# Introduction


16-385 Computer Vision
Fall 2024, Lecture 1

# Overview of today's lecture

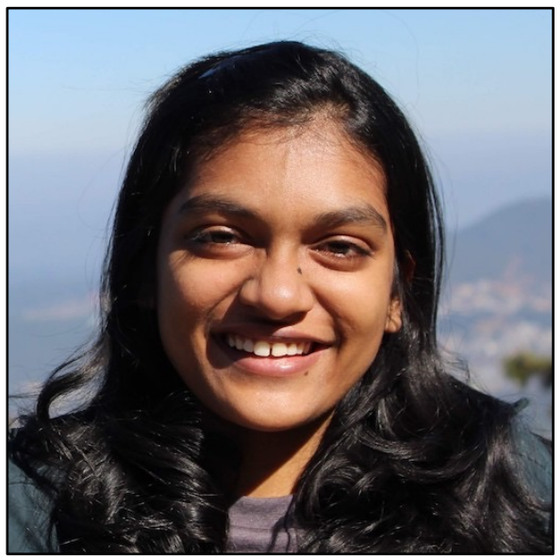
- Teaching staff introductions
- What is computer vision?
- Course fast-forward and logistics

# Teaching staff introductions

# Hi!

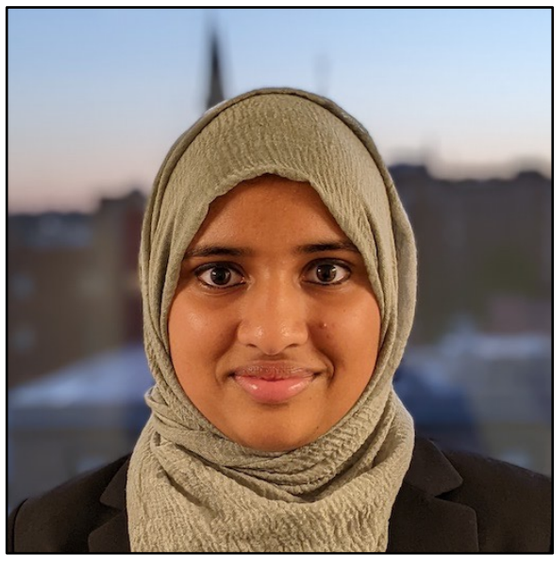
**Matthew O’Toole (Instructor)**


**Jinhyung (David) Park**



**Shubhika Garg**

**Zhinan (Sam) Wang**

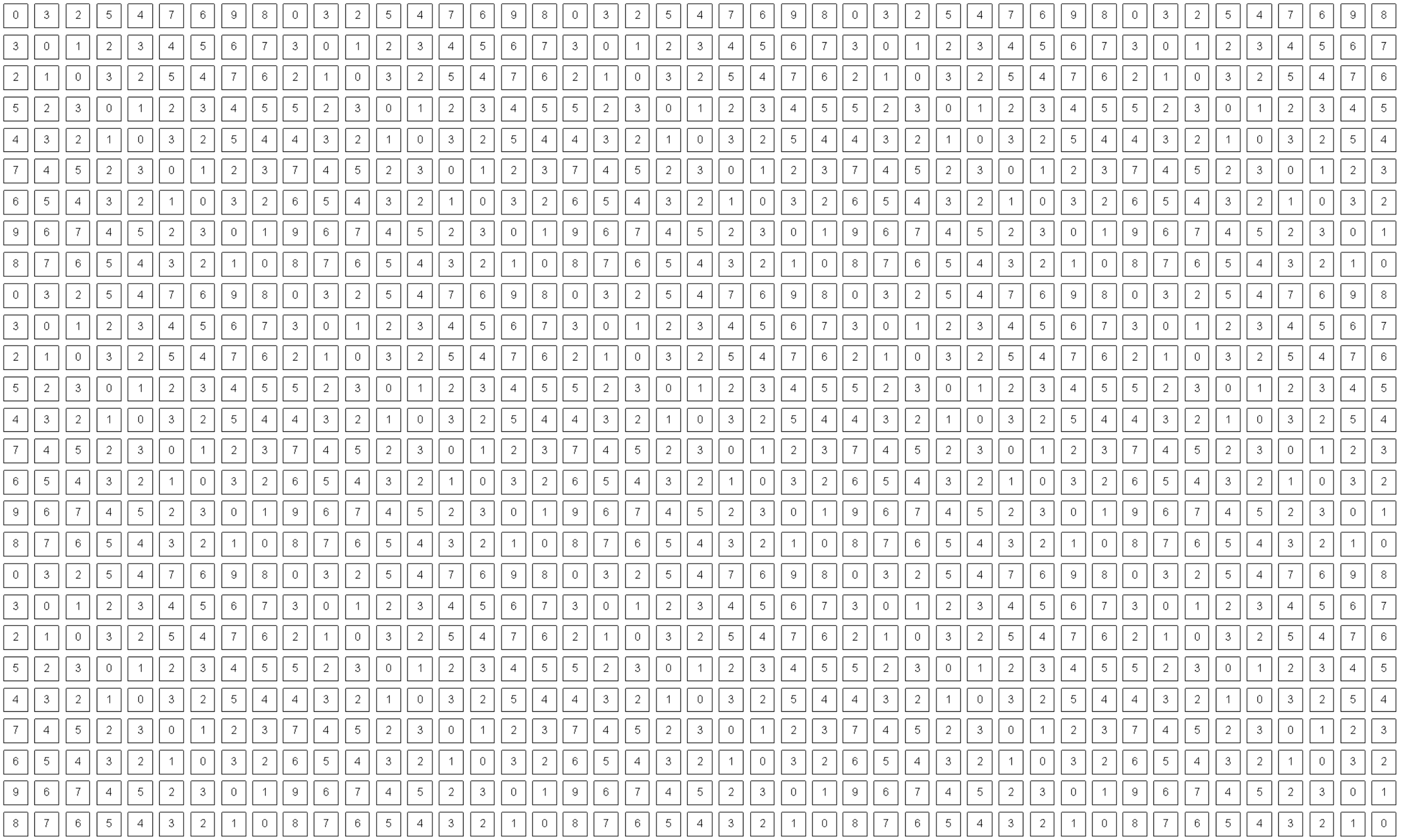


