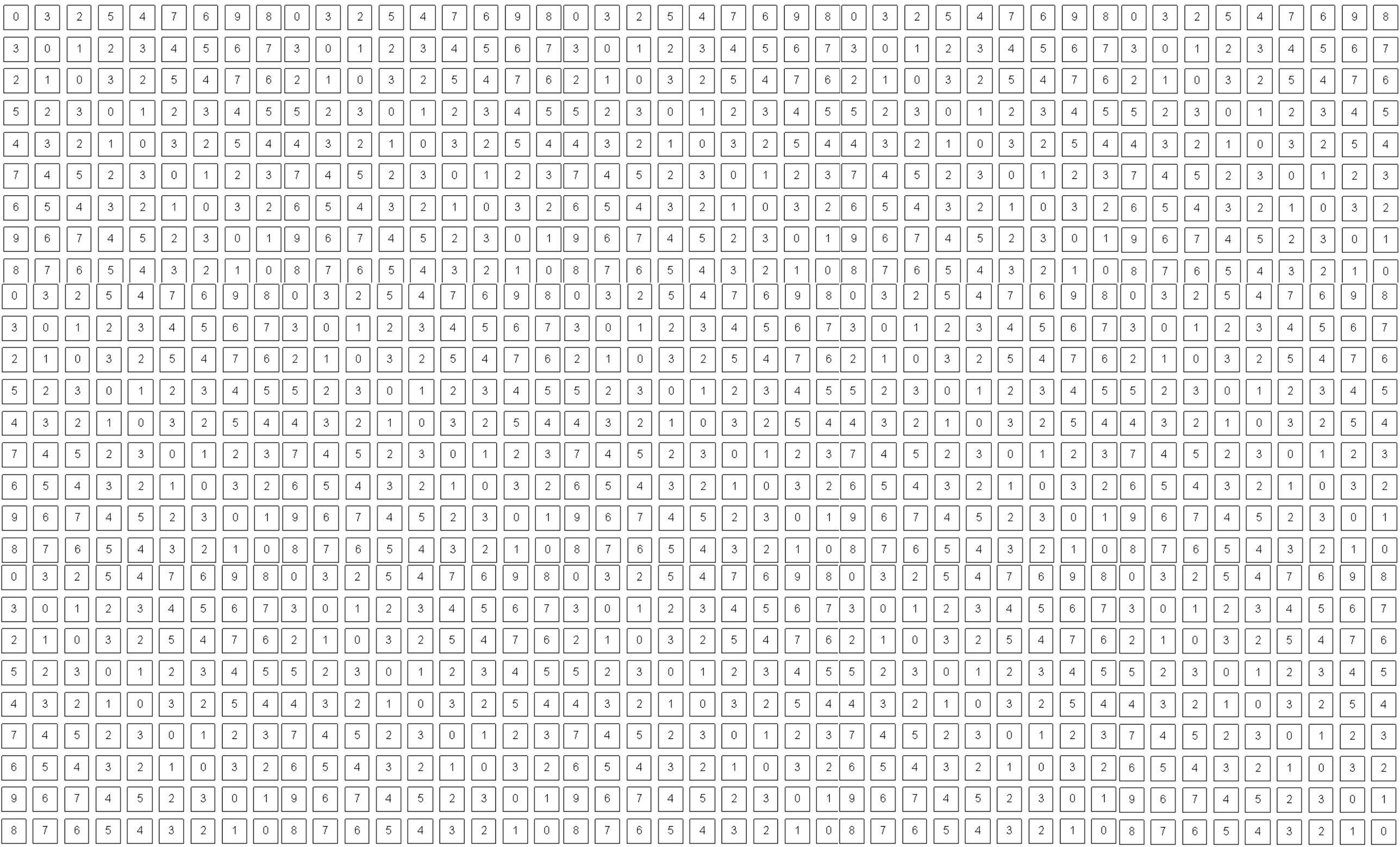
Introduction

Slides Credit:
Adithya Pediredla (Dartmouth), Joseph Redmon (UW), Matt O'Toole/Ioannis Gkioulekas (CMU)

ICS483 Computer Vision
Fall 2025, Lecture 1

What is computer vision?

What a person sees



What a computer sees

Why are we able to interpret this image?

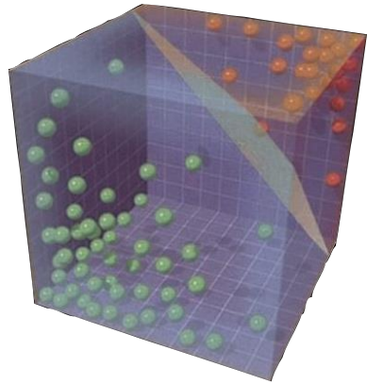
The goal of computer vision is

to give computers

(super) human-level perception

typical perception pipeline

representation



‘fancy math’

output

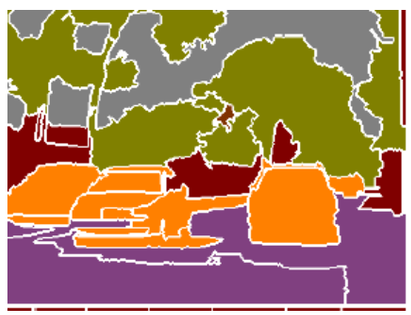
typical perception pipeline

representation

‘fancy math’

what should we look at?
(image features)

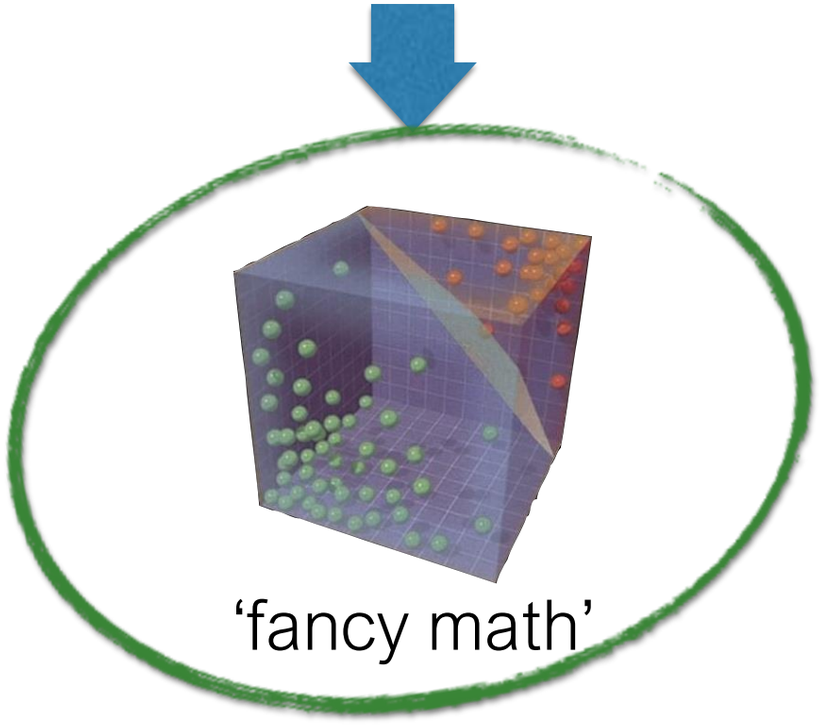
output



what can we understand?
(semantic segmentation)

typical perception pipeline

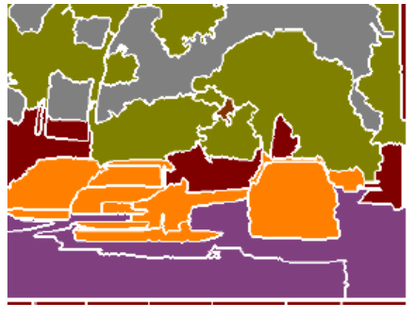
representation



easy to get lost in the techniques

output

what should we look at?
(image features)

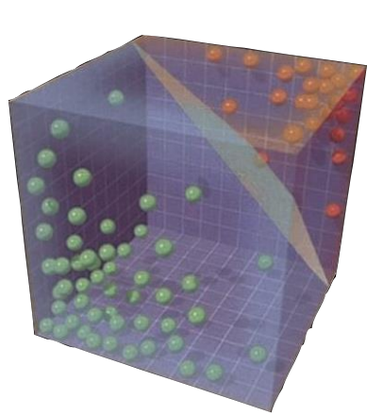


what can we understand?
(semantic segmentation)

typical perception pipeline

representation

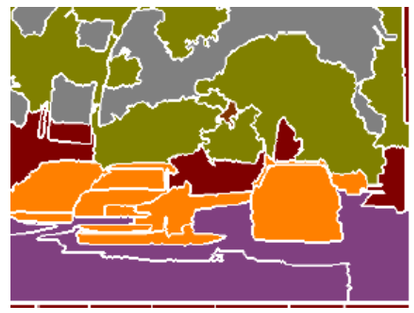
The parts that we are most interested in



‘fancy math’

output

what should we look at?
(image features)



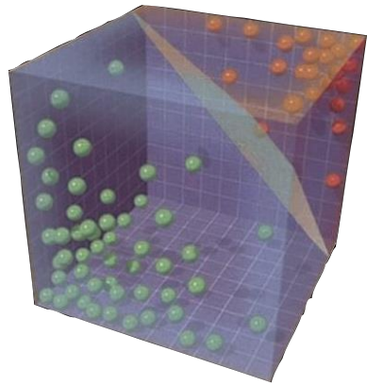
what can we understand?
(semantic segmentation)

typical perception pipeline

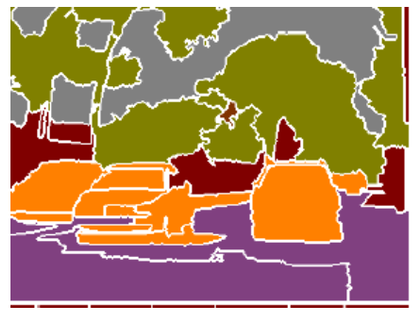
representation

what should we look at?
(image features)

The parts that we are most interested in



‘fancy math’



what can we understand?
(semantic segmentation)

output

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

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".

