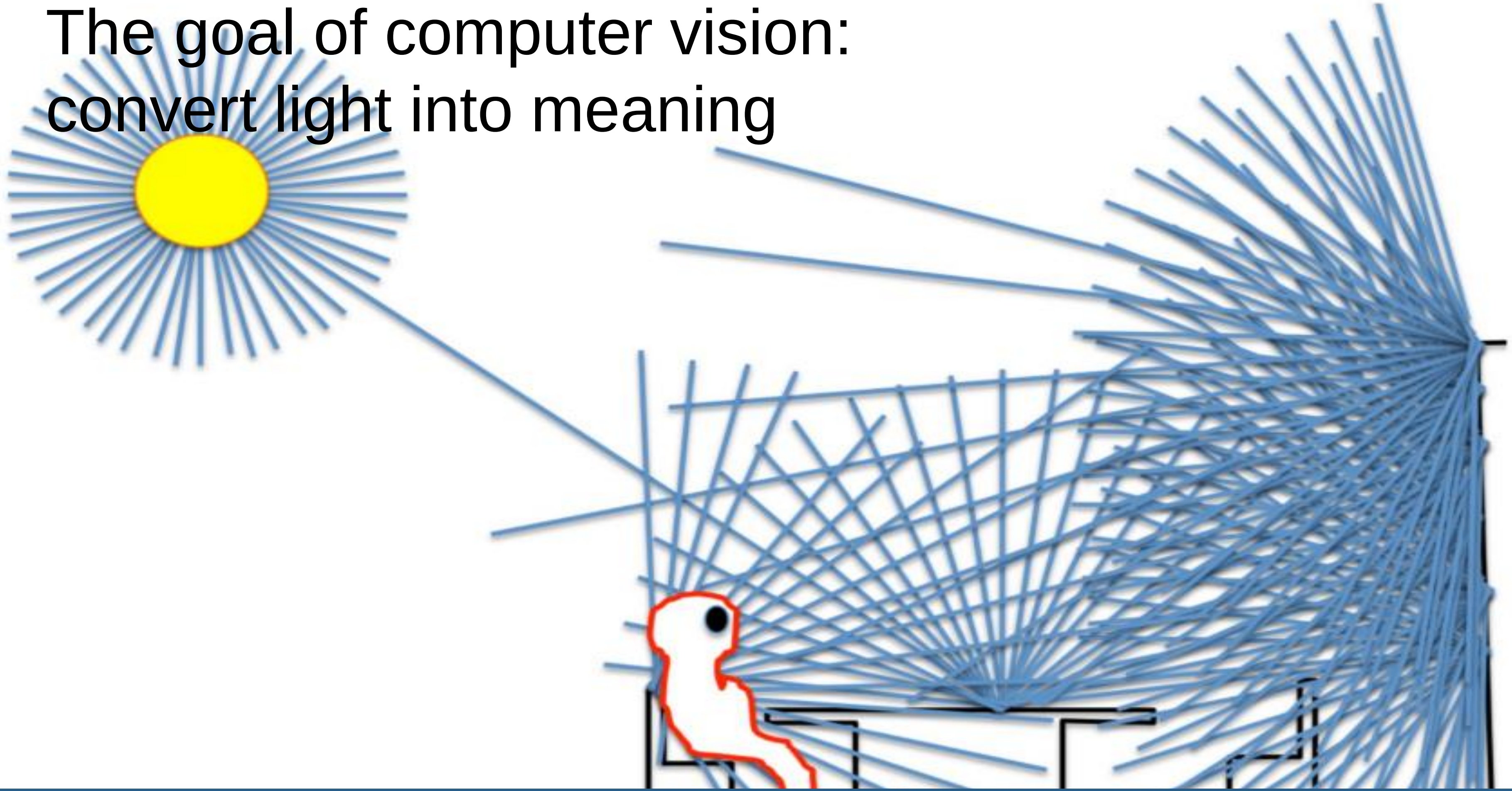
Today's agenda

- History of understanding perception
- Introduction to computer vision
- Course overview

Today's agenda

- History of understanding perception
- Introduction to computer vision
- Course overview

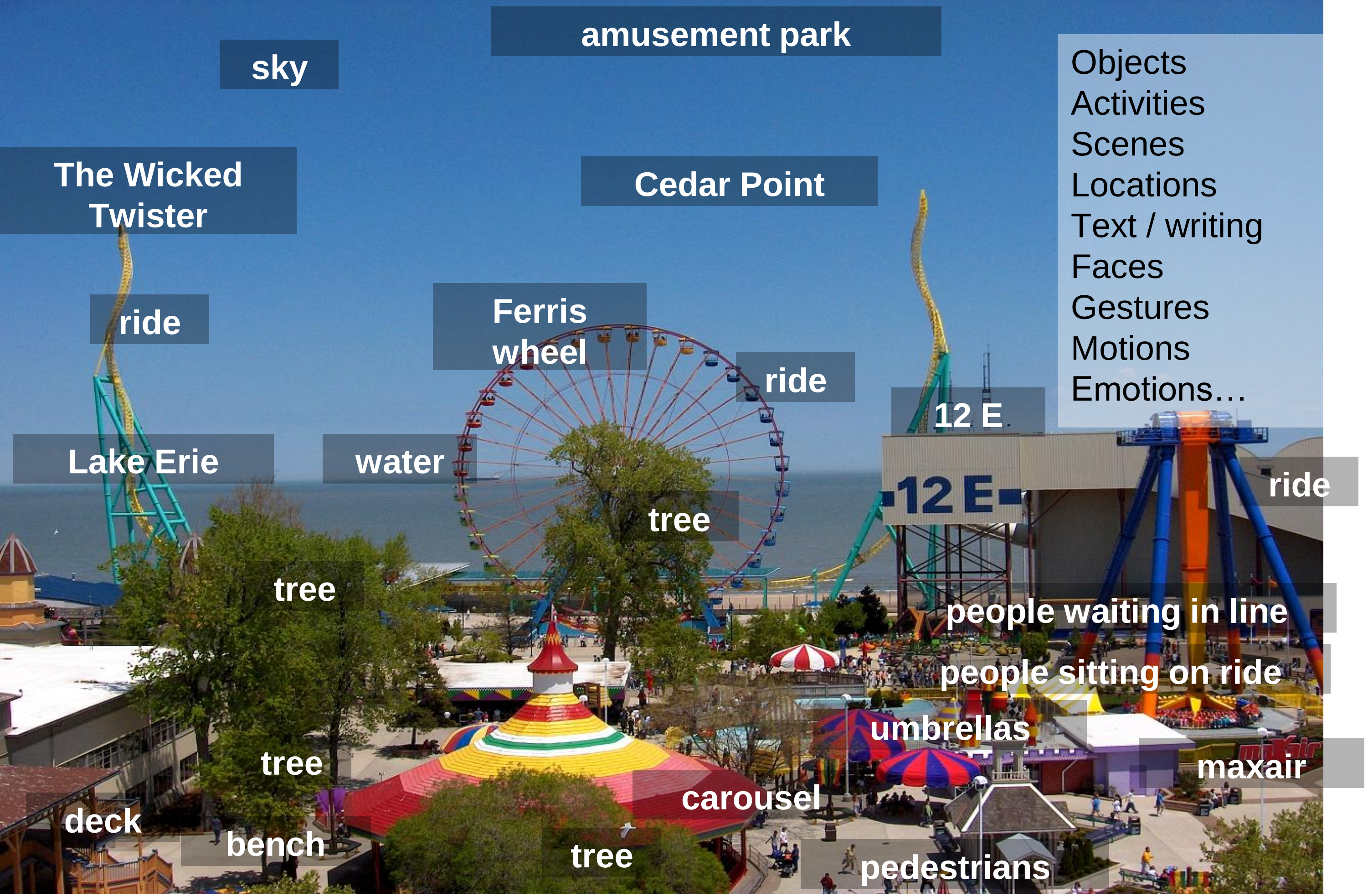
The goal of computer vision: convert light into meaning



What kind of information can we extract from an image?