**Tahaseen Shaik**

**Felipe Zanforlin Mautner**

# What is computer vision?


# What a person sees




# What a computer sees


# Why are we able to interpret this image?

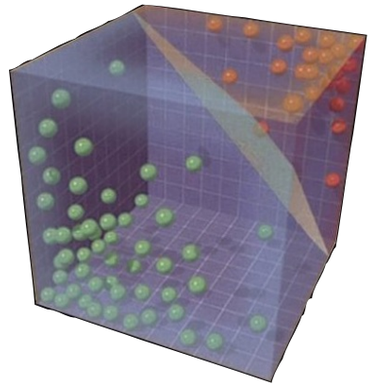
The goal of computer vision is

to give computers

**(super) human-level perception**

# typical perception pipeline

**representation**

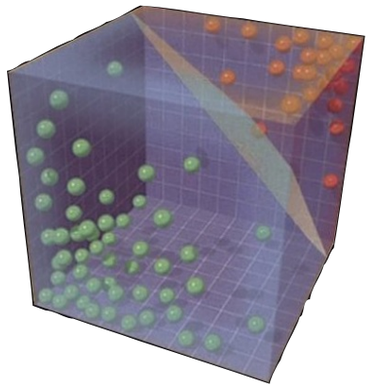


‘fancy math’

**output**

# typical perception pipeline

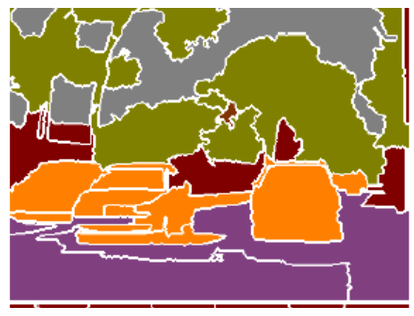
**representation**



‘fancy math’