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

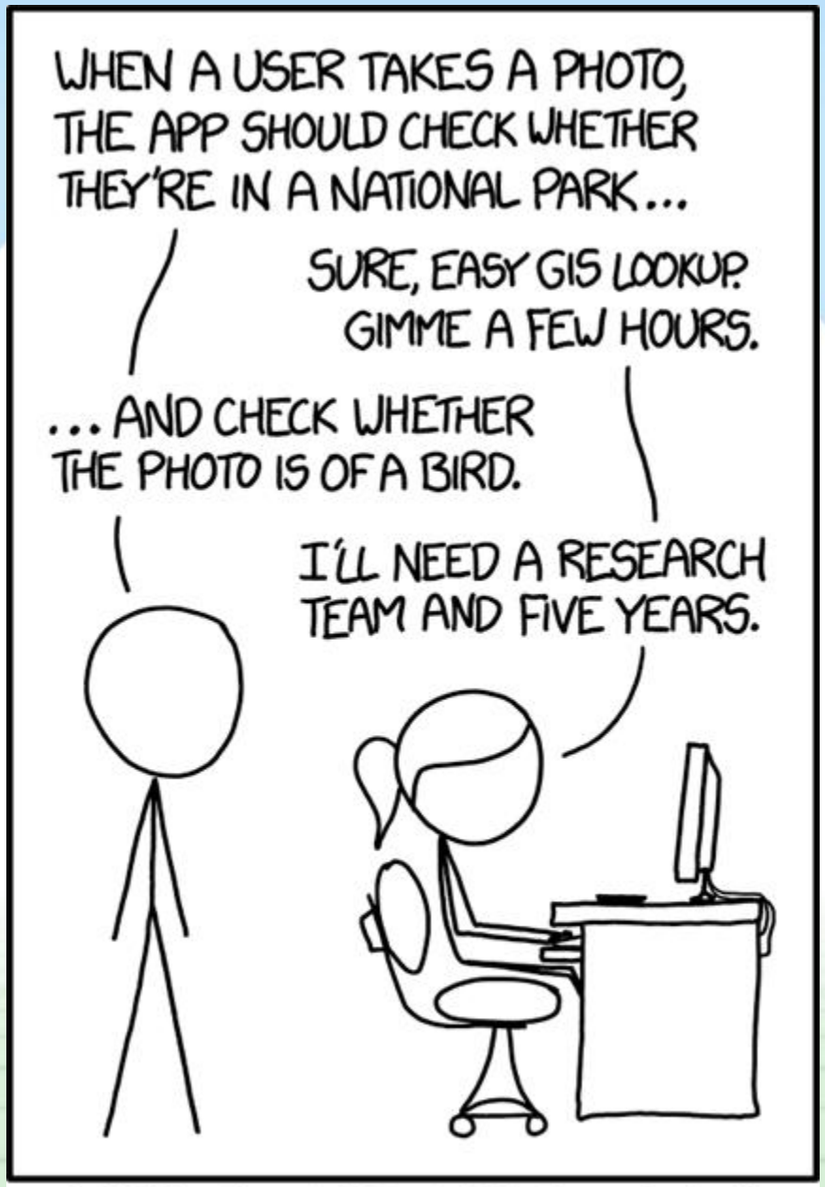
Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

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".



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

Computer Vision

What does WALL-E see?

Low-Level: Resizing

Low-Level: Resizing

Low-Level: Image Adjustments

Low-Level: Grayscale

Low-Level: Exposure

Low-Level: Saturation

Low-Level: Hue

Low-Level: Edges

Low-Level: Oriented Gradients

Low-Level: Oriented Gradients

Low-Level: Segmentation (color)

Low-Level Vision

Photo manipulation
- Size
- Color
- Exposure
- X-Pro II

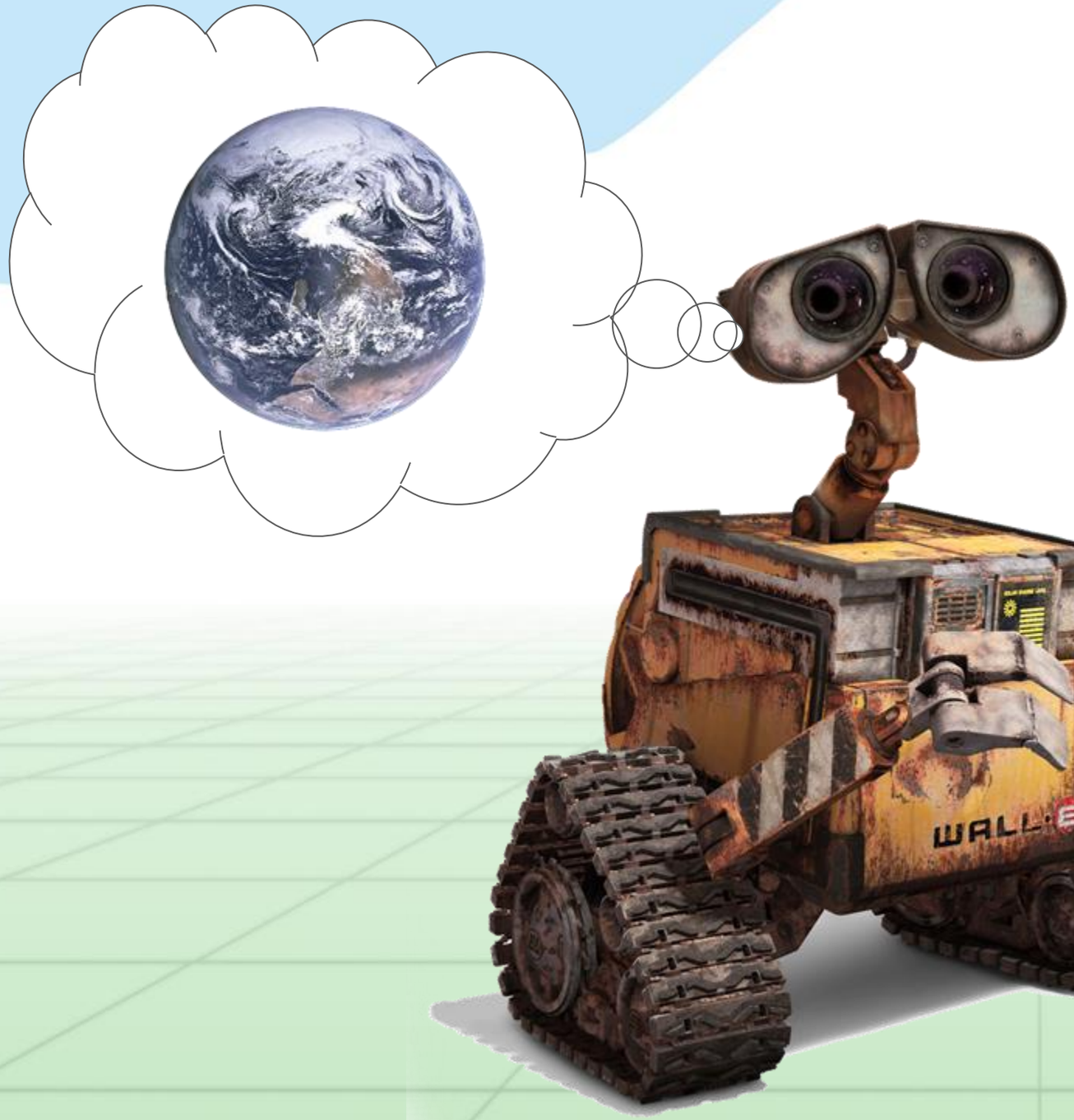
Feature extraction
- Edges
- Oriented gradients
- Segments

Low level vision is exciting!!! #latergram #nofilter

Low-Level Vision Applications?

Anyone?

