THE UNIVERSITY OF TEXAS AT DALLAS

Introduction to Computer Vision

CS 4391 Computer Vision

Professor Yapeng Tian

Department of Computer Science

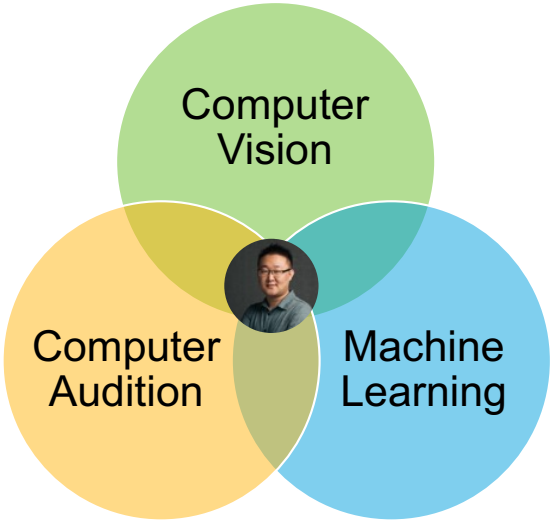
Instructor

Dr. Yapeng Tian

Assistant Professor in UTD@CS

Research area:

- Computer Vision
- Computer Audition
- Machine Learning



yapeng.tian@utdallas.edu

yapengtian.com

Today

- What is computer vision?
- Why study computer vision?
- Why is computer vision difficult?
- Course overview

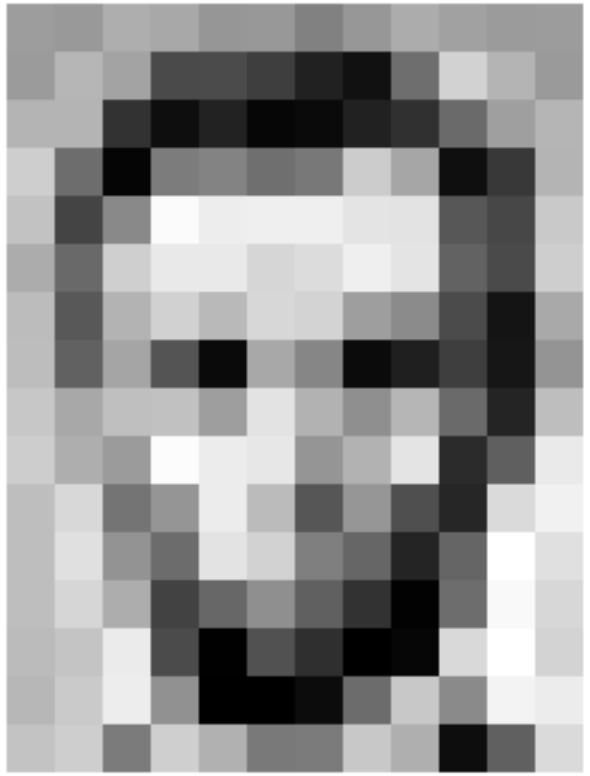
Every Image Tells a Story

https://unwritten-record.blogs.archives.gov/2019/04/11/a-picture-is-worth-a-thousand-words/

- Primary goal of computer vision: perceive the “story” behind the picture by machines
- Compute properties of the world
  - 3D shape
  - Names of people or objects
  - What happened?

Slide Credit: Noah Snavely

The Goal of Computer Vision



157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

A gray image is represented by a 2D matrix in computer, and each pixel value is in [0, 255].