**output**

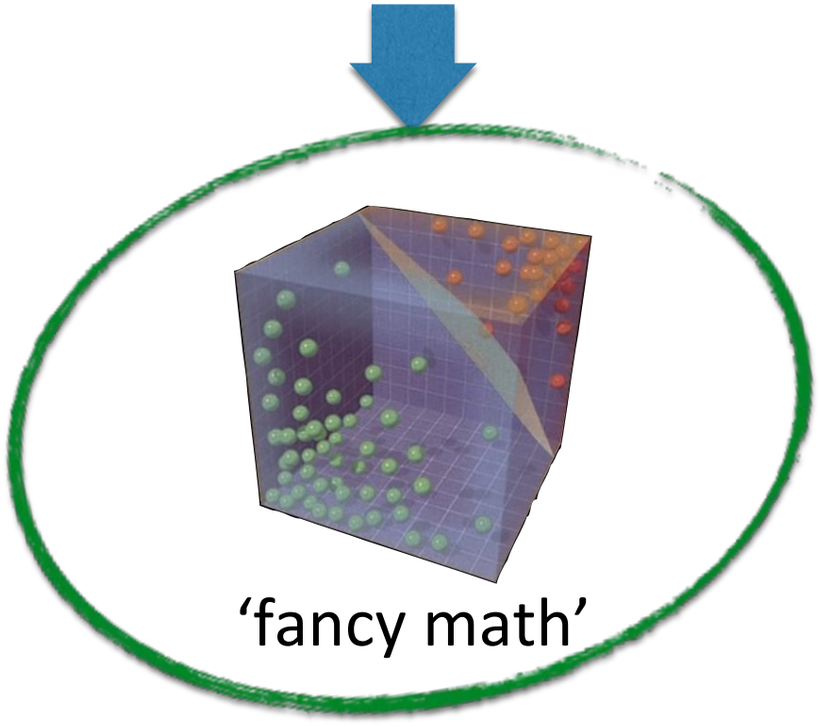
what should we look at?
(image features)



what can we understand?
(semantic segmentation)

# typical perception pipeline

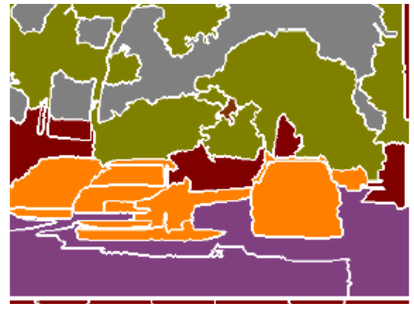
**representation**




**output**

easy to get lost in
the techniques

what should we look at?
(image features)



what can we understand?
(semantic segmentation)

Important note:

**In general, computer vision does not work**

Important note:

**In general, computer vision does not work**

(except in certain situations/conditions)