1. Semantic information
2. Geometric 3D information

Vision as a source of semantic information

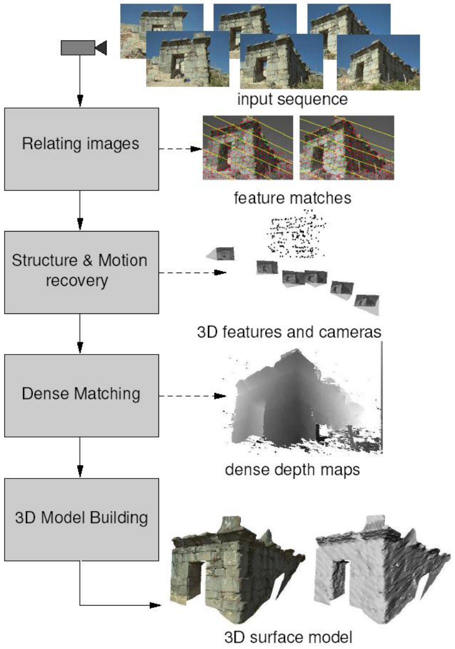


Extracting Semantics

Segment Anything
2023

Extracting geometric information

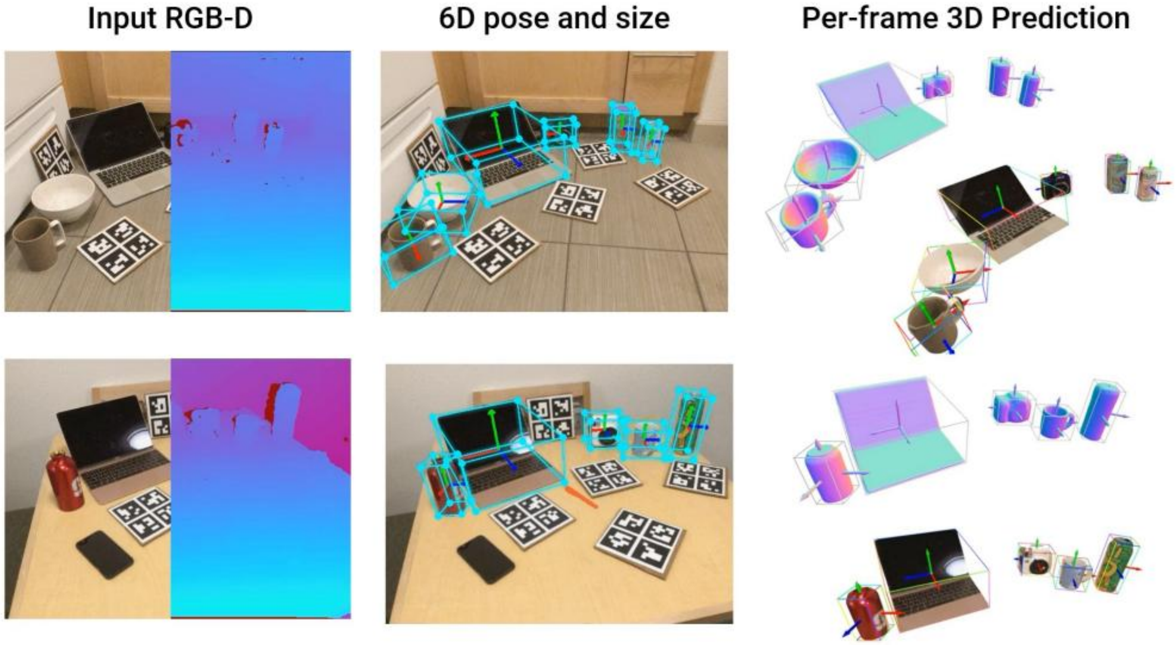
Real-time stereo



Pollefeys et al.

Goesele et al.

Geometric 3D information from 2D images

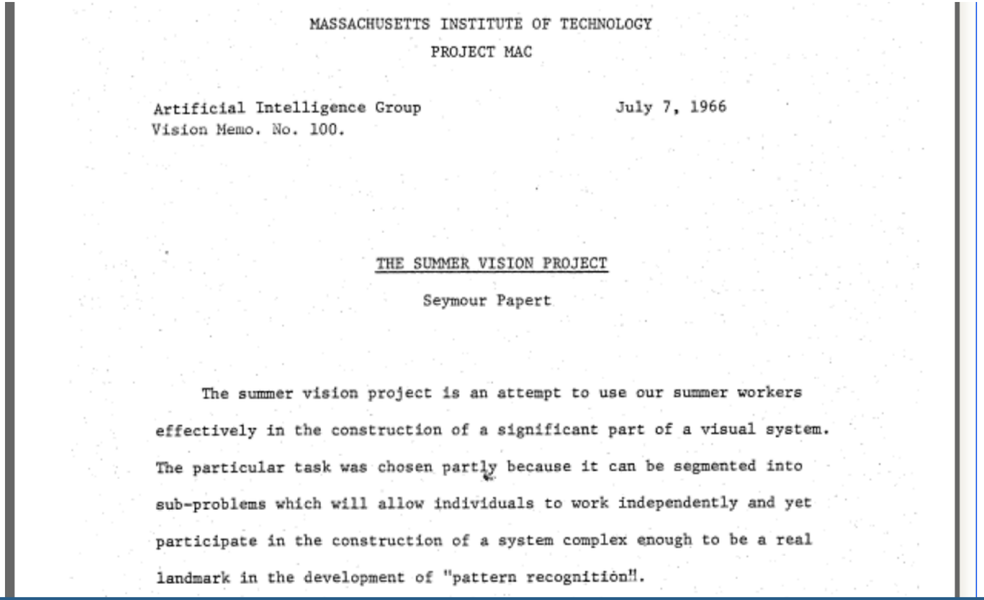


TRI & GATech’s ShaPO (ECCV’22): https://zubair-irshad.github.io/projects/ShAPO.html

MIT thought that computer vision would be solved as an undergraduate summer project

"The primary goal of the project is to construct a system of programs which will divide a [...] picture into regions such as likely objects, likely background areas and chaos."

"The final goal is OBJECT IDENTIFICATION which will actually name objects by matching them with a vocabulary of known objects."


MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

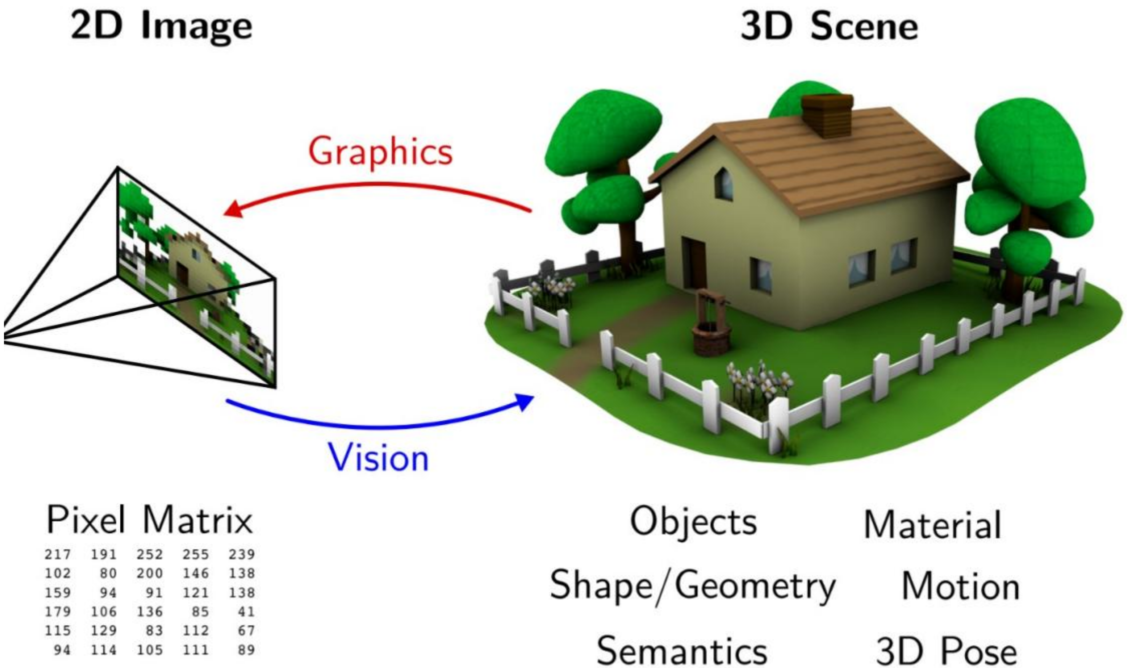
Artificial Intelligence Group July 7, 1966
Vision Memo. No. 100.