The Goal of Computer Vision

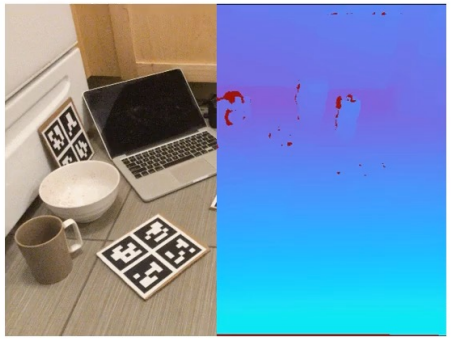
- Compute the 3D shape of the world

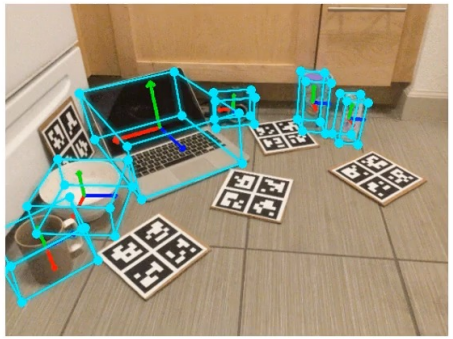
Kinect RGB-D camera

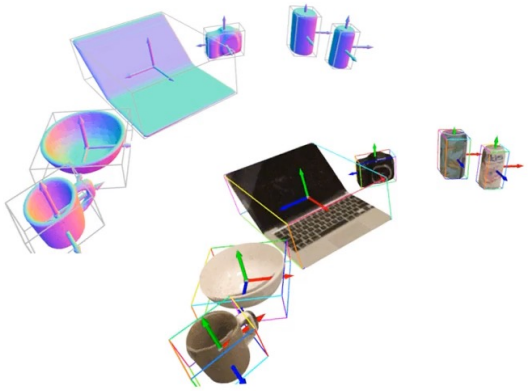
Input RGB-D

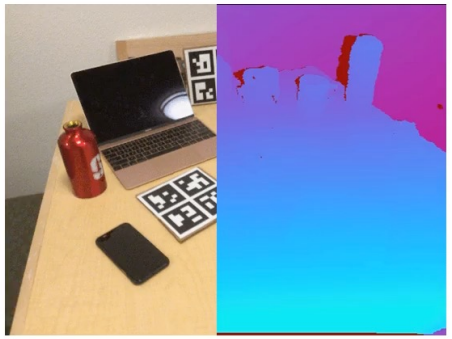
6D pose and size

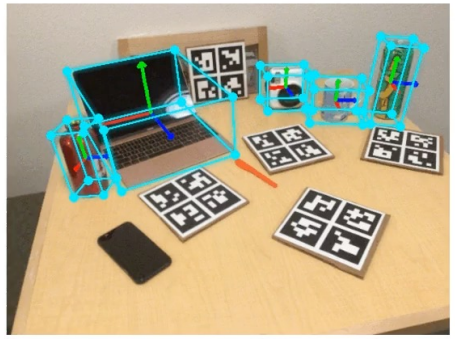
Per-frame 3D Prediction

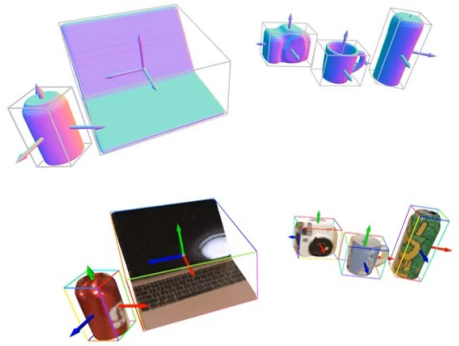












https://zubair-irshad.github.io/projects/ShAPO.html

The Goal of Computer Vision

- Recognize persons and objects

The Goal of Computer Vision

- Describe visual content

a man on a skateboard jumping over a railing.

https://huggingface.co/spaces/SRDdev/Image-Caption

The Goal of Computer Vision

- Enhance photo quality

Image super-resolution

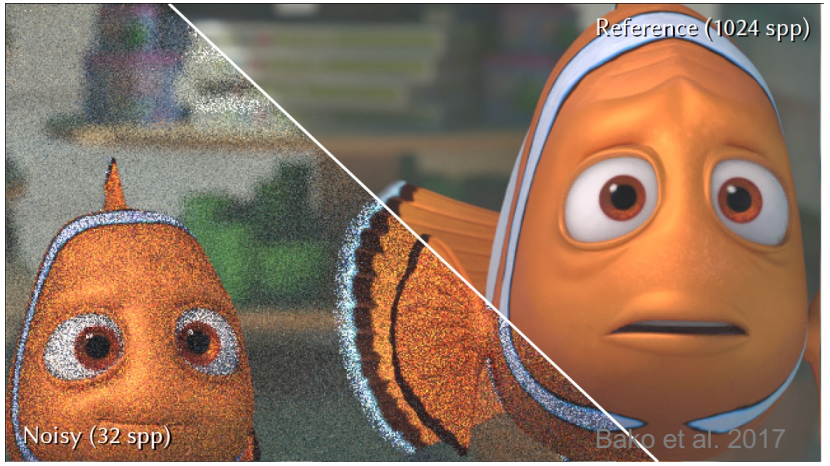


Image denoising

Low-light image enhancement

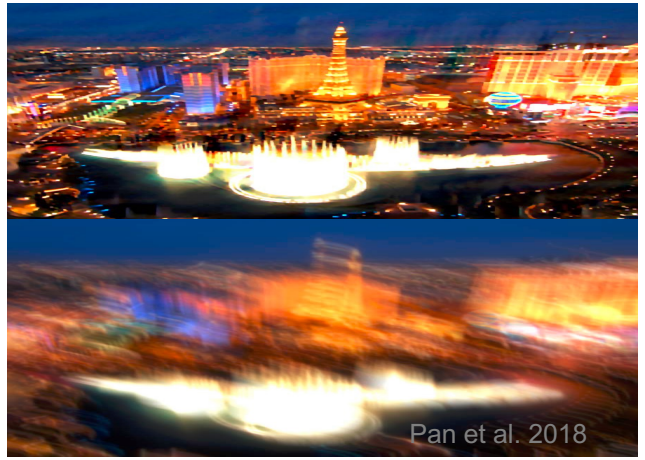


Image deblurring

Image deraining

The Goal of Computer Vision

- Manipulate photos

Image inpainting [theinpaint.com]

Style transfer [Gatys et al. 2016]

The Goal of Computer Vision

Which image was produced by humans?