# Applications of computer vision

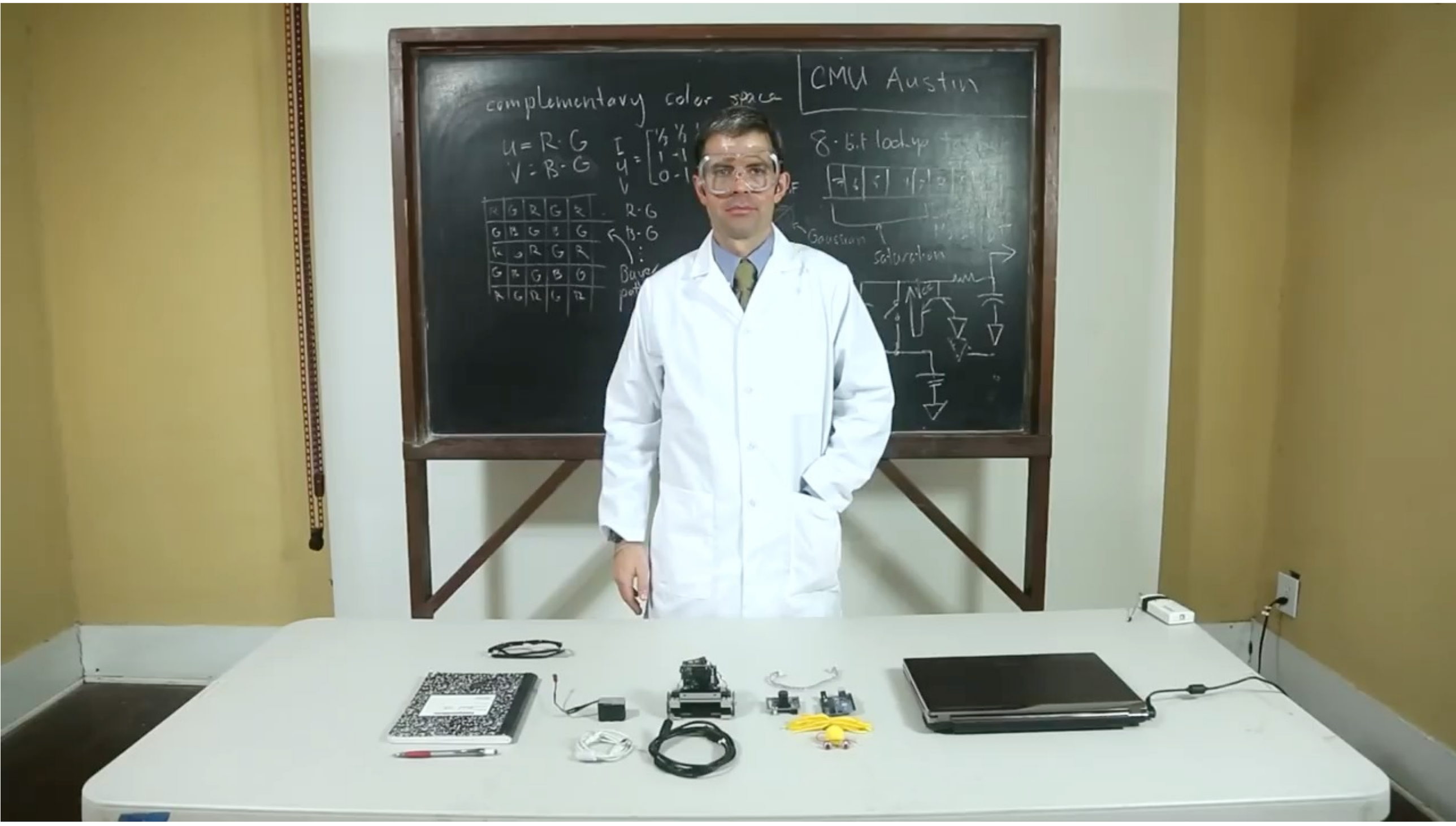
# Object Recognition



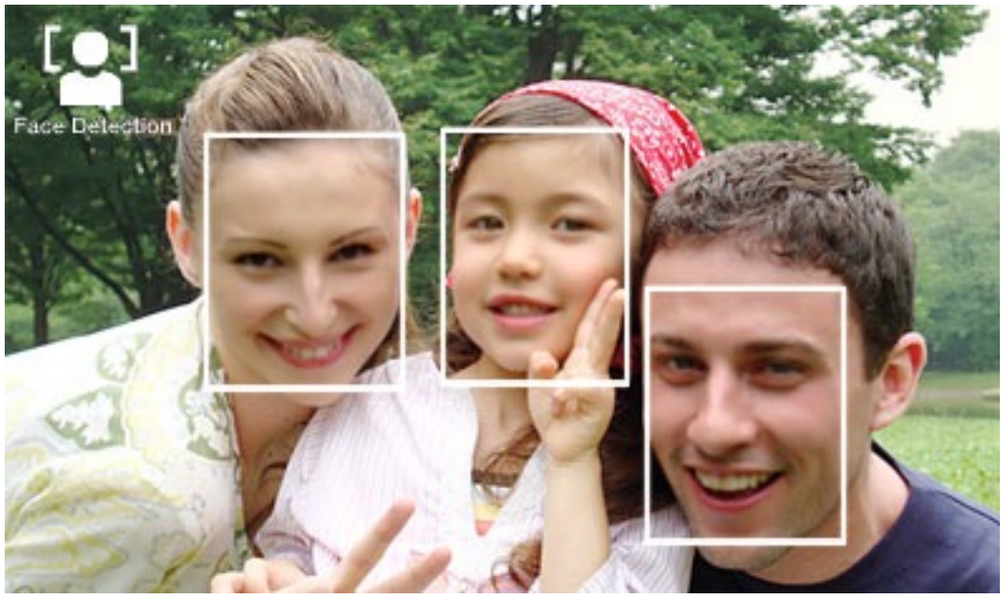
Toshiba Tech IS-910T 2013