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".

But why is computer vision so hard?



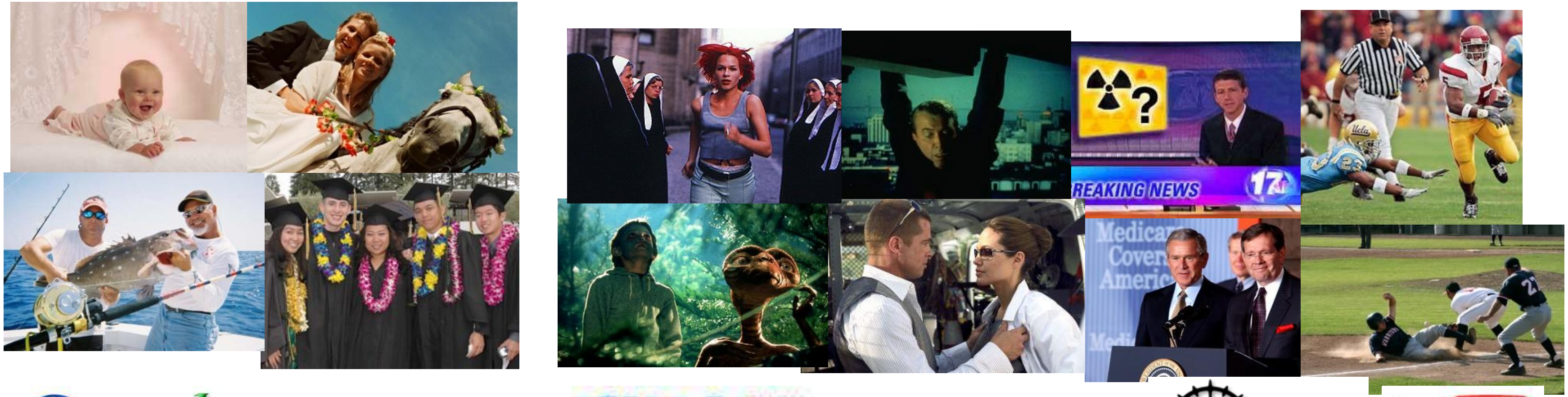
It is an ill posed problem

Computers need to convert pixel intensities into meaning

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

Why study computer vision?

Vision is useful: Images and video are everywhere!



Google Image Search

Google Photos

flickr

webshots

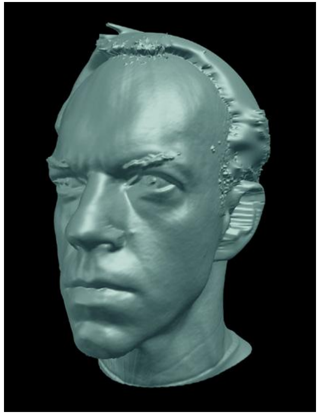
picsearch

YouTube Broadcast Yourself

80% of all web traffic is images and videos

Majority of the internet is dark matter without computer vision

Special effects: shape and motion capture



3D urban modeling

Google Streetview - custom campus tours

3D urban modeling: Microsoft Photosynth

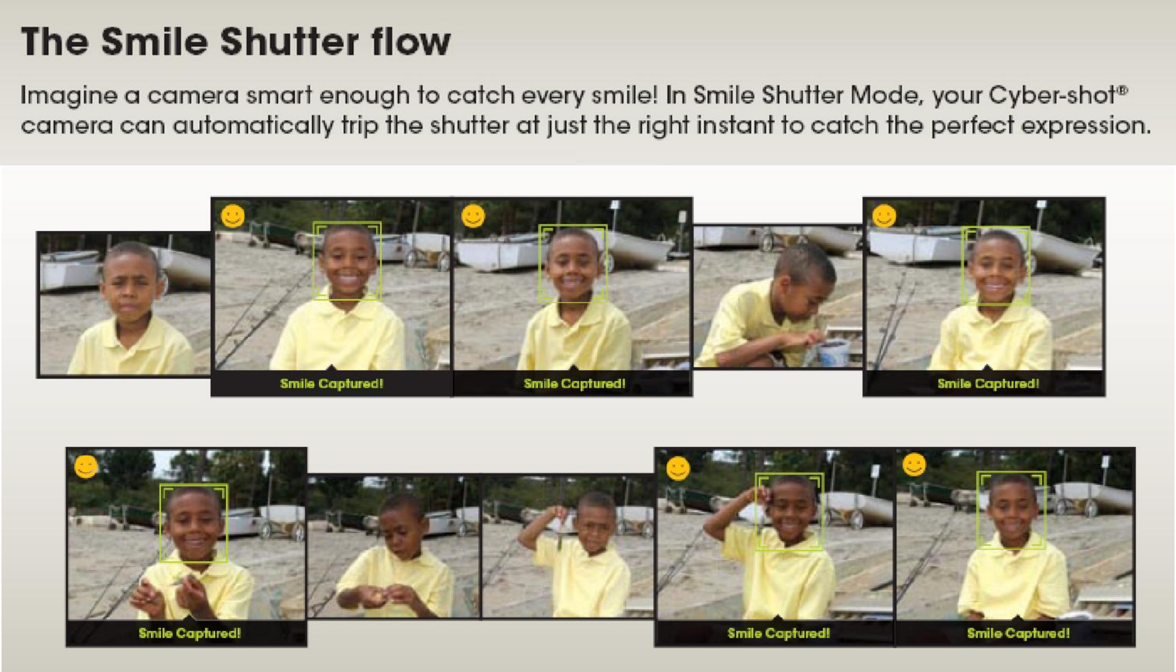
Formerly PhotoTourism
Developed at UW.

http://photosynth.net

Face detection

- Many digital cameras now detect faces
    - Canon, Sony, Fuji, …

Smile detection

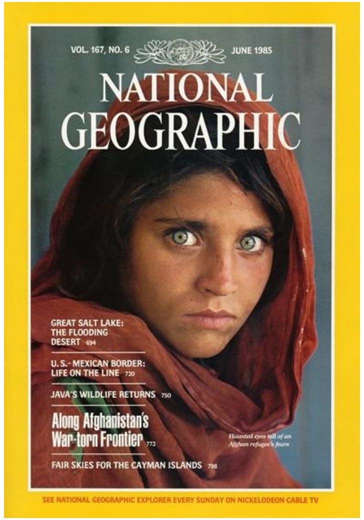


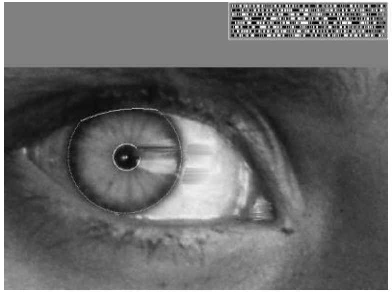
Sony Cyber-shot® T70 Digital Still Camera

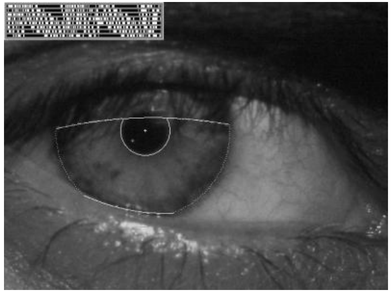
Face recognition: Apple iPhoto software

Biometrics

How the Afghan Girl was Identified by Her Iris Patterns







Biometrics

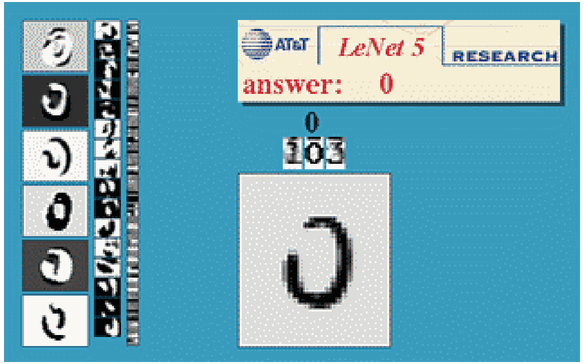
Fingerprint scanners on many new laptops, other devices