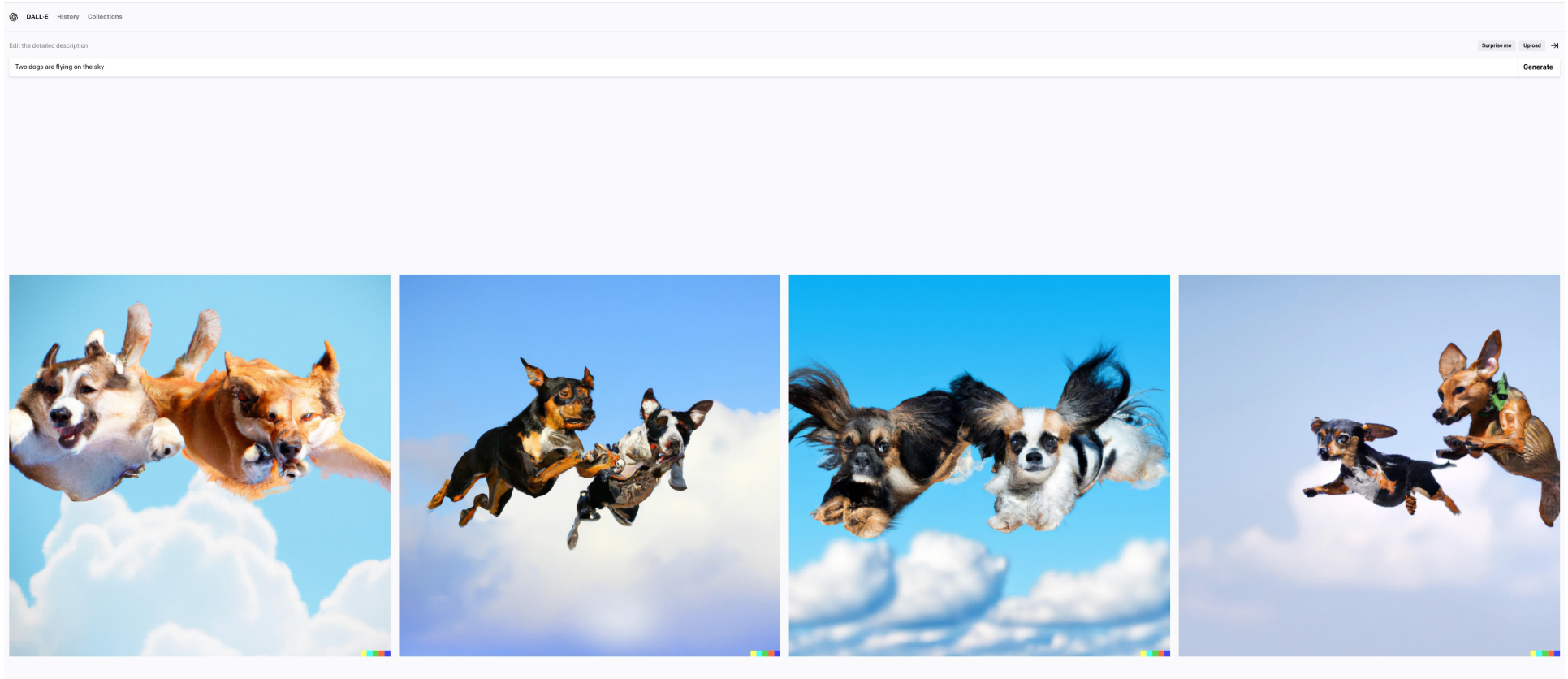
The Goal of Computer Vision

- Generate visual content by AI

DALL-E 2@OpenAI

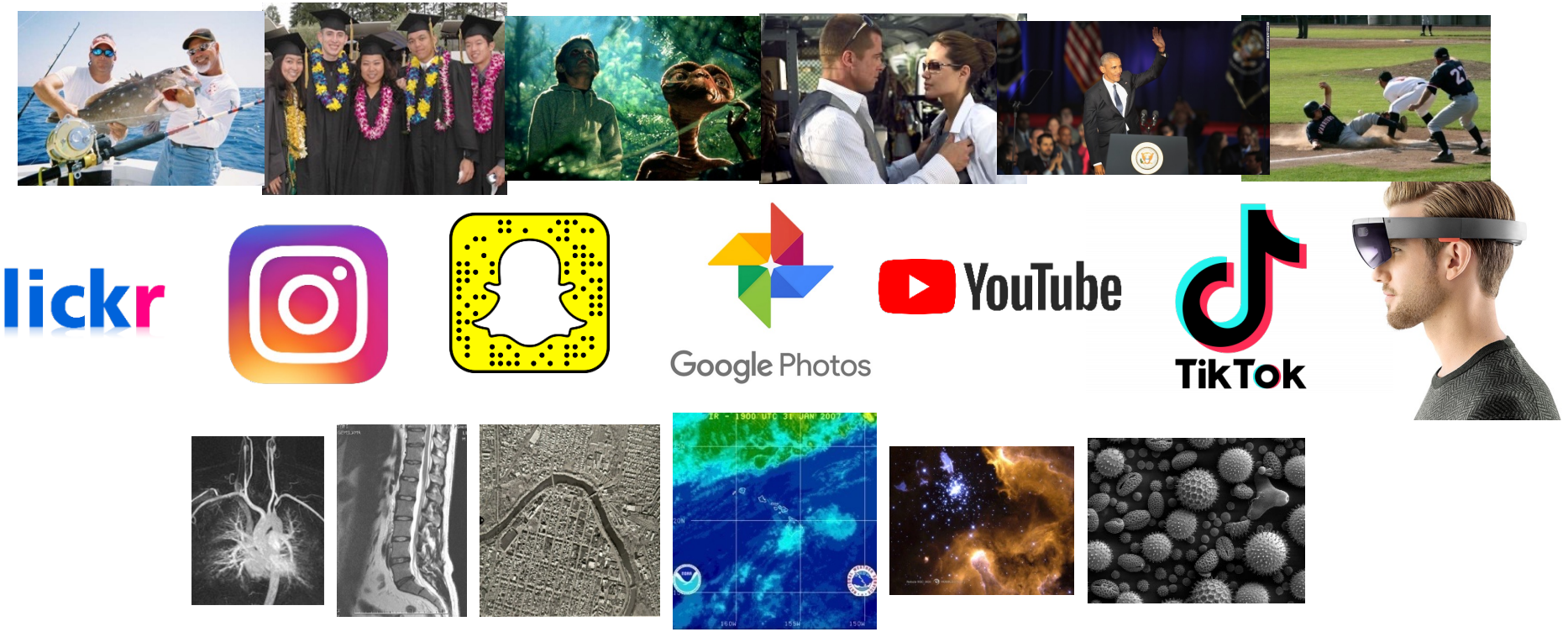
The Goal of Computer Vision

- Generate visual content by AI



Why study computer vision?