Mid-Level Vision

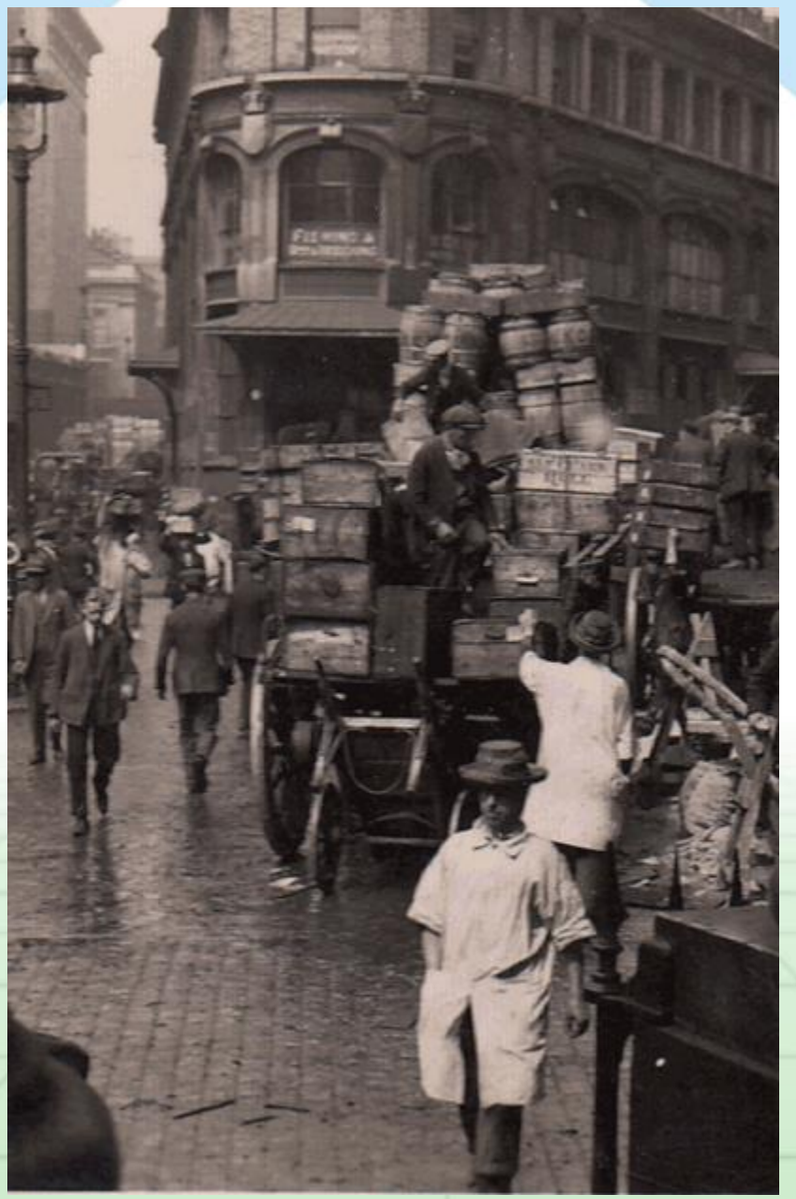


Mid-Level: Panorama Stitching

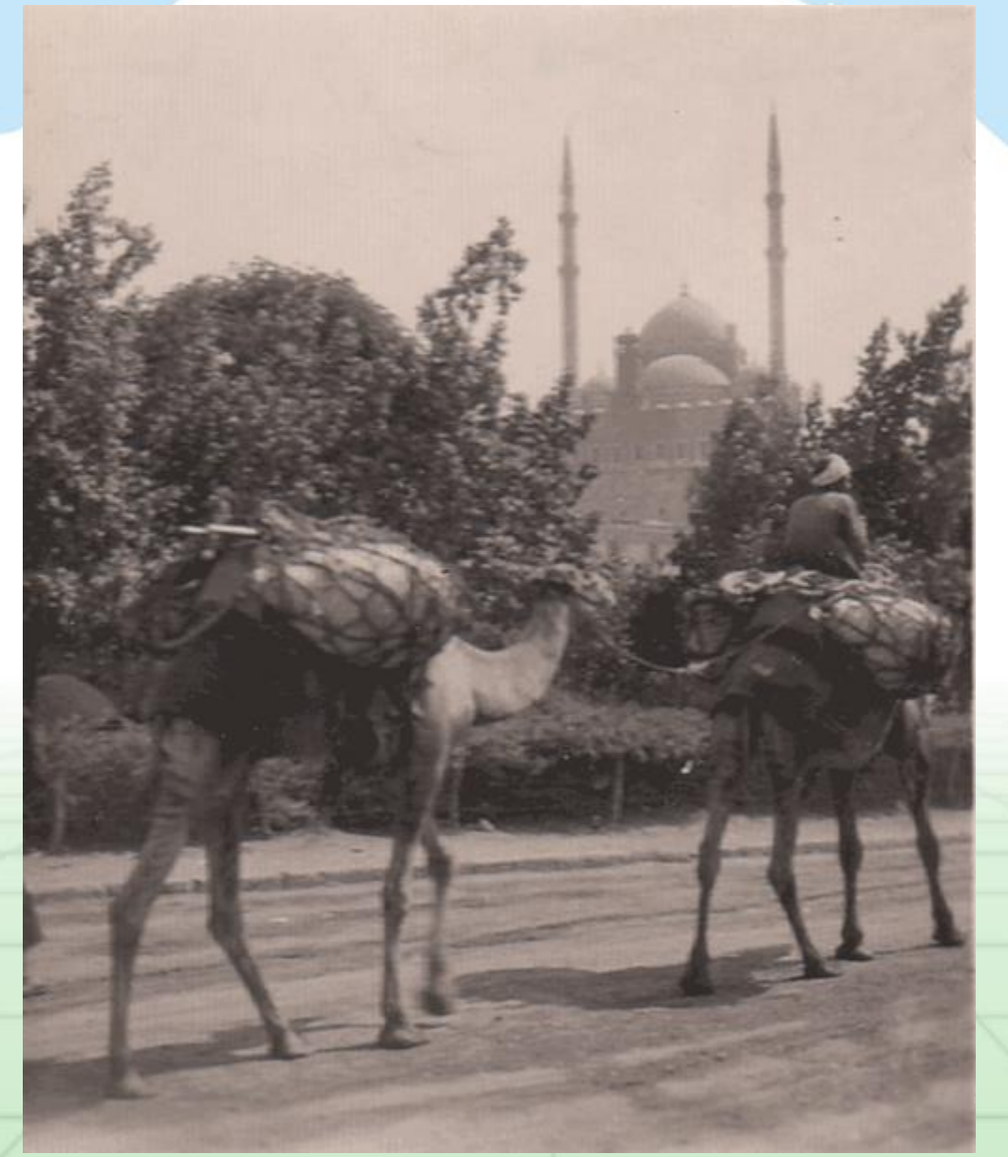
Mid-Level: Panorama Stitching

Mid-Level: Panorama Stitching

Mid-Level: Multi-View Stereo



Mid-Level: Multi-View Stereo



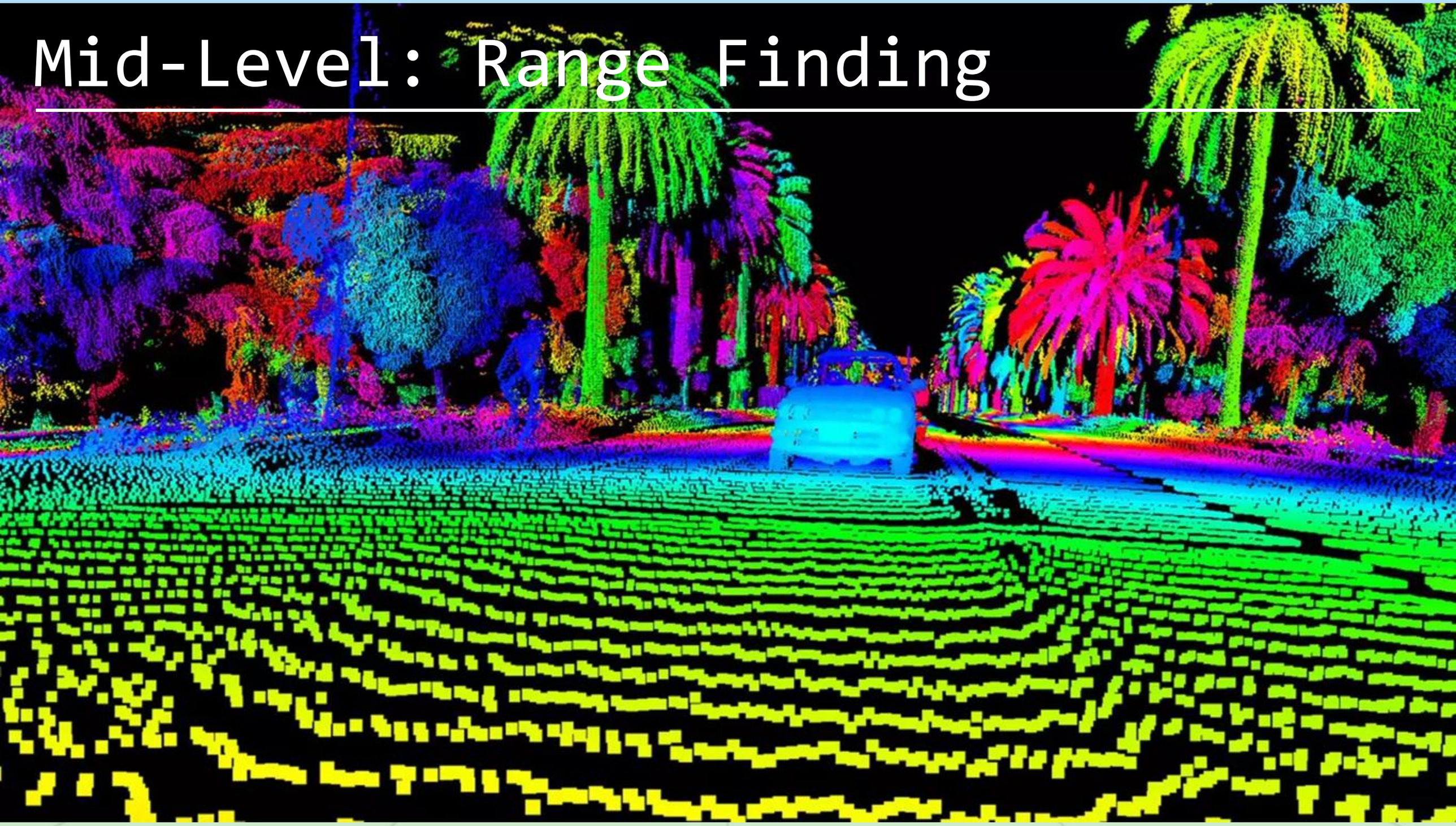
Mid-Level: Multi-View Stereo

Mid-Level: Multi-View Stereo

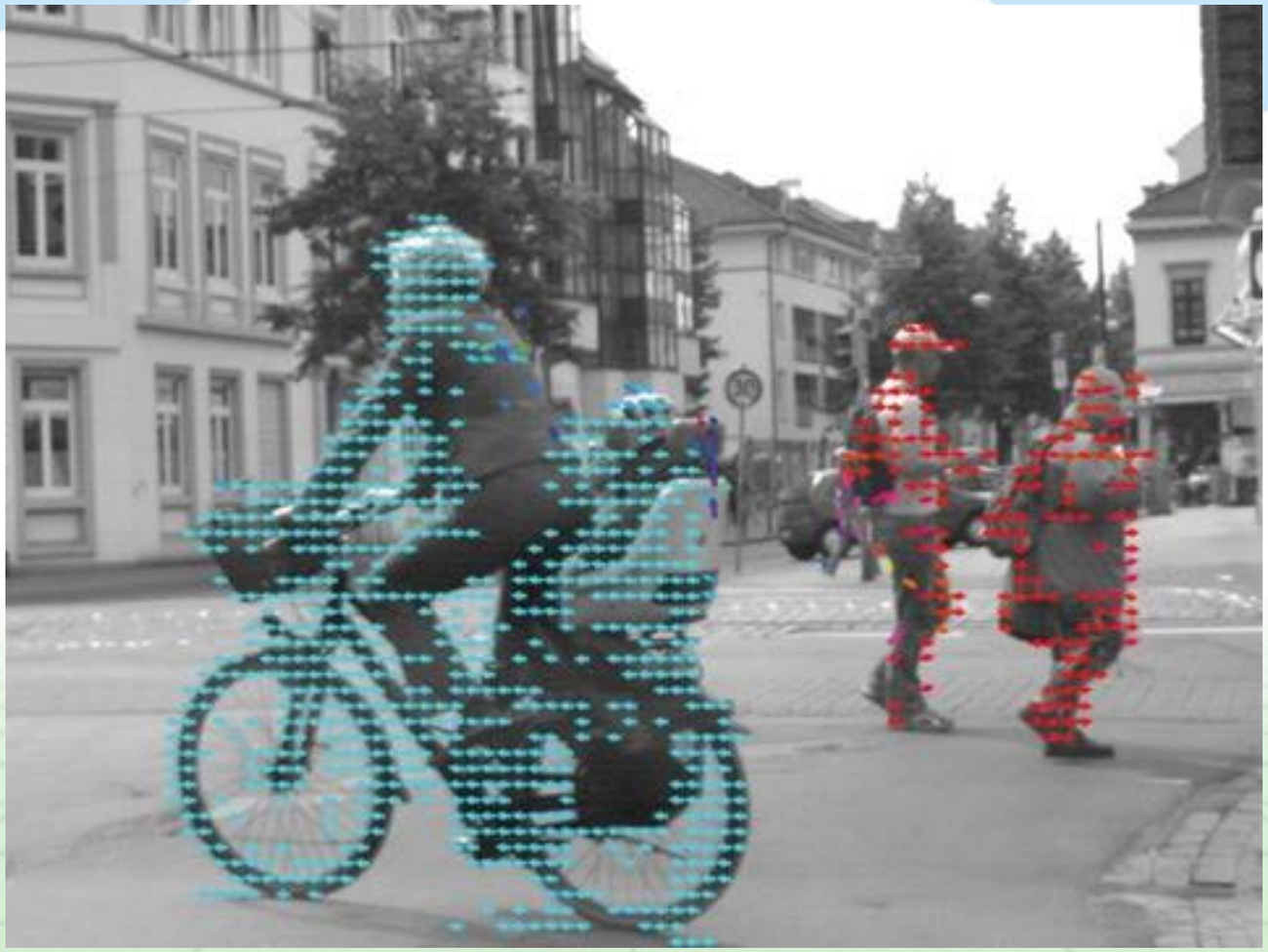
Mid-Level: Structured Light Scan

Mid-Level: Structured Light Scan

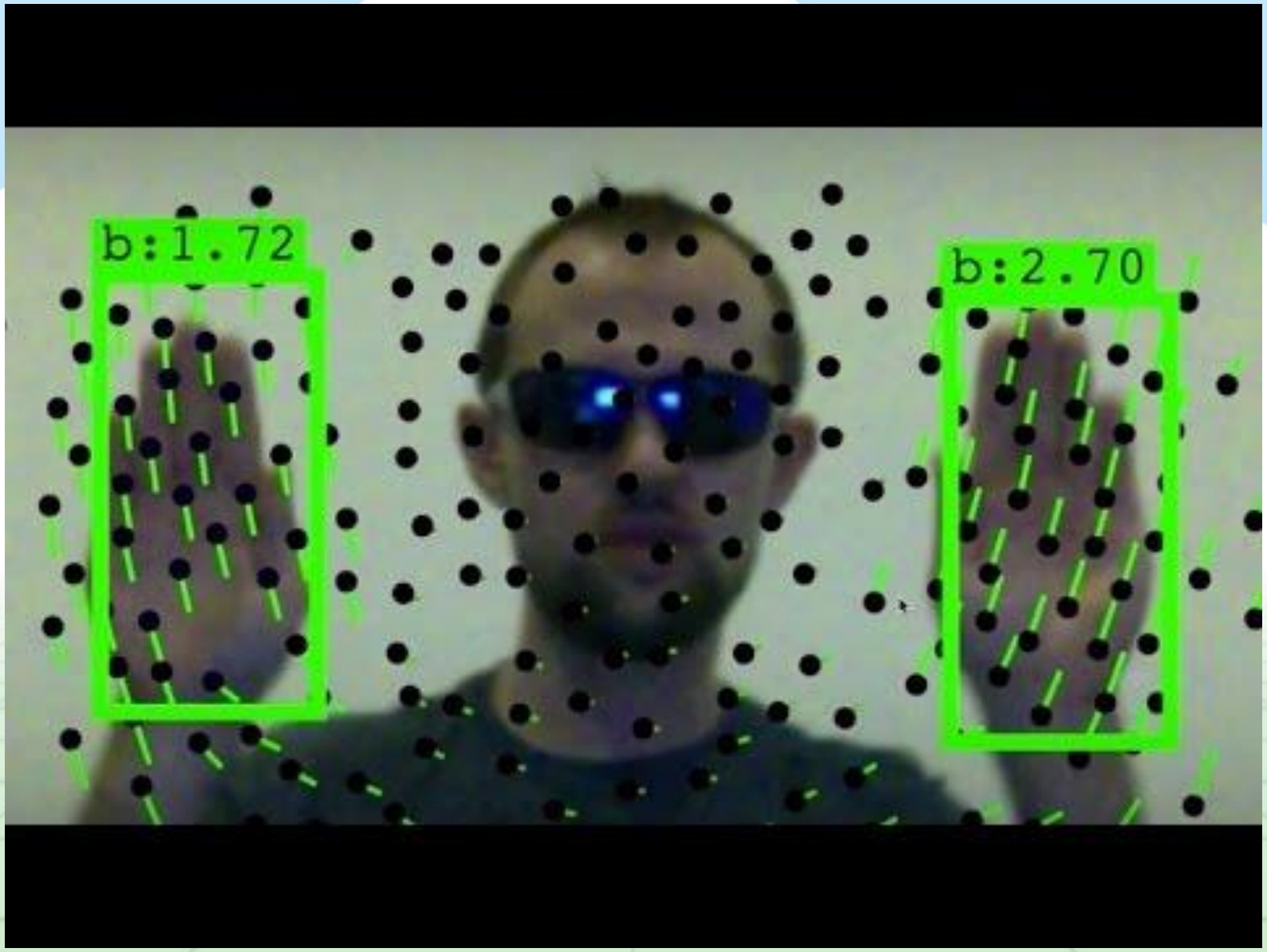
Mid-Level: Range Finding



Mid-Level: Optical Flow



Mid-Level: Optical Flow



Mid-Level: Time Lapse

Mid-Level Vision