Face recognition systems now on iphones and samsungs

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs

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

Google maps: Annotate all houses and streets

Avenue des Sapins

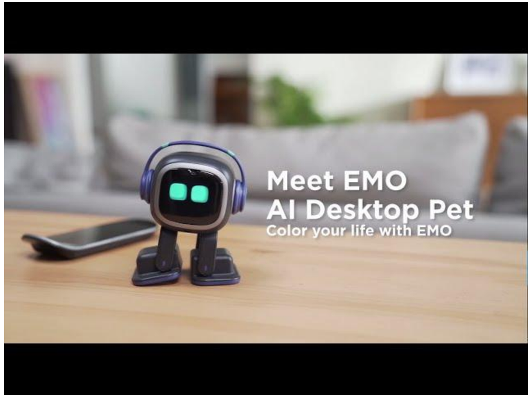
Goodfellow et al. 2014

Vision-powered toys and robots in the 2000s

Vision-powered toys and robots in the 2020s

Scout home security robot - monitors your house

Unitree Go1's companion robot - like a dog



Apple Vision Pro, Snapstacles and Google glasses

Ranjay in undergrad ->

Automotive safety

Mobileye: Vision systems in high-end BMW, GM, Volvo models

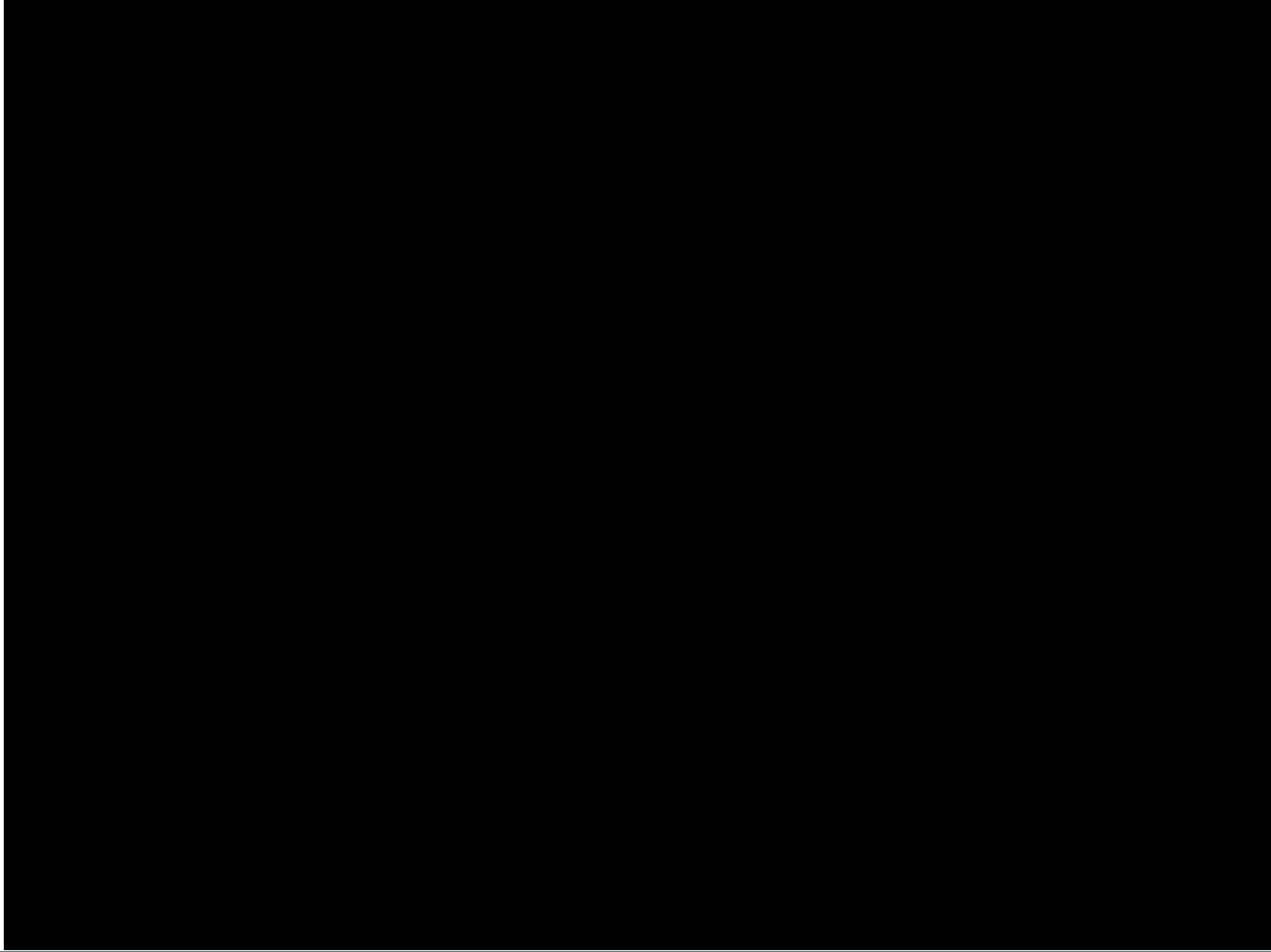
Claimed that they would release self-driving cars by 2015. Still not there.

Vision in supermarkets

LaneHawk by EvolutionRobotics
(acquired by iRobot for $74M in 2012)

"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it… "

Amazon Go



Vision-based interaction (and games)

Microsoft's Kinect

Assistive technologies

Sony EyeToy

Augmented Reality

Virtual Reality

Vision for robotics, space exploration

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

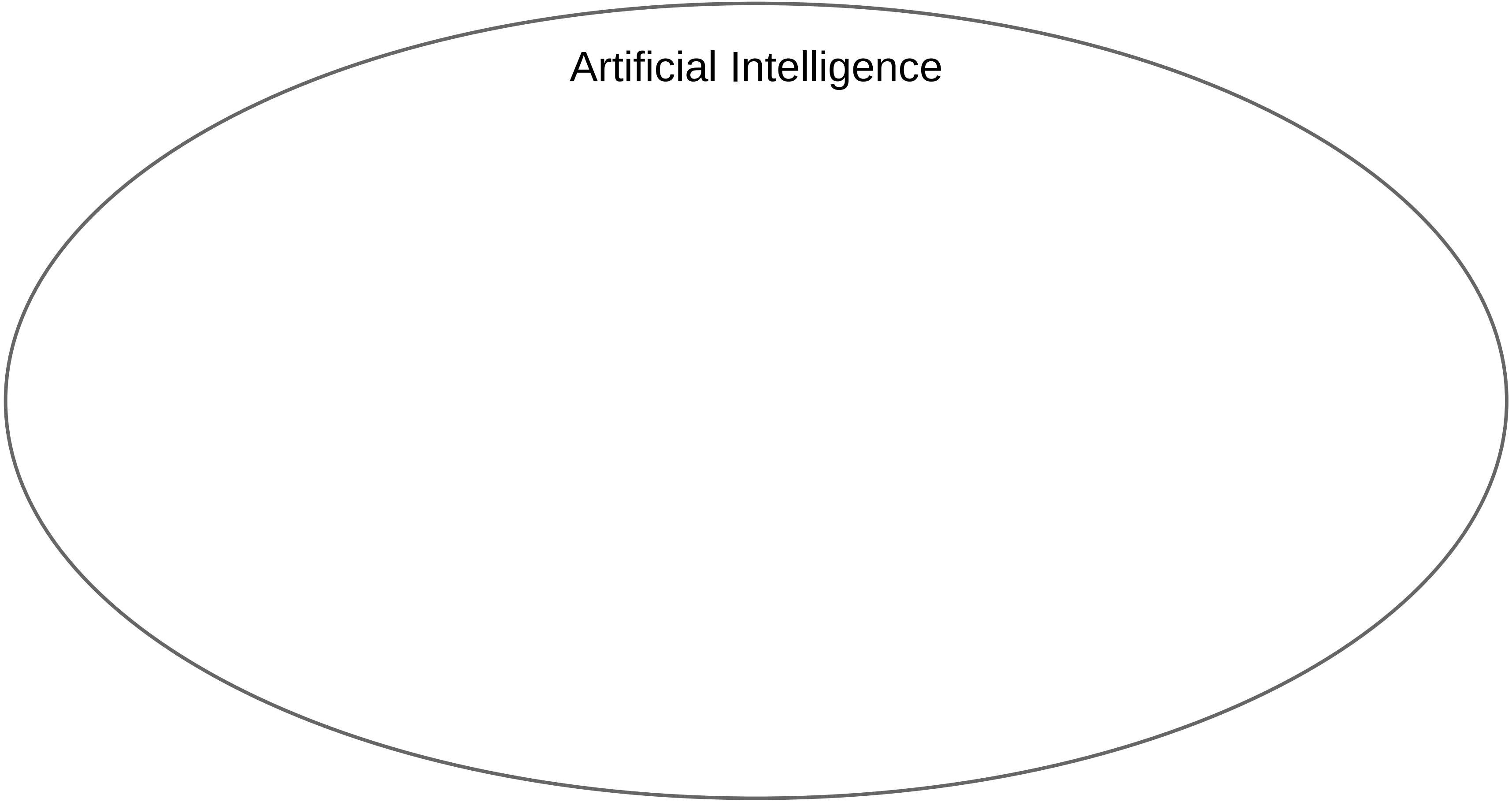
- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “Computer Vision on Mars” by Matthies et al.