- Billions of images/videos captured per day

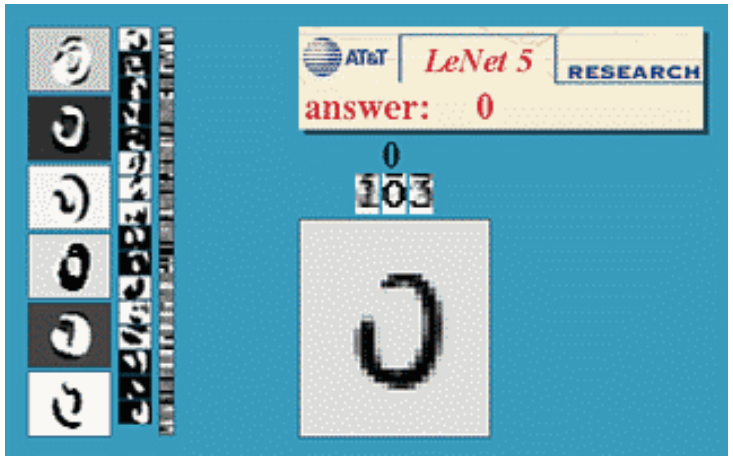


- Huge number of potential applications

Slide Credit: Noah Snavely

Optical Character Recognition (OCR)

- recognize text from scanned images and documents



Digit recognition, Yann LeCun. (1990's)

Automatic check processing

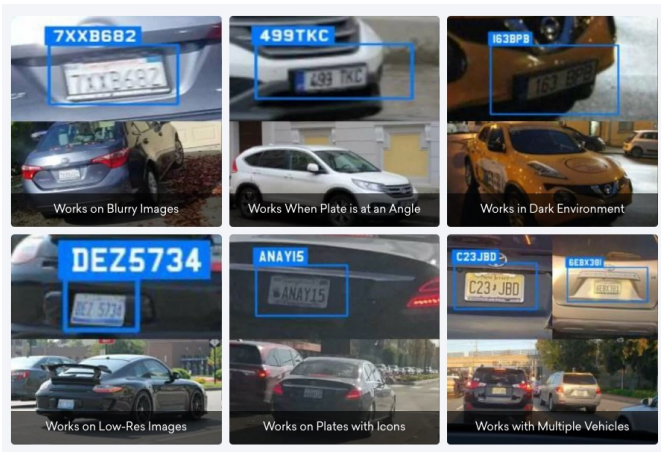
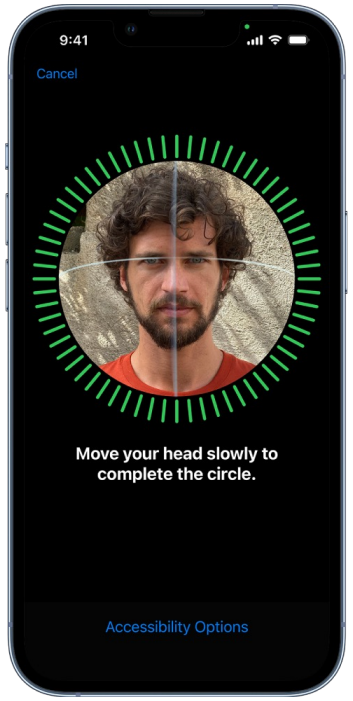


Plate recognition

Biometric

Biometric techniques are methods used to identify individuals based on their physical or behavioral characteristics.

https://modelcards.withgoogle.com/face-detection



FaceID

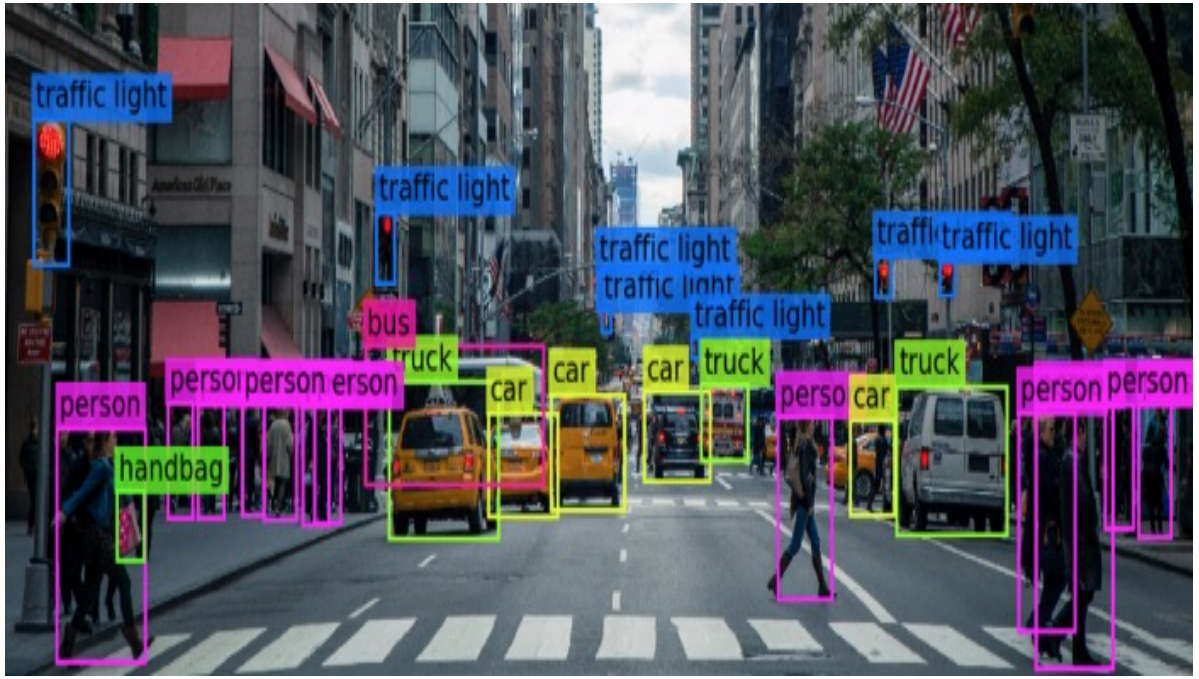
A facial recognition system for law enforcement [Credit: Saul Loeb]