Image <-> Image
- Panoramas

Image <-> World
- Multi-view stereo
- Structure from motion
- Structured light
- LIDAR

Image <-> Time
- Optical flow
- Time lapse

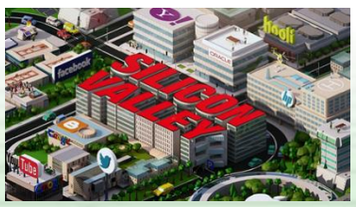
Instructor: Huaijin (George) Chen

Pronunciation: Hawaii-Gin. Hanzi: 陈怀瑾/陳懷瑾

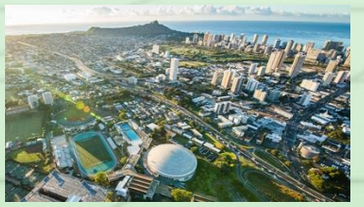
Born in Shanghai, raised in Shenzhen

Rochester Institute of Technology (2009-13)

Rice University (2014-19)

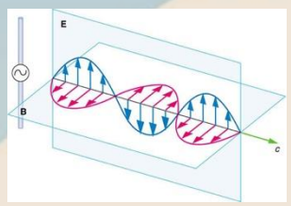


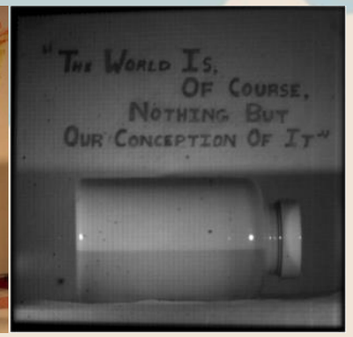
SenseBrain Techonology (2019-22)
Smartphone photography
Vayu Robotics (2022-2024)
Vision-based perception system