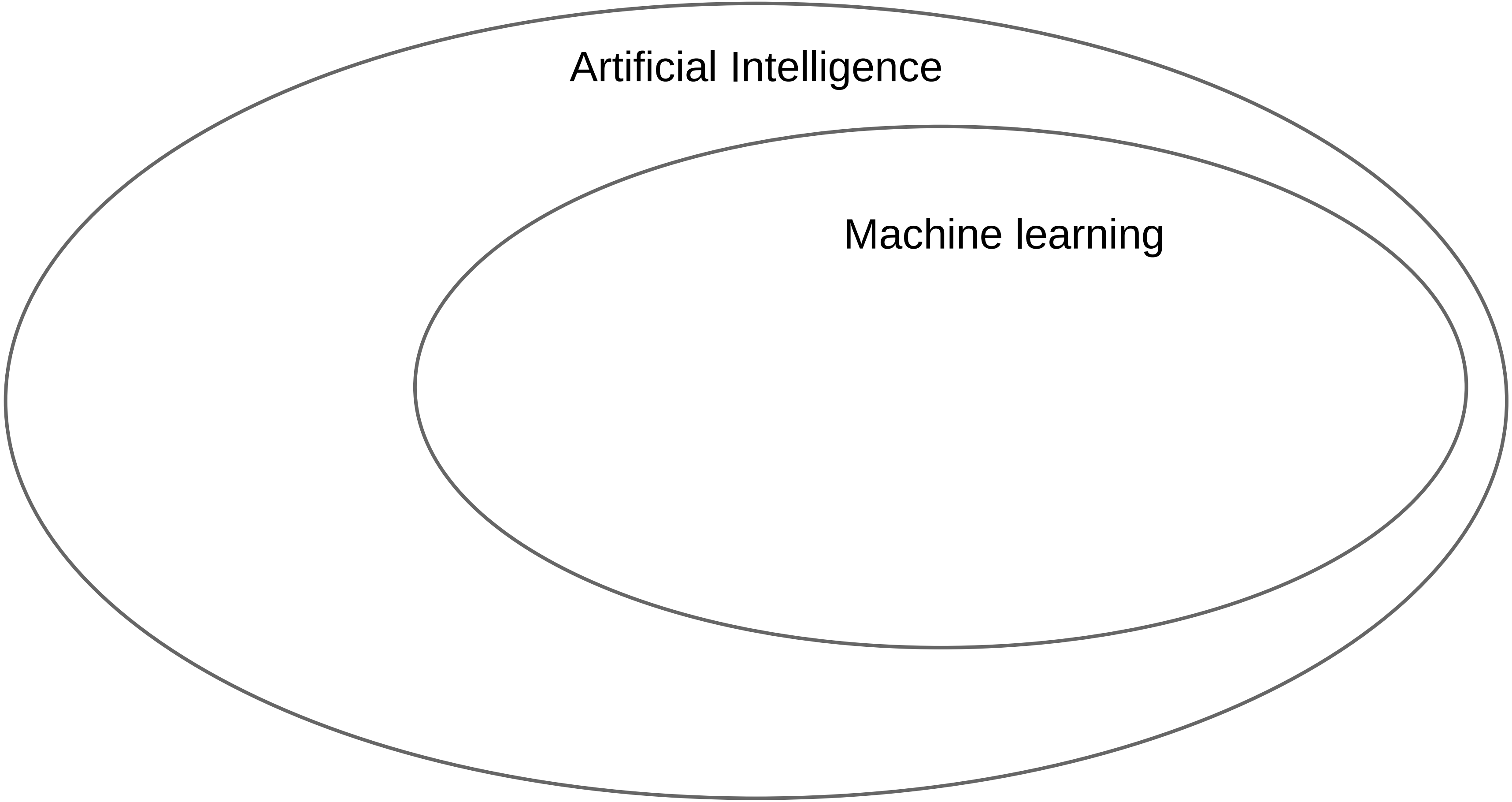
e: S. Seitz

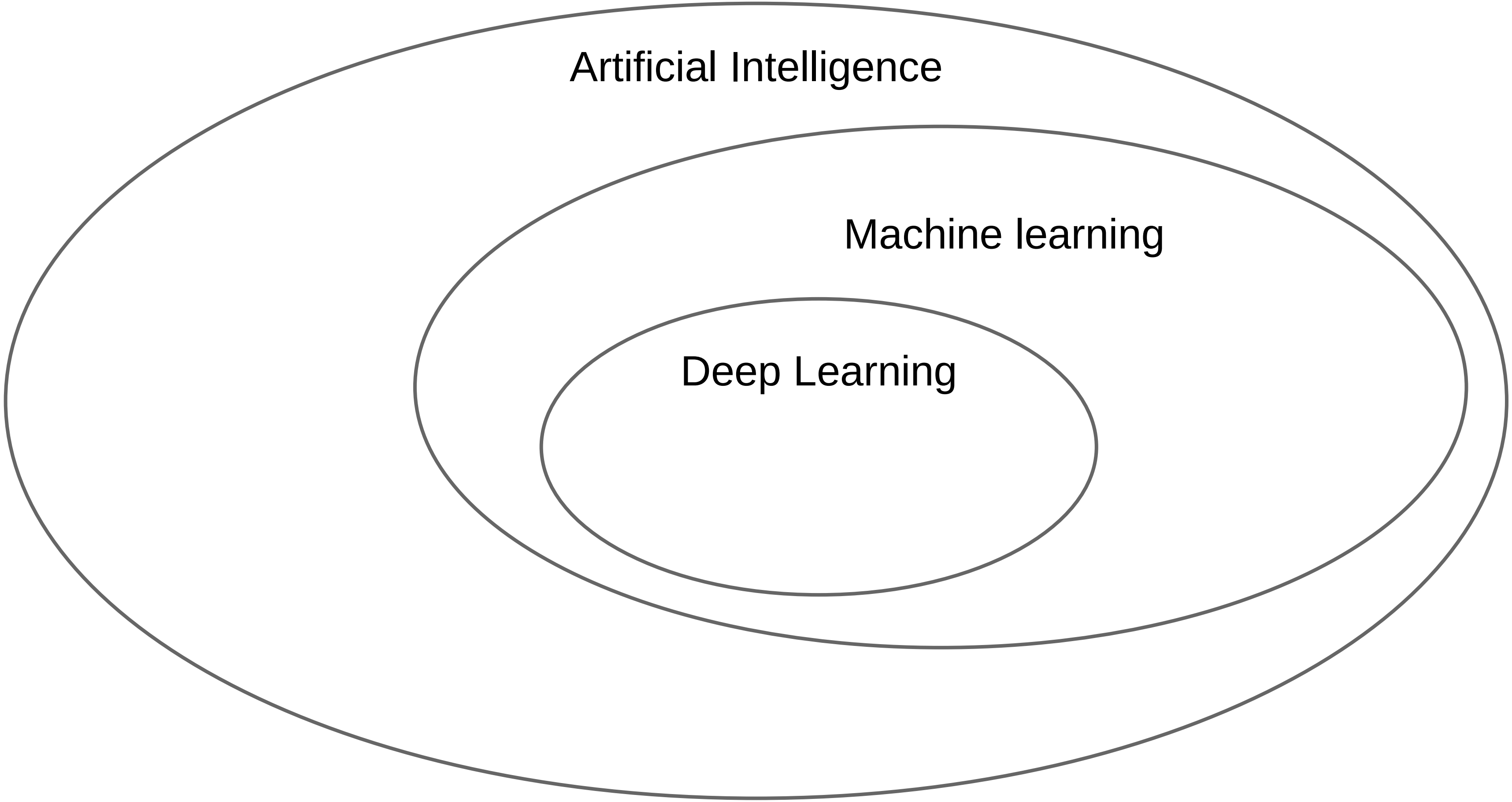

Machine Bias
There's software used across the country to predict future criminals. And it's biased against blacks.