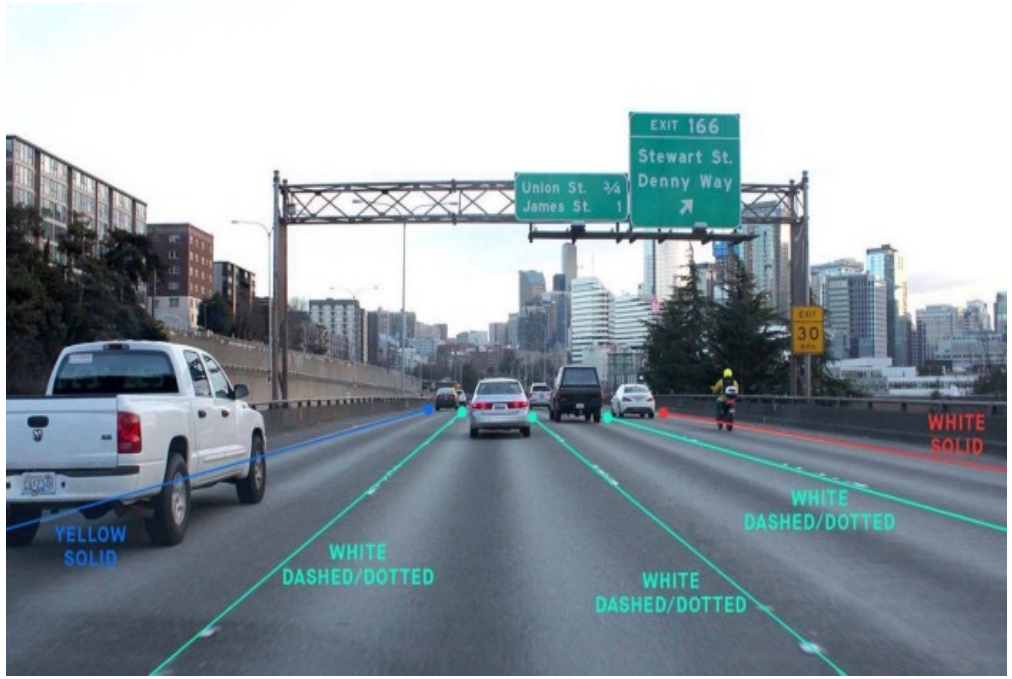
Access control

Face Detection and Recognition

Autonomous Driving

- Detect persons, cars, and lanes on roads and streets





Source: https://medium.com/@safk8899/computer-vision-in-autonomous-vehicles-21dffa873b23

Tesla Self-Driving Car

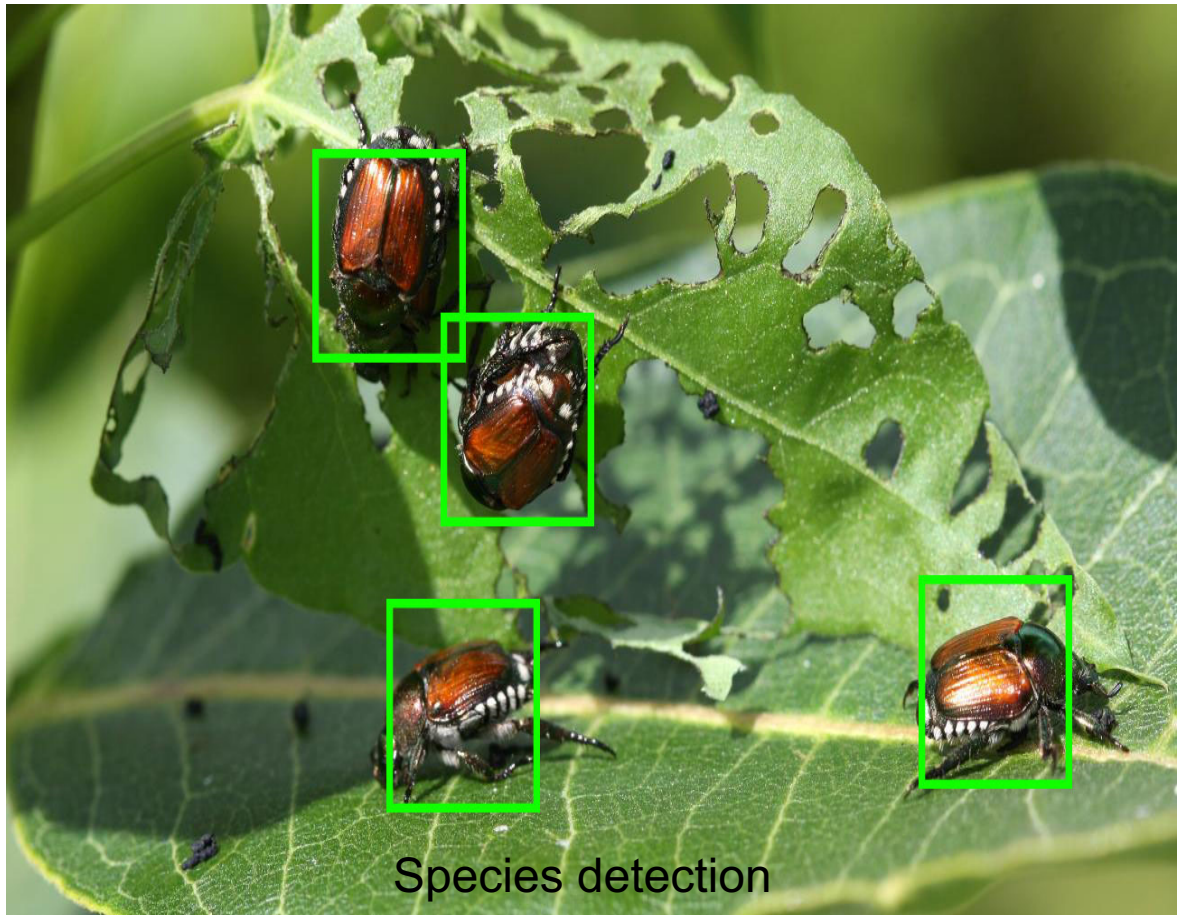
https://vimeo.com/192179726?embedded=true&source=vimeo_logo&owner=128712855