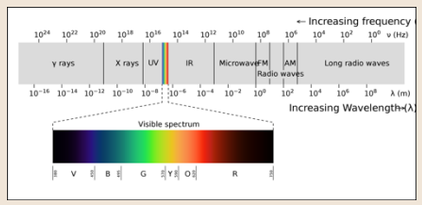
UH Manoa (2024-onwards)
Computational Imaging and Robotic Perception Lab

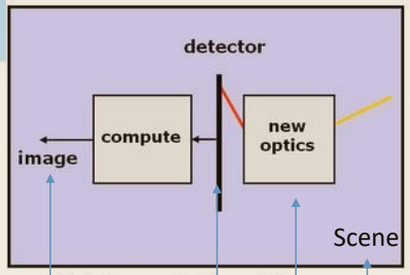
My website: https://hgchen.com/ (will be updated soon)



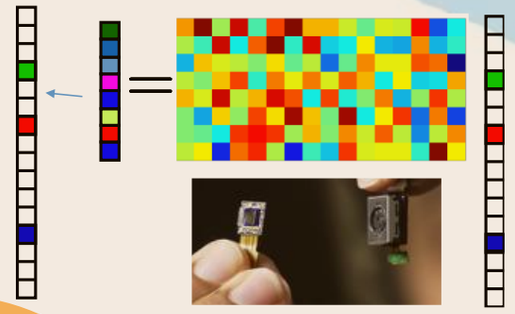


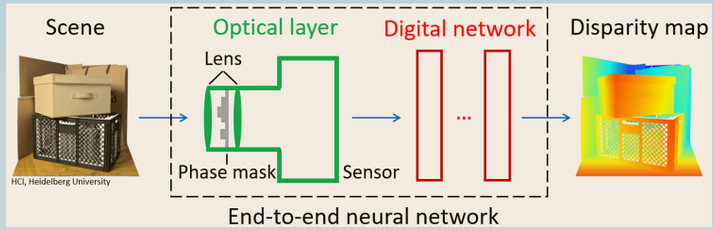
$$L(\theta, \varphi, x, y, z, \lambda, t, \psi)$$

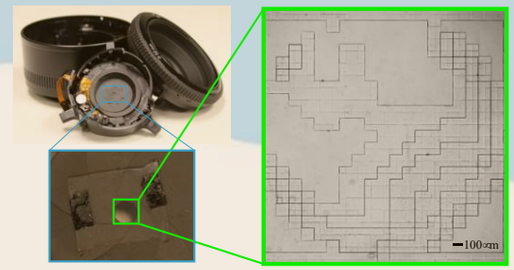




$$\hat{x} \leftarrow \text{Solver} \quad y = Kx$$



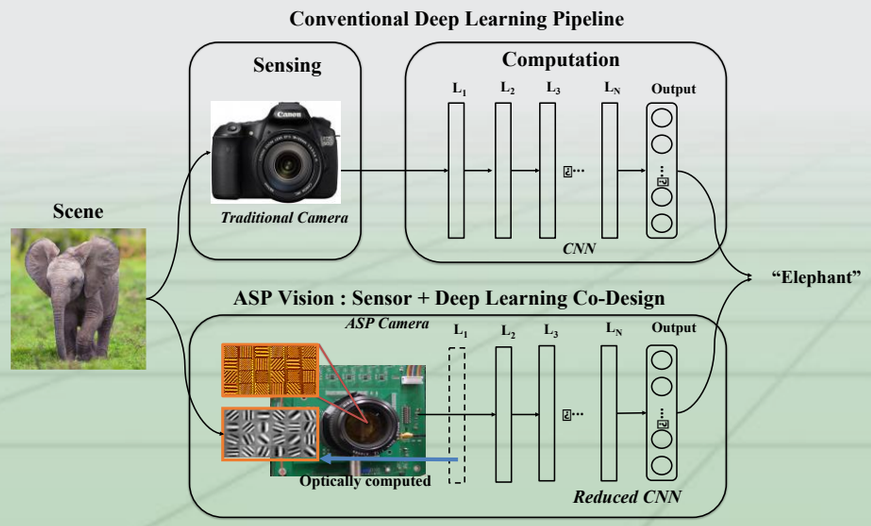




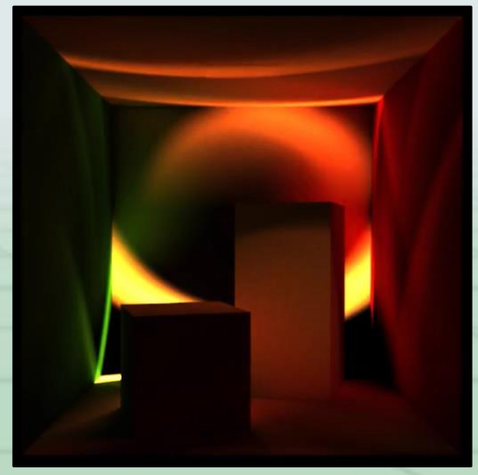
Camera: Super-human sensing & low-cost

System: End-to-end optimized system

Building the foundation for next-gen CV



Computation: Efficient sensing & on-camera optical computation



Model & Data: Physics-based self-supervision and rendering

High-Level Vision

High-Level: Classification

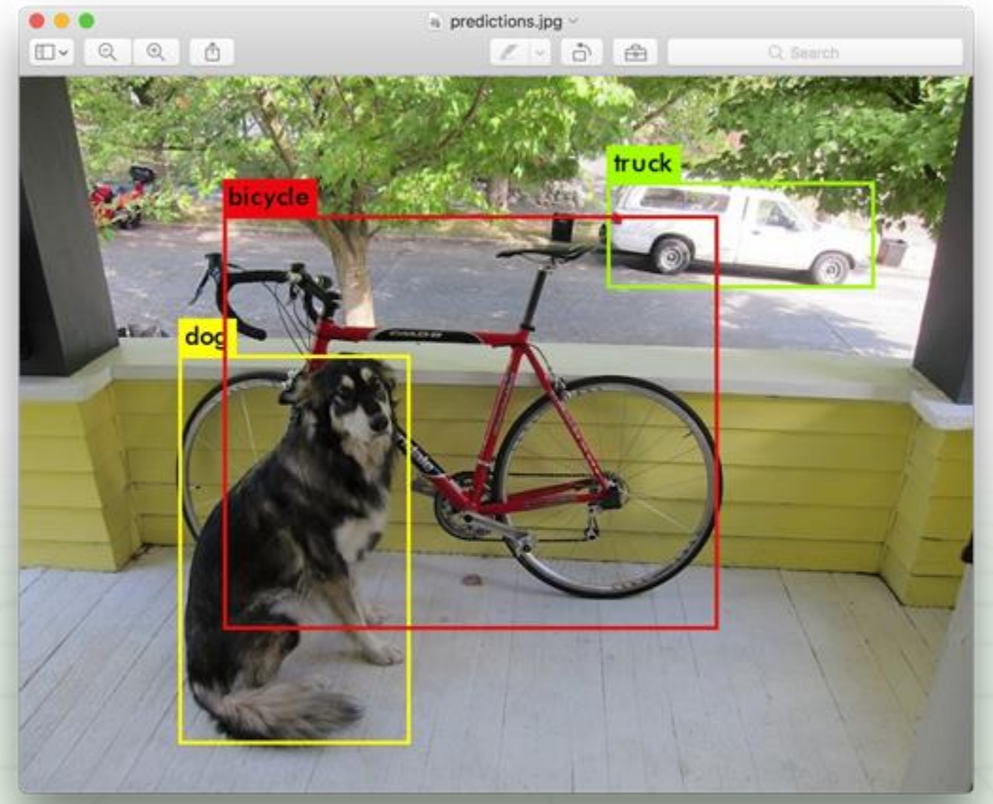
- What is in the image?

High-Level: Tagging

- What are ALL the things in the image?

High-Level: Detection

- What are ALL the things in the image?
- Where are the?



High-Level: Semantic Segmentation

High-Level: Instance Segmentation

High-Level: So many other things

- Single image 3D
- Game playing
- Retrieval
- Other cool things, yay!