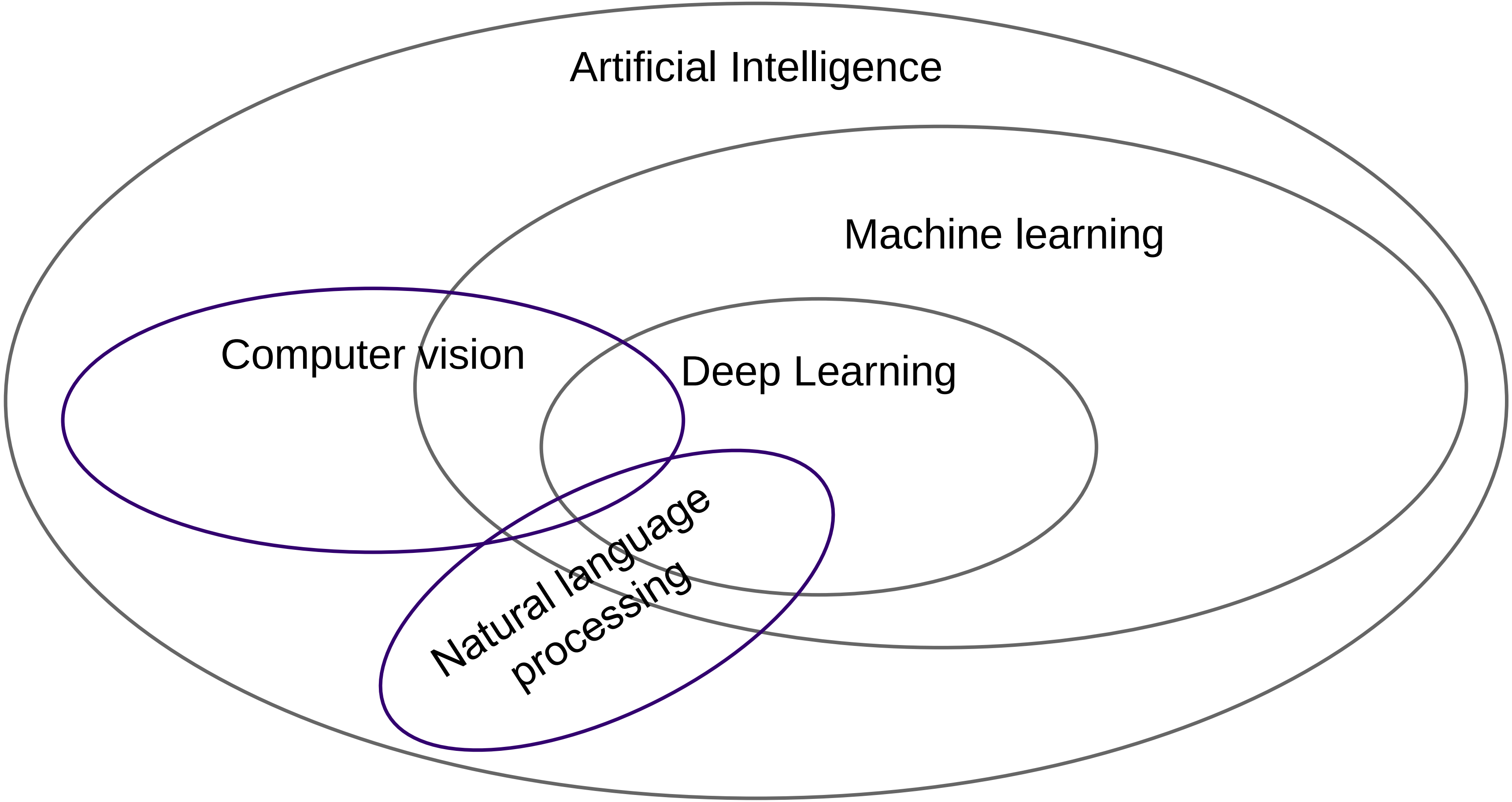
How should you make sense of computer vision as a field?

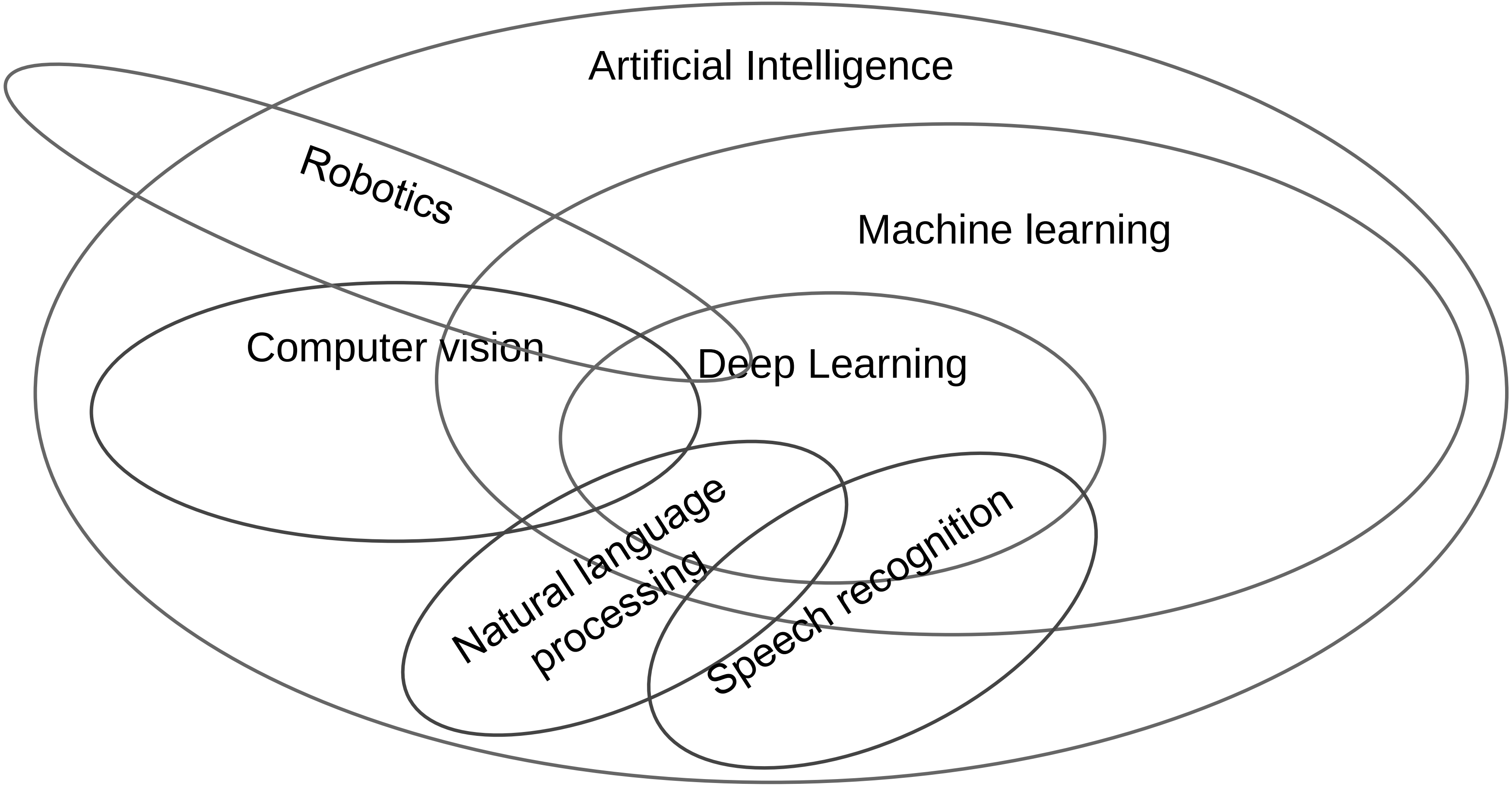
Let’s situate computer vision in the broader context of AI