DataLogic LaneHawk LH4000 2012

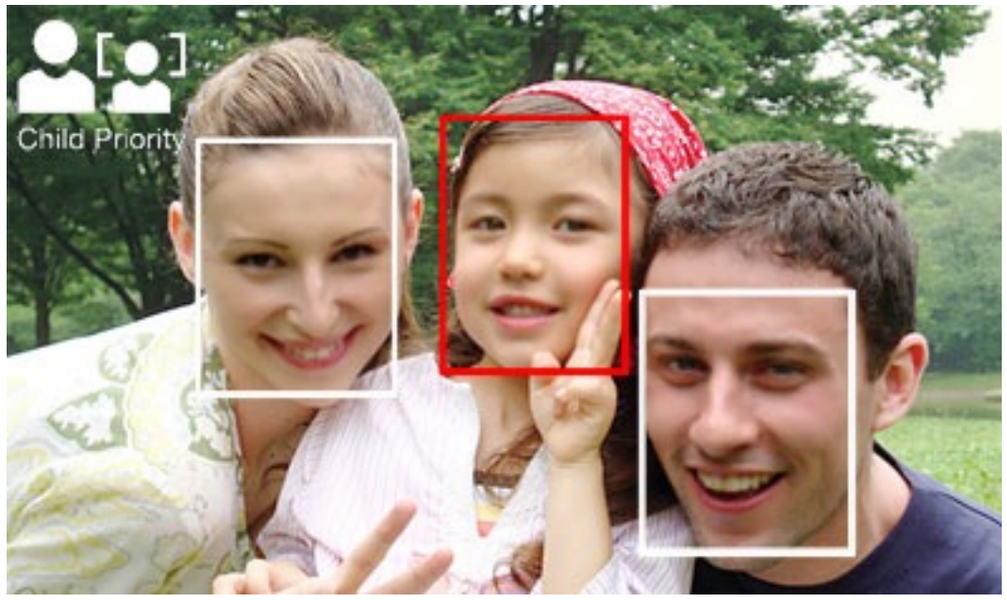
# Object Recognition



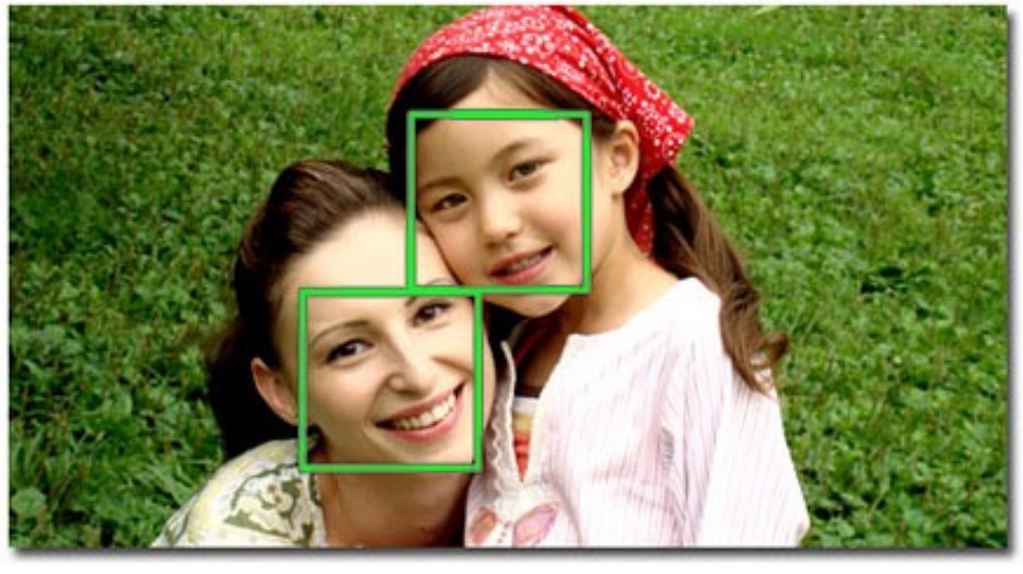
# Face detection




Sony Cyber-shot