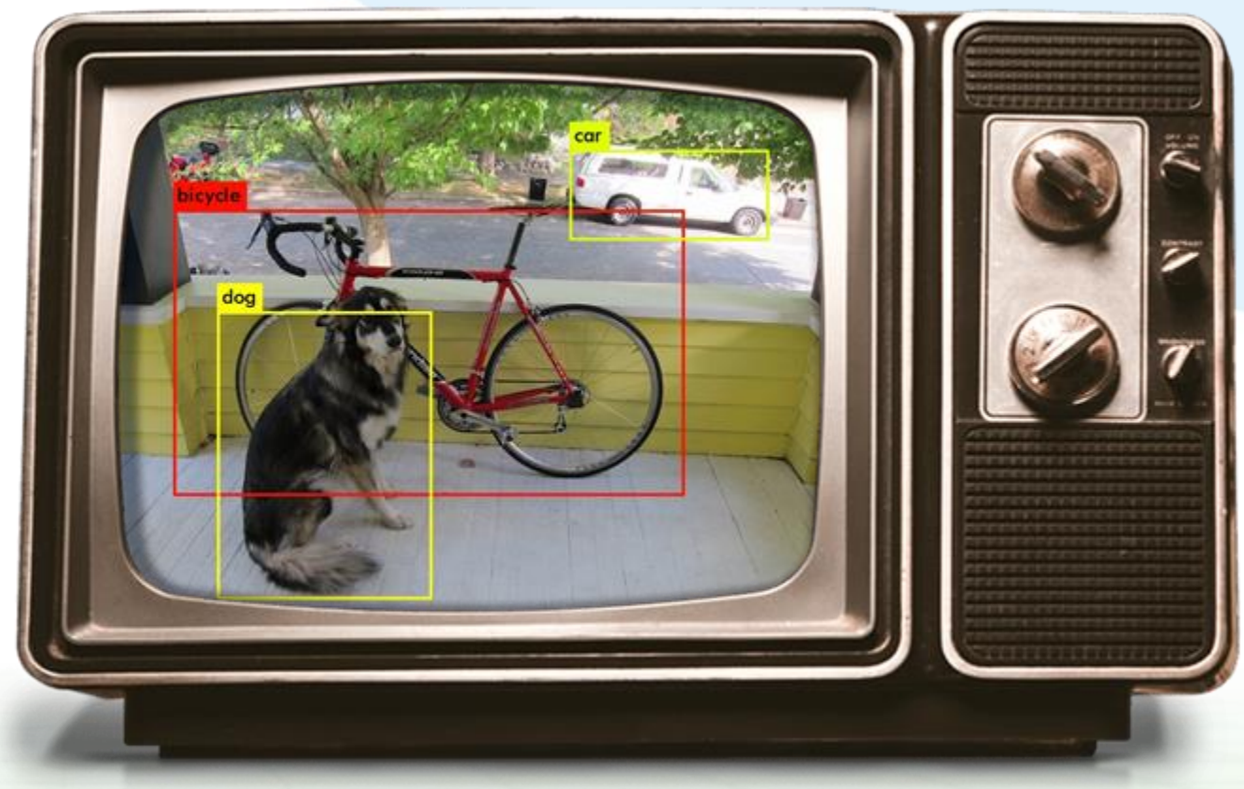
High-Level Vision

Semantics!

- Image classification
- Object detection
- Segmentation

Applications

- Retrieval
- Robots?
- and…????



It’s a good time to do computer vision

Industry aggressively hiring CV faculty from universities


UW


Berkeley


CMU


Toronto

UBER

UCLA


USC


Dropbox

UCSD


Columbia


NYU



CMU


CMU

NVIDIA


Toronto

amazon.com


CalTech


UW

Stanford

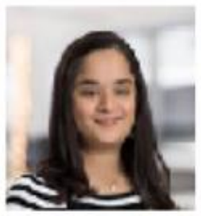

CMU


Stanford


NYU


GTech


GTech


UT Austin


NYU


Google

Zillow

Oculus VR

ARGO AI

SFU


CMU


CMU CMU GTech

CMU


MIT

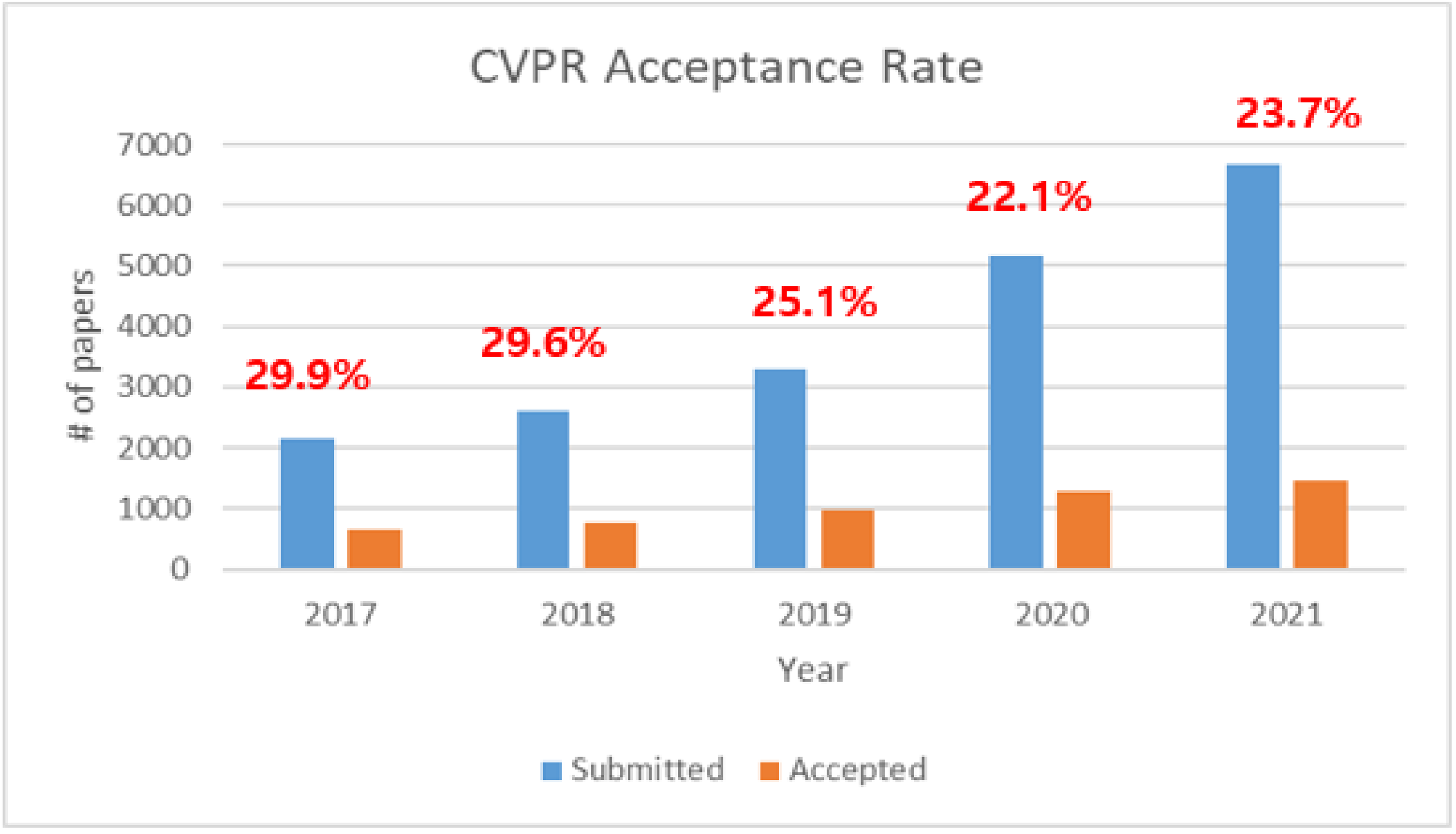

MIT Toronto UW

Industry hiring CV graduates, or even students!

(~~strong~~ dominant industrial presence at conferences for recruitment)

https://www.linkedin.com/jobs/search-results/?geoId=103644278&keywords=computer%20vision

Stats for CVPR (Computer Vision and Pattern Recognition)



Other premier CV venues: ICCV (in Honolulu 2025!), ECCV, ICCP, NeurIPS, SIGGRAPH, IJCV

Course logistics

Announcements & Assignments

Lamaku

Discussions & course materials

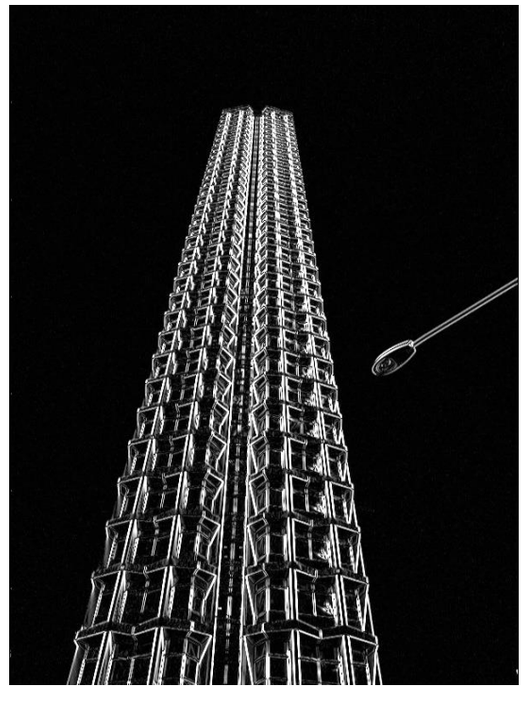
Lamaku? **Discord**?

https://discord.gg/S5Gbr3GCWj

Topics to be covered

Image processing:

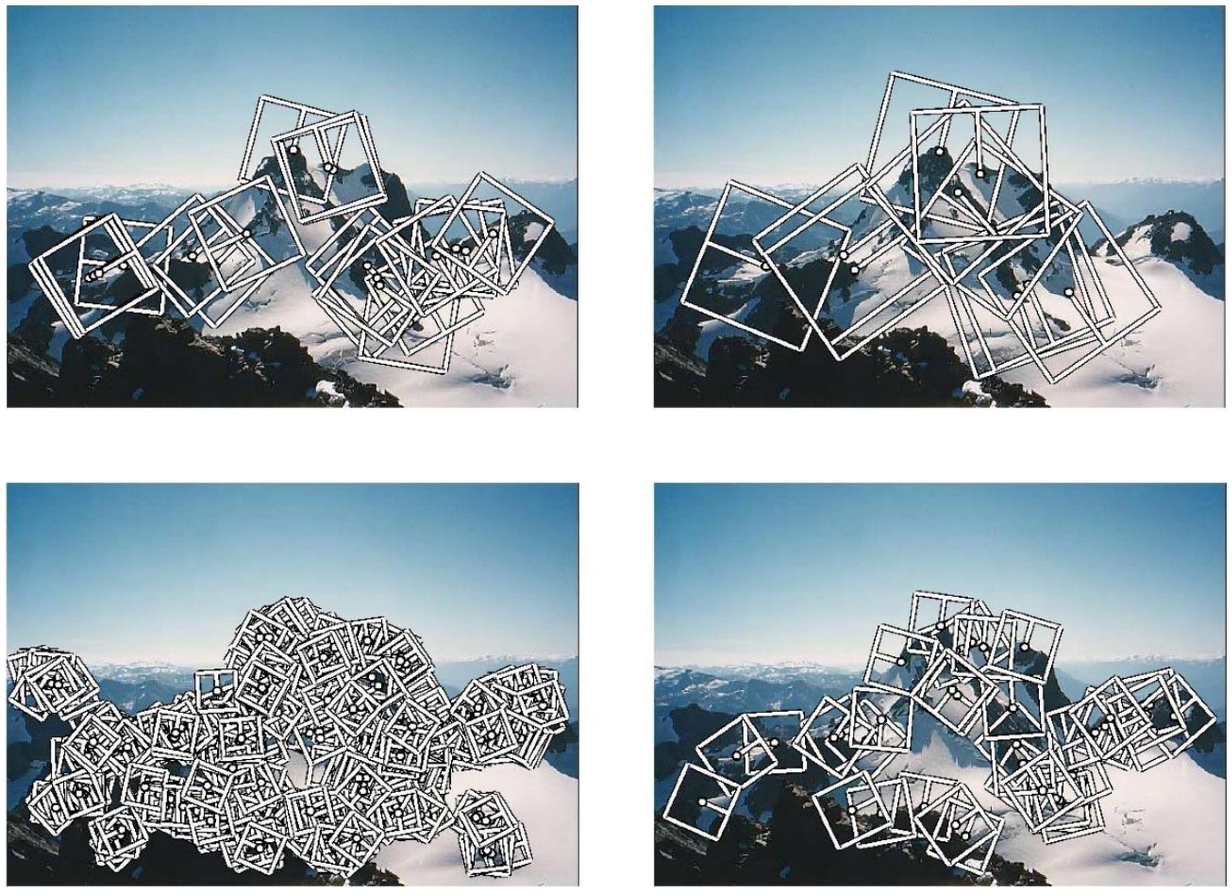
- Basics of filtering.
- Image pyramids.
- Gradients and lines.
- Hough transforms.



Topics to be covered

Feature detection and correspondences:

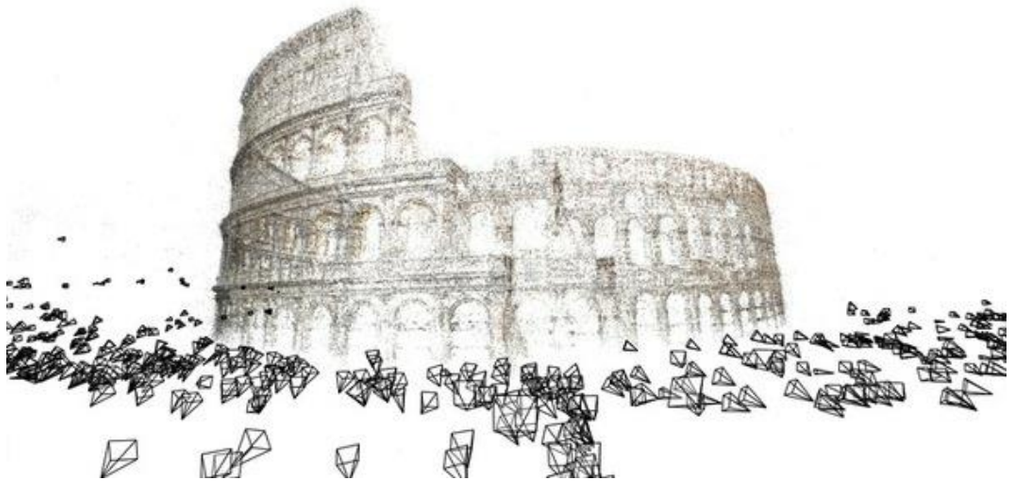
- Corner detection.
- SIFT et al.
- Feature descriptors.
- RANSAC.



Topics to be covered

Transformations and geometry:

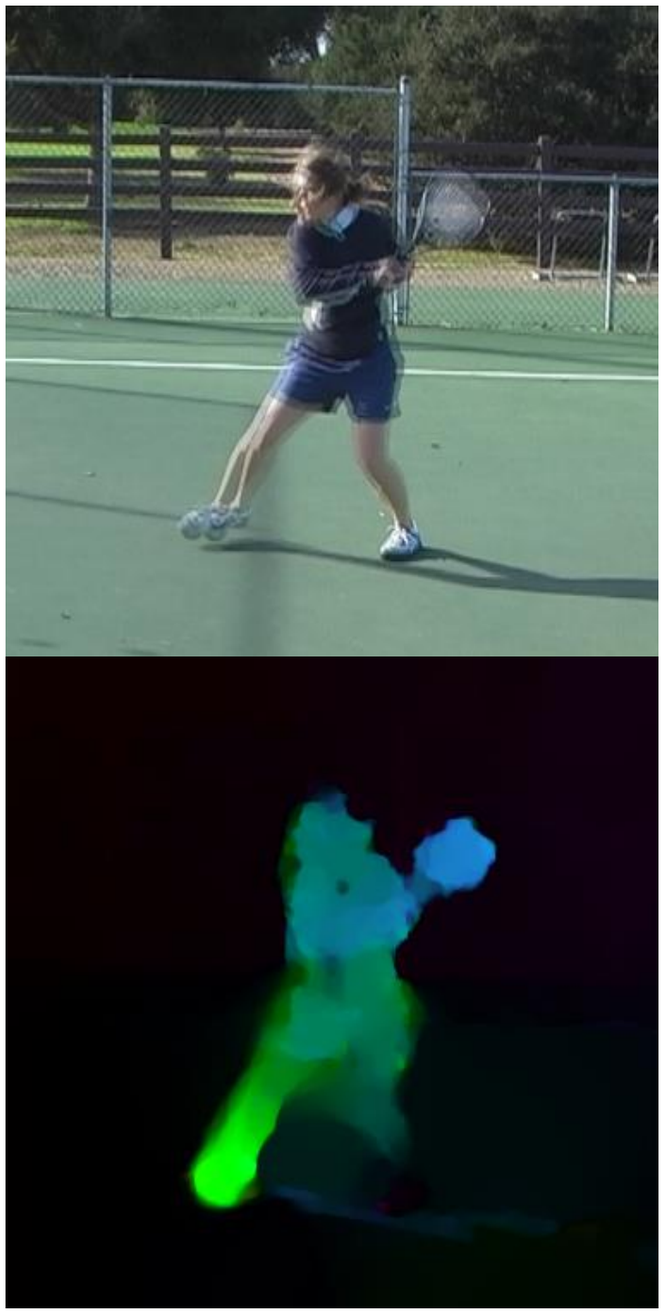
- Homographies and image alignment.
- Camera models.
- Fundamental matrix.
- Epipolar geometry and stereo.
- Structure from motion.



Topics to be covered

Dealing with motion:

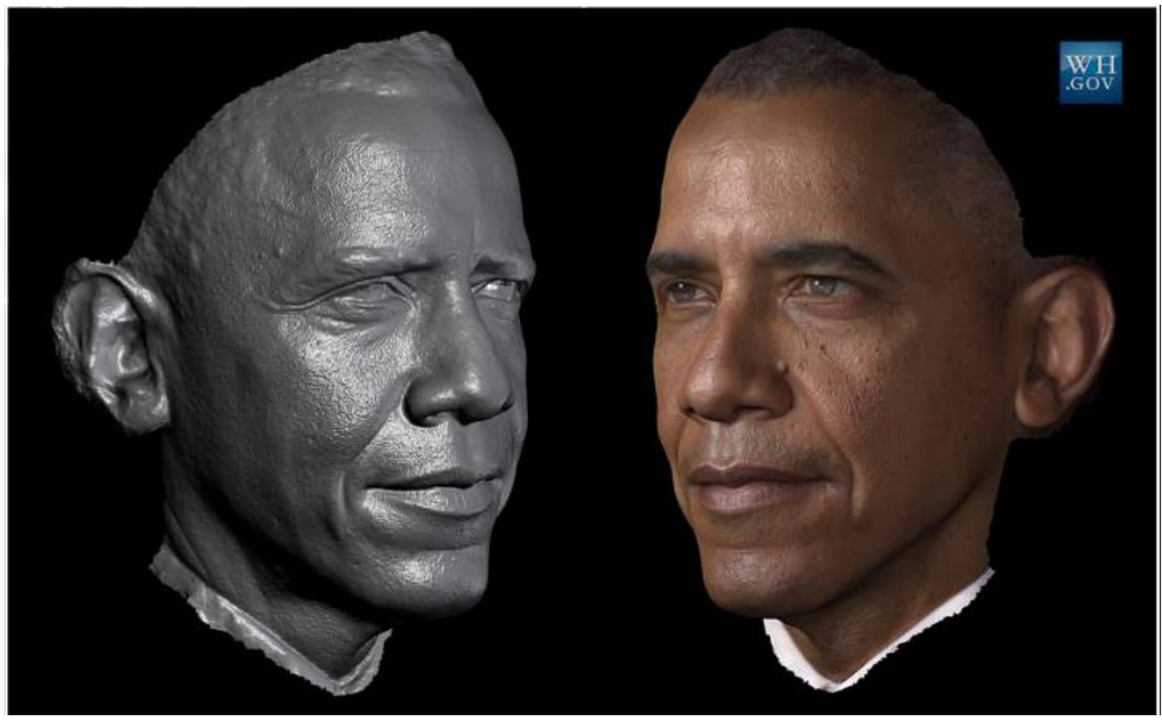
- Optical flow (LK, HS).
- Image registration.
- Kalman Filtering.
- Tracking (KLT, Mean-Shift).