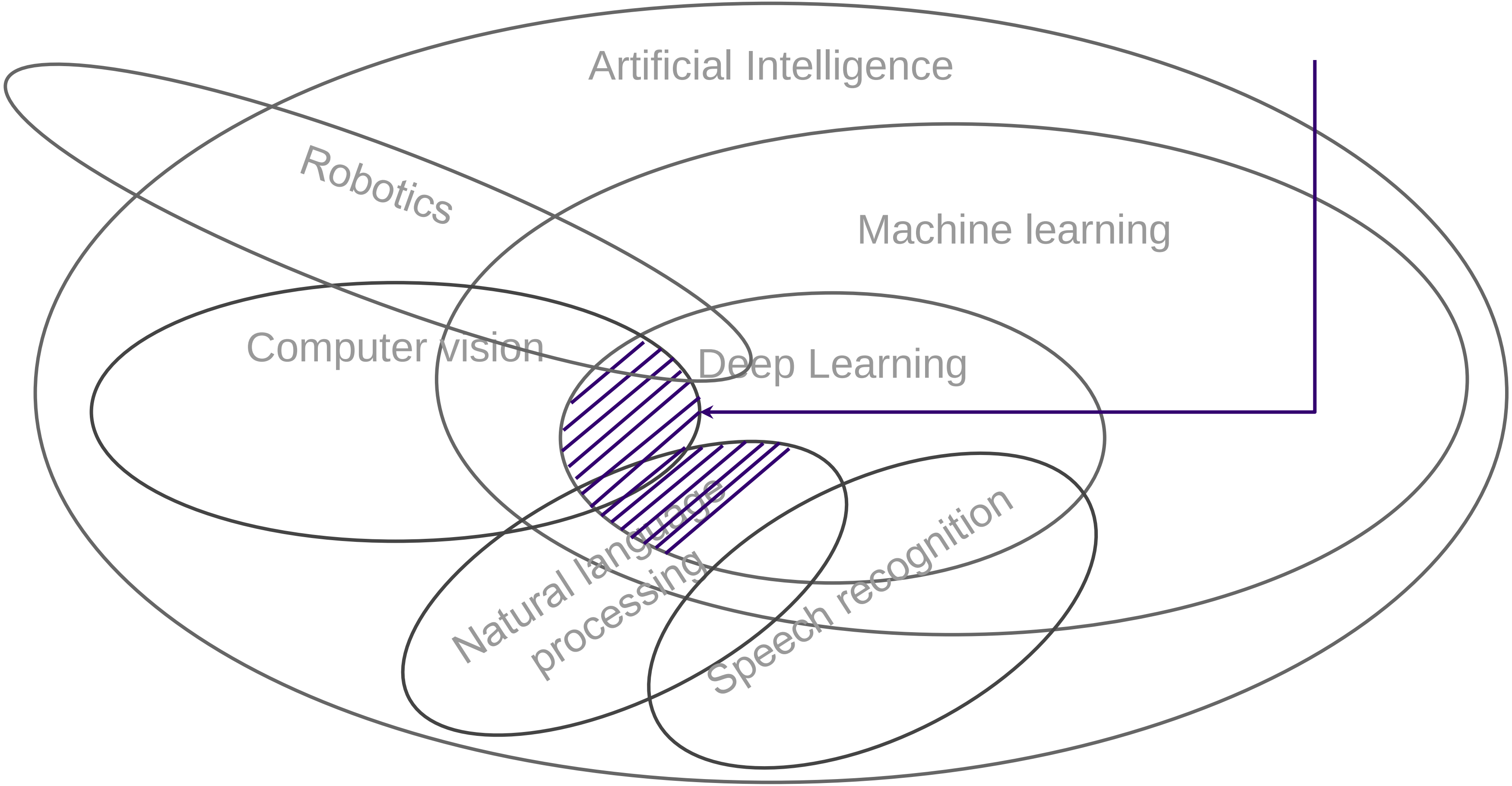

Artificial Intelligence


Artificial Intelligence
Machine learning

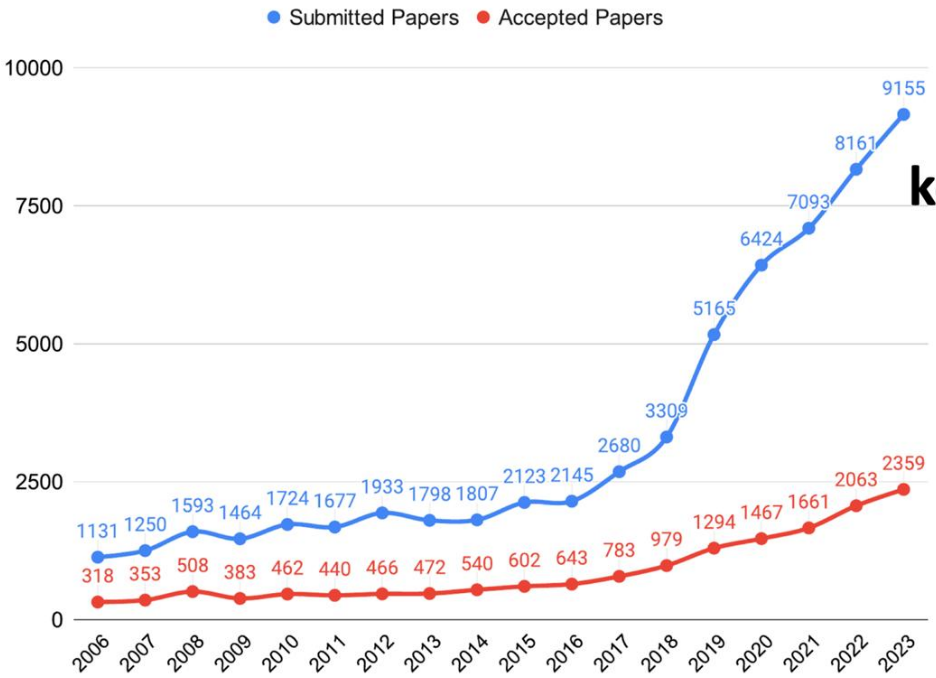

Artificial Intelligence
Machine learning
Deep Learning


Artificial Intelligence
Machine learning
Computer vision
Deep Learning
Natural language processing


Artificial Intelligence
Robotics
Machine learning
Computer vision
Deep Learning
Natural language processing
Speech recognition


Artificial Intelligence
Robotics
Machine learning
Computer vision
Deep Learning
Natural language processing
Speech recognition