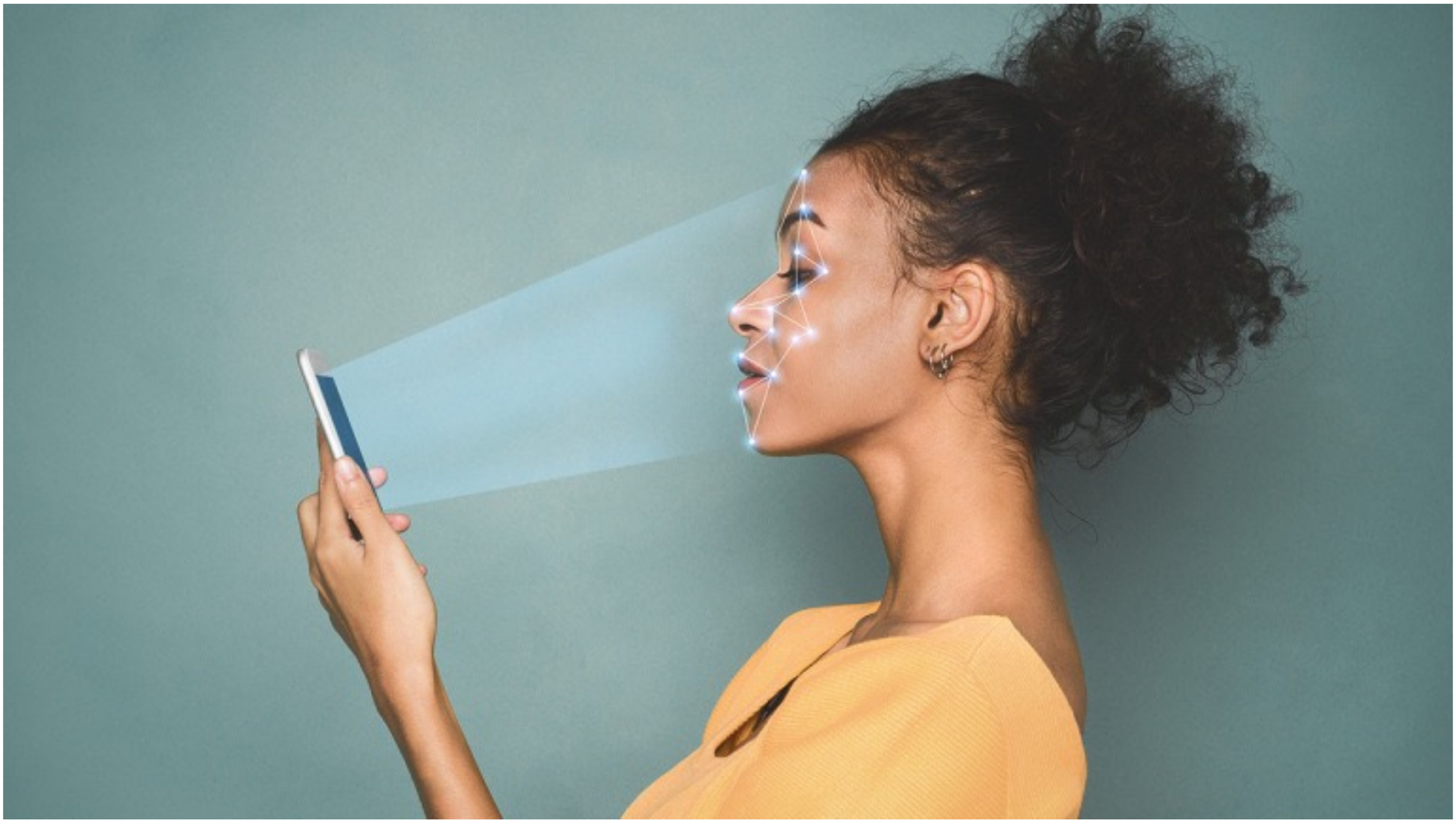



Age recognition



Smile recognition

# Face ID



# Face ID

# Identifying plants

# Google translate

# First-down line

# Vision in Cars

# Image stitching