Robotics

Automotive Manufacturing

Delivery Robots

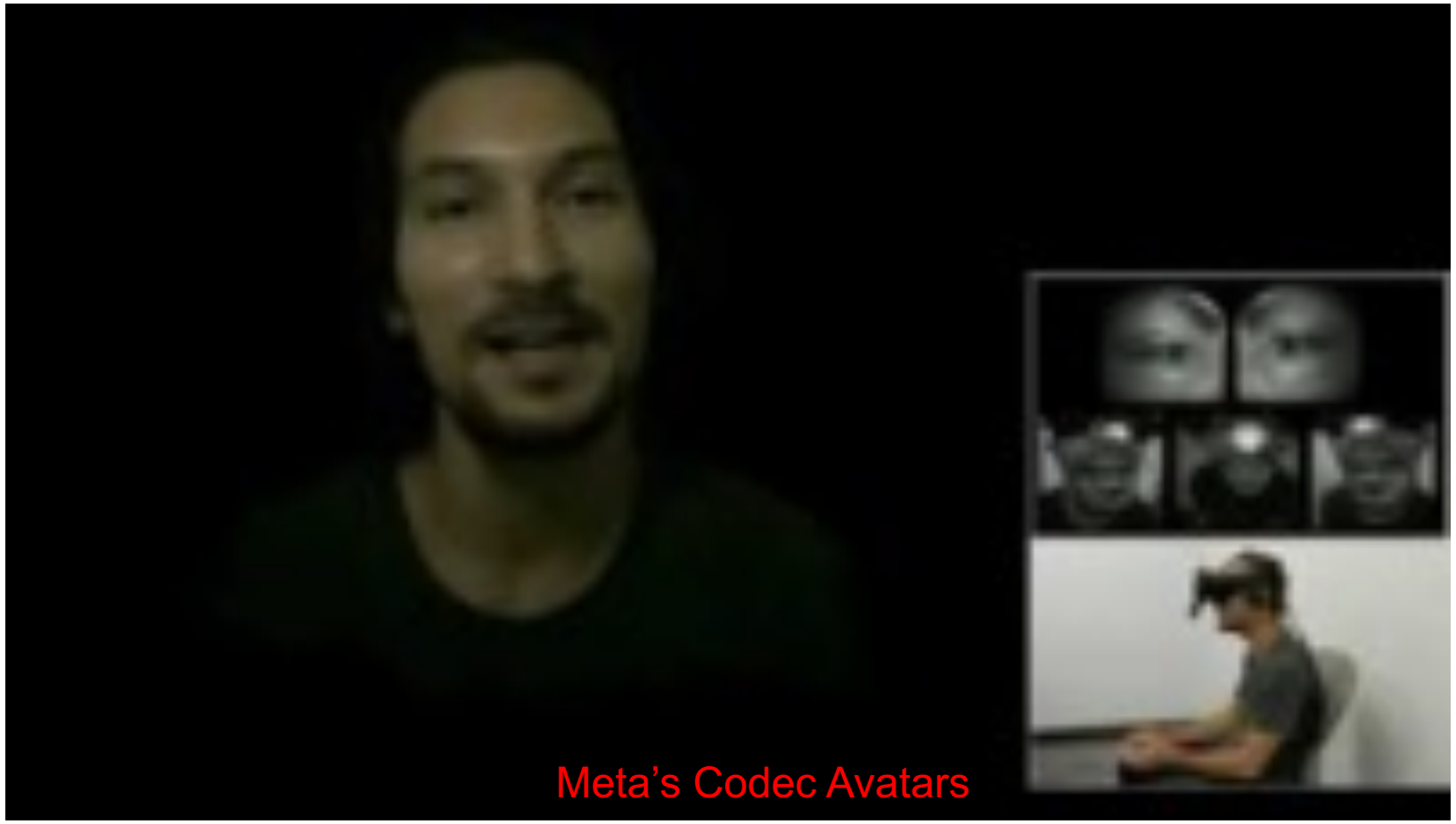
Agriculture



Healthcare

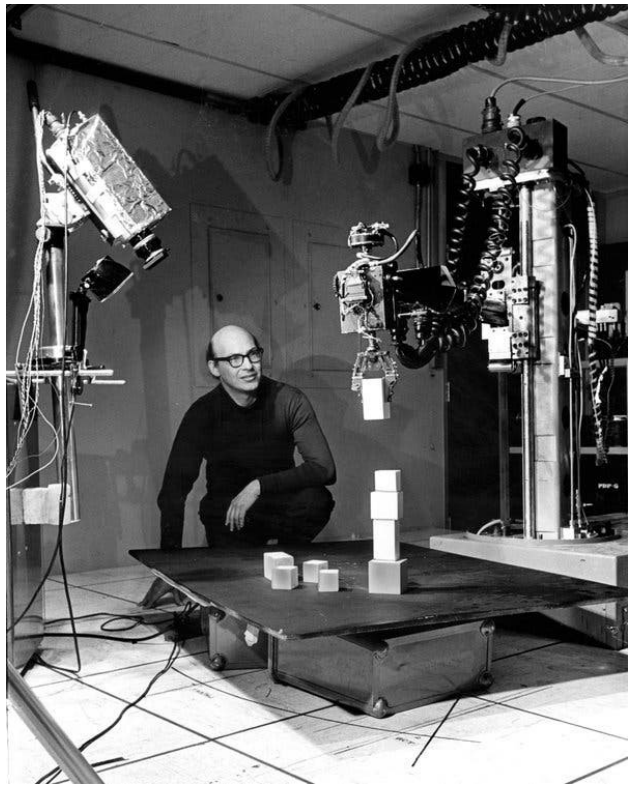
https://www.youtube.com/watch?v=nezqrfAP-g8