CVPR has seen a large number of deep learning people enter



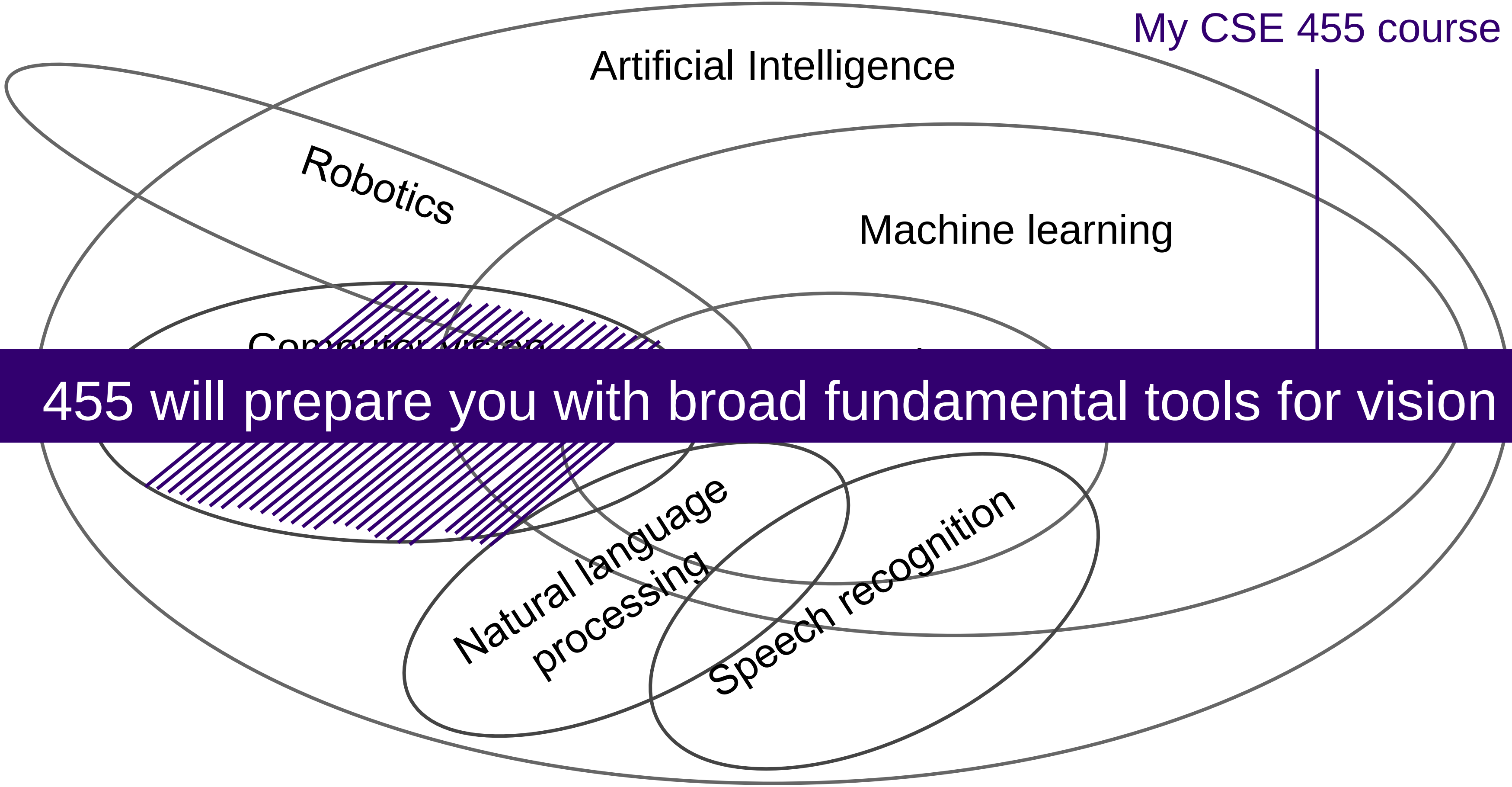
CVPR

kept on growing...

What is CVPR? Best computer vision conference.

What does it mean? Computer Vision and Pattern Recognition

Who helped start it? Professors Shapiro and Tanimoto



455 will prepare you with broad fundamental tools for vision

Decade by decade

- **1960s**: Image processing and pattern recognition, blocks world
- **1970s**: Key recovery problems defined: structure from motion, stereo, shape from shading, color constancy. Attempts at knowledge-based recognition
- **1980s**: Fundamental and essential matrix, multi-scale analysis, corner and edge detection, optical flow, geometric recognition as alignment
- **1990s**: Multi-view geometry, statistical and appearance-based models for recognition, first approaches for (class-specific) object detection
- **2000s**: Local features, generic object recognition and detection
- **2010s**: Deep learning, big data

Adapted from J. Malik

CVPR 2024 was here in June 2024 (2025 was in Nashville)

https://cvpr.thecvf.com/

Why should you go to CVPR? It is ranked #4 amongst all scientific publications across all disciplines

	Publication	h5-index	h5-median
1.	Nature	467	707
2.	The New England Journal of Medicine	439	876
3.	Science	424	665
4.	IEEE/CVF Conference on Computer Vision and Pattern Recognition	422	681
5.	The Lancet	368	688
6.	Nature Communications	349	456
7.	Advanced Materials	326	415
8.	Cell	316	503

Source: Google scholar

Today's agenda

- History of computer vision
- Introduction to computer vision
- Course overview

Course staff

Instructor: Linda Shapiro

TAs

Nishat Khan Wisdom Ikezogwo Rustin Soraki Raymond Yu Weikaih Huang

Class times

Lectures

- Tuesdays and Thursdays
  11:30-12:50am @ CSE2 G01

Recitations

- Friday afternoons 12:30-1:20pm @ SIG 134

Lecture recordings

Will be made available on canvas:

https://canvas.uw.edu/courses/1844262

Come to class!

Contacting instructor and TAs

- All announcements, Q&A in EdStem
  - https://edstem.org/us/courses/84905/discussion
  - All course related posts should be public.

- All private correspondences to course staff should post private (instructors only) post on EdStem.
  - Use this for personal problems, and debugging help to avoid showing other people your solutions.
  - If you have questions that others can benefit from, do a public post.

How to think about computer vision?

Breadth

- Computer vision is a huge field
- It can impact every aspect of life and society
- It is driving the current generative AI revolution
- Pixels are everywhere in our lives and cyber space
- CSE455 is meant as an broad overview course,
  - we will not cover all topics of CV
- Lectures are mixture of detailed techniques and high level ideas

Depth

- …

How to think about computer vision?

Breadth

- …

Depth

- Computer vision is a highly technical field, i.e. know your math!
- Master bread-and-butter techniques: face recognition, corners, lines, features, optical flows, clustering and segmentation
- Programming assignments: be a good coder AND a good writer
- Math problem questions: know your concepts!
- Final Project: your chance to shine!

Syllabus

Official website

https://courses.cs.washington.edu/courses/cse455/25au/

Winter and Spring 2025 we changed the entire course. We will continue to improve it

All assignments, lectures, etc. changed in 2025.

What is new?

All assignments are new!

- No more coding in C. Everything is in Python
- We moved everything to Google Colab

All slides are all new.

- There is some overlap with topics from previous years but taught differently
- This quarter will be a mix of Prof. Krishna's slides and some of Prof. Shapiro's.

Grading policy (tentative)

75%: 5 Assignments.

- **0%** for Assignment 0
- **20%** for Assignment 1
- **15%** each for Assignment 2, 3, & 4
- **10%** for Assignment 5

25%: 1 Final Project.

10%: Extra Credit - in assignments.

Grading policy – Assignments (tentative dates)

- **Assignment 0** (Using Colabs, Python basics)
  - Recommended Due by Oct 2 (Ungraded)
- **Assignment 1** (Filters, Convolutions, Edges)
  - Due Oct 14, 11:59 PST
- **Assignment 2** (Keypoints, Panoramas)
  - Due Oct 28, 11:59 PST
- **Assignment 3** (Cameras, Clustering, Segmentation)
  - Due Nov 12, 11:59 PST
- **Assignment 4** (kNN, PCA, LDA, Detection)
  - Due Nov 25, 11:59 PST
- **Assignment 5** (CNNs)
  - Due Dec 4, 11:59 PST

Grading policy - assignments

- Most assignments will have an extra credit worth 1% of your total grade.

- Late policy
    - 5 free late days – use them in your ways
    - Maximum of 2 late days per assignment
    - Afterwards, 10% off per day late

- Collaboration policy
    - Read the student code book, understand what is ‘collaboration’ and what is ‘academic infraction’
    - We have links to this on the course webpage

Submitting homeworks

- **Homeworks** will consist of python files with code and jupyter notebooks.

- **Jupyter notebooks**:
  - Will guide you through the assignments.
  - Might contain written questions
  - Once you are done, convert the ipython notebook into a pdf and submit on Gradescope (https://www.gradescope.com/courses/1104428).
    - Access code:

- **Python files**:
  - All code must be submitted to Gradescope as well.
  - Check our course website for details on submissions.

- **A0** will be live soon, you can start working on it immediately. We will try and get all the assignments out to you as soon as they are ready.

Final Project

- There will be a final project worth 25% of your final grade.
- The project should be done in **teams of 3-4.**
- **Projects must involve MACHINE LEARNING, but otherwise are up to you to define.**
- It will be assigned Oct 16.
- Groups and proposals due Oct 23.
- Reports due Dec 8 and presentations Dec 10 during final exam slot.

CSE 455 Roadmap

Pixels	Segments	Images	Videos	Web
Convolutions Edges Descriptors	Resizing Segmentation Clustering	Recognition Detection Machine learning	Motion Tracking	Neural networks Convolutional neural networks

From Convolutions to Convolutions

Who is Linda Shapiro?



1975 2008 2025

- Boeing Endowed Professor in Computer Science & Engineering
- Professor of Electrical & Computer Engineering
- Adjunct Professor of Biomedical Informatics & Medical Education

- 51 years of teaching and research in computer vision
- 53 Ph.D. graduates who are professors, industrial researchers, or retired!

- Previous Chair of CVPR (twice) and also Program Chair
- Previous Editor or Associate Editor of several computer vision journals

- First and only woman to be Chair of the IEEE PAMI Technical Committee that runs the major vision conferences

- Working mainly in medical image analysis but also robot vision