Topics to be covered

Physics-based vision:

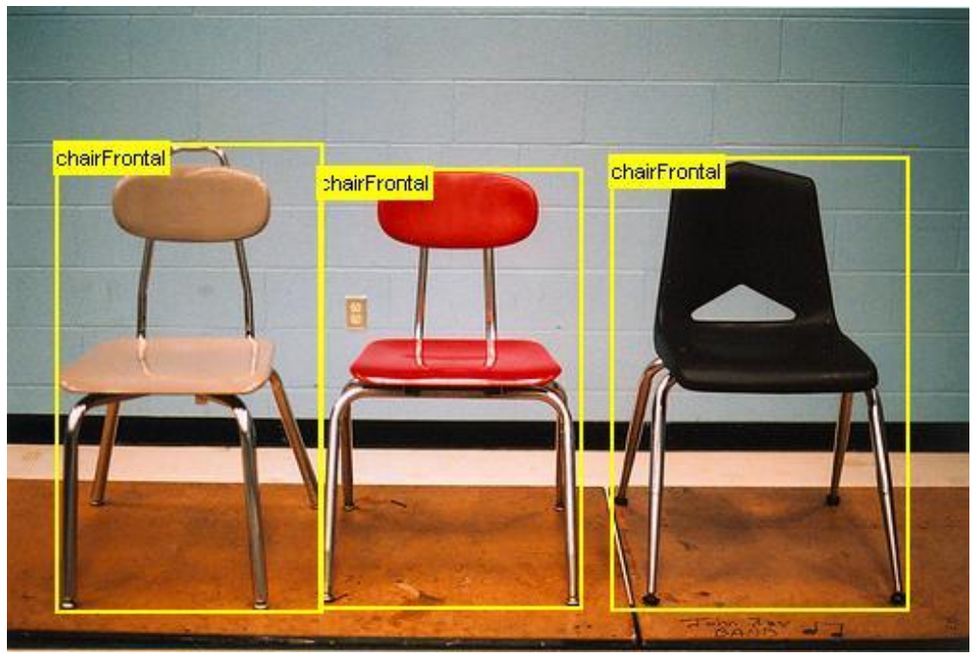
- Reflectance and image formation.
- Radiometry.
- Shape from shading.
- Photometric stereo.
- Computational photography



Topics to be covered

Objects, faces, and learning:

- Basics of probability.
- K-means, KNN, PCA, SVM.
- Bag of words.
- Viola-Jones face detection.
- Perceptron, backpropagation.
- Convolutional neural networks.



Other Good Resources

- Noah Snavely's CS5670 - Introduction to Computer Vision class at Cornell Tech (Spring 2023)
- Bill Freeman, Antonio Torralba, and Phillip Isola's 6.8300/6.8301: Advances in Computer Vision class at MIT (Spring 2023)
- Yasutaka Furukawa's CMPT 412 - Computer Vision class at Simon Fraser University (Spring 2023)
- David Fouhey's EECS 442: Computer Vision class at the University of Michigan (Winter 2023)
- Alyosha Efros' CS194-26/294-26: Intro to Computer Vision and Computational Photography class at Berkeley (Fall 2022)
- James Hays' CS 4476-A / 6476-A Computer Vision class at Georgia Tech (Fall 2022)
- James Tompkin's CSCI 1430 Computer Vision class at Brown (Spring 2023)
- Ioannis Gkioulekas's 15-463, 15-663, 15-862 Computational Photography class at CMU (Fall 2023)
- Matthew O'Toole's 16-385 Computer Vision class at CMU (Fall 2022)
- Justin Johnson's EECS 498.008 / 598.008: Deep Learning for Computer Vision class at the University of Michigan (Winter 2022), which is an outstanding introduction to deep learning and visual recognition
- Yann LeCun and Alfredo Canziani's DS-GA 1008: Deep Learning class at NYU (Spring 2021)
- Luiz Velho's Fundamentals and Trends in Vision and Image Processing class at IMPA (Spring 2021)
- UC Berkeley's CS294-158-SP20: Deep Unsupervised Learning class (Spring 2020)
- Scott Wehrwein's CSCI 497P/597P - Introduction to Computer Vision class at Western Washington University (Spring 2020)
- Andrew Owens' EECS 504: Foundations of Computer Vision class at the University of Michigan (Winter 2020)

List curated by Rick Szeliski

Grading

Items	Points	Remarks
3-5 Programming Assignment	55	Hands-on programming problem sets.
1 Final Project	20	Individual or group project (maximum 3-person group is allowed). Details of the final project requirements will be released soon.
4-8 Take-home Quizzes	25	2-3 math and theory questions.
Attendance & Extra Credit Opportunities	Up to 10	There are extra credit opportunities throughout the semester. Including a) 10-minute paper presentations (5 pts each), b) mini exercises (2 pts), and c) class participation (3 pts max). The total extra pts are capped at 10 pts maximum.
Total	130	Final Grade: 90-100: A; 80-89: B; 70-79: C; 60-69: D; 59 and below: F

Programming assignment topics

Filtering / Hough Transform

Homography

Stereo

Lucas-Kanade Tracking

Deep learning & Generative AI

seriously, a lot of

programming, so start early!

- a lot of programming in Python.
- hours and hours of programming.
- days and days of debugging.
- generous grading policy with extra credits (like grad school)
- score per assignment gets clipped to maximum points

Schedule

- Tentative schedule
- Likely to change.
- Always check Lakuma page
- Course material is subject to be changed

https://docs.google.com/spreadsheets/d/15Zh_J_C1sjxu03i_r-GV_N3lFoYaEH-8OzveNOGPJgA

Late submission

Late days for programming assignments & take-home quizzes:

- 10% reduction of points per late day
- Maximum 2 late days allows per submission
- use only for unforeseen circumstances

No late days for final project

Zero tolerance for cheating

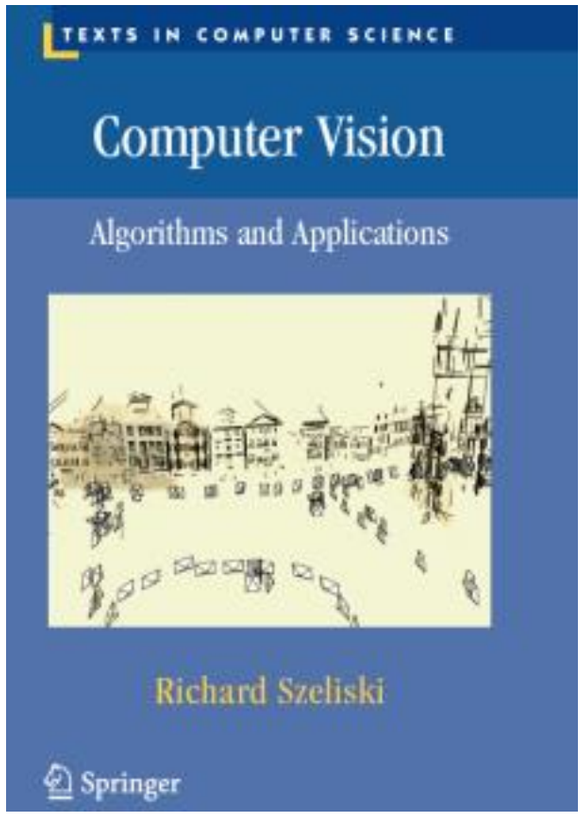
- ICS department policy: Instructors **must** file a report to the VC of Student Affairs as soon as they suspect/detect cheating.

- Discussions with peers **are encouraged**, but you should declare them and you should write your own code/report.

- Large Language models (ChatGPT, etc) LLMs can be used, but **not to write codes or your reports directly**. If you use any LLMs in your assignments/quizzes, please attach a report including the prompts you used and responses you got, as well as your own reflection on the answers with your submission. If time permits, we will have a class session to review the prompts.

Book

We will be posting readings after each lecture



PDF online

http://szeliski.org/Book/

Also check out the “Slide sets and lectures” section

Prerequisites

We assume familiarity with calculus, linear algebra, basic probability, and programming.

Formal prerequisites:

- ICS 212 Program Structure
- ICS 311 Algorithms
- ICS 314 Software Engineering I

Other prerequisites

- Programming in Python
- MATH 311 Introduction to Linear Algebra
  - Will be official pre-requitesite starting next year
  - YouTube: “Essence of linear algebra” by 3Blue1Brown

Contact information and office hours

- Email me about administrative questions only.
    - Please use [ICS 483] in email title.
    - Allow up to one day for reply and email again after 2 days not receiving the reply.
- **Technical questions should be asked on in the discussion tool we choose.**
    - We won't answer technical questions through email.
- Office hours
    - POST 312 TR 115pm-2pm (right after class)

Questions?

Mini Quiz

Mandatory, but worth 1 pts extra points
will serve as your enrollment verification as well

1. Solve a Linear System

Solve for x, y:

$$\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

---

2. Matrix–Vector Multiplication

Compute

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

---

3. Integration

Evaluate the integral:

$$\int_0^2 (3x^2 + 2x)\, dx$$

---

4. Derivative

Find $\dfrac{d}{dx}\left(x^3 - 5x^2 + 2x\right)$.

---

5. Python Programming (Basics)

Write a Python function that:

- Creates a list of 5 random numbers between 0 and 1
- Multiplies each number by 10
- Returns the new list

Hints:

- Use `import random`
- Use `random.random()` to generate a random number in [0,1)
- Use a `for` loop or a list comprehension