Virtual/Augmented Reality



https://www.youtube.com/watch?feature=oembed&v=w52CziLgnAc

The Origin of Computer Vision

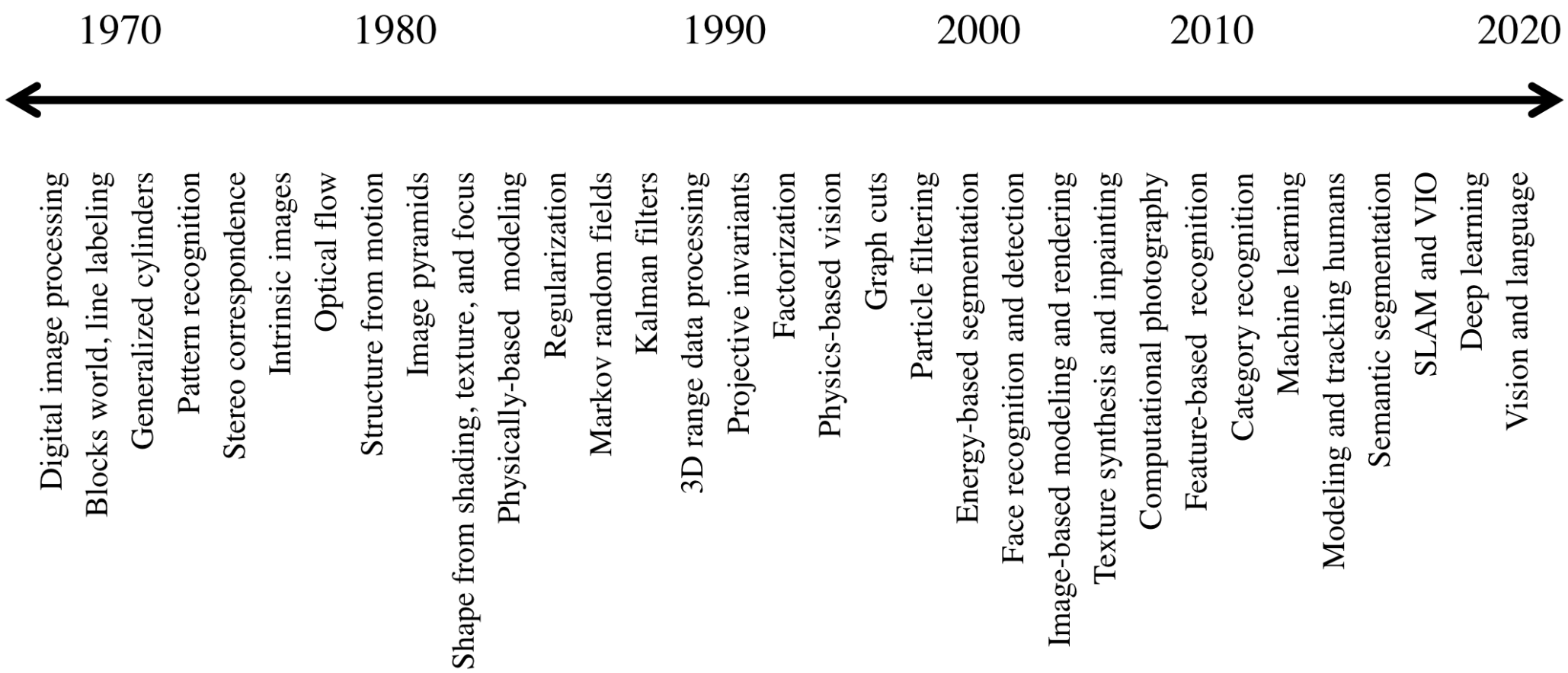


Marvin Minsky in a lab at M.I.T. in 1968.

- 1970s: Viewed as the visual perception component of an ambitious agenda to mimic human intelligence and to endow robots with intelligent behavior.
- At the time, it was believed by some of the early pioneers of AI and robotics that solving the "visual input" problem would be an *easy step* along the path.
- An undergraduate project assigned by Marvin Minsky in 1966: "*spend the summer linking a camera to a computer and getting the computer to describe what it saw*."

Computer vision is far more complex than they thought!

Timeline of Active Topics in Computer Vision



An active research area, rapidly changing, big progress made in recent 10 years!

Why is computer vision difficult?

Viewpoint variation

Illumination

Scale

Slide Credit: Noah Snavely

Why is computer vision difficult?

Intra-class variation

Motion (Source: S. Lazebnik)

Background clutter

Occlusion

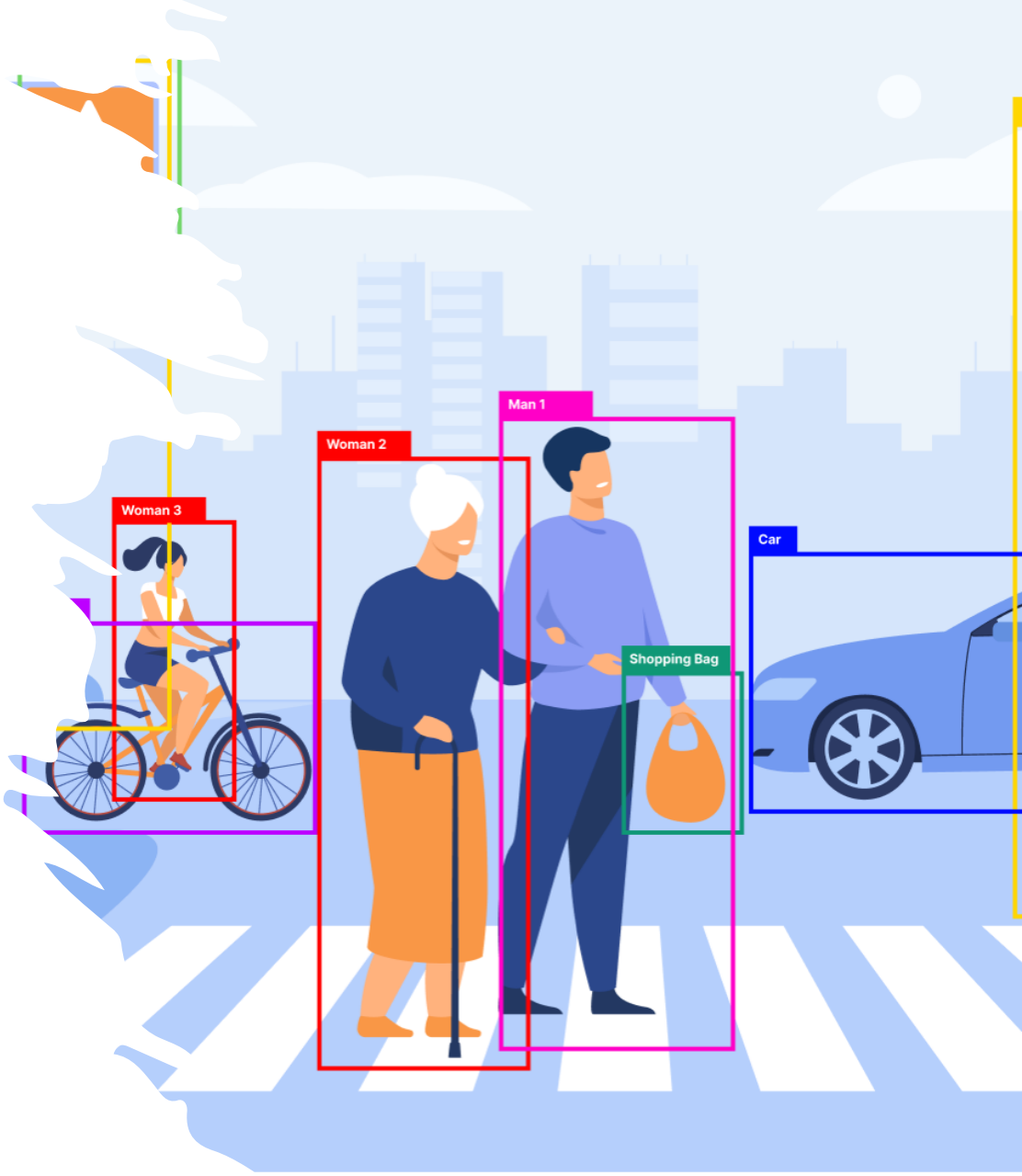
Slide Credit: Noah Snavely

Challenges: local ambiguity

Slide Credit: Fei-Fei, Fergus & Torralba

What will you learn?

- Image Formulation and Processing
- Feature Detection and Matching
- Deep Learning in Computer Vision
- Visual Recognition
- Visual Motion
- 3D Vision
- Advanced Topics in Computer Vision
  - e.g., NeRF, visual representation learning, vision and language, vision and audio



Course Details

Richard Szeliski. Computer Vision: Algorithms and Applications. Second Edition, 2022.
Available online: https://szeliski.org/Book/

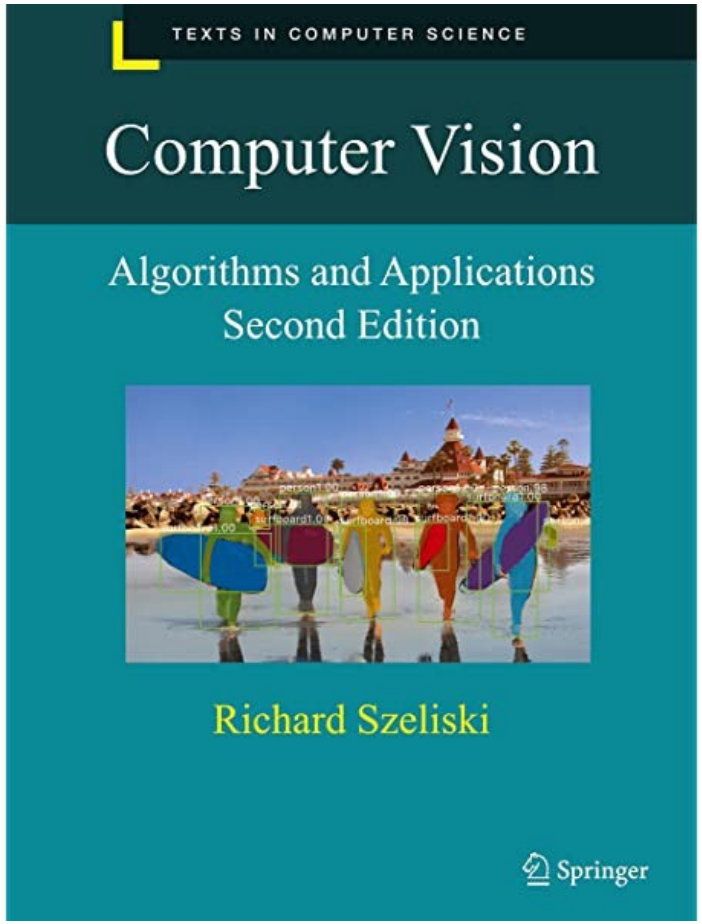
My office hour:
Monday 2pm - 3pm, ECSS 4.211 or By Appointment

TA office hour: TBD

Course website: www.yapengtian.com/t/4391F23

Course access and navigation: eLearning

Course Pre-requisites: *Linear Algebra* and *Python Programming*



Grading Policy

Homework (30%)

- 5 homework in total
- Submit all homework assignments on time. Collaboration is allowed but final work is done independently, and all collaborators should be acknowledged.
- Individual submission

Midterm (30%)

Team Project (30%):

- Develop and implement a method to solve a vision-related problem
- Maximum 4 students for a project
- Project proposal (5%)
- Project presentation (10%)
- Project final report (15%)

In-class Activity (10%)

- 10 quizzes

Late Submission

- For the assignments (not including your final project report), students will be allowed a total of **five** late days per semester
  - No additional late days will be given

- After you use up the free late days, your late submissions will be penalized as follows:
  - Assignments turned in within 24 hours of the due date will receive 90% of its score.
  - Assignments turned in within 48 hours of the due date will receive 70% of its score.
  - Assignments more than 48 hours late will not be accepted.

Questions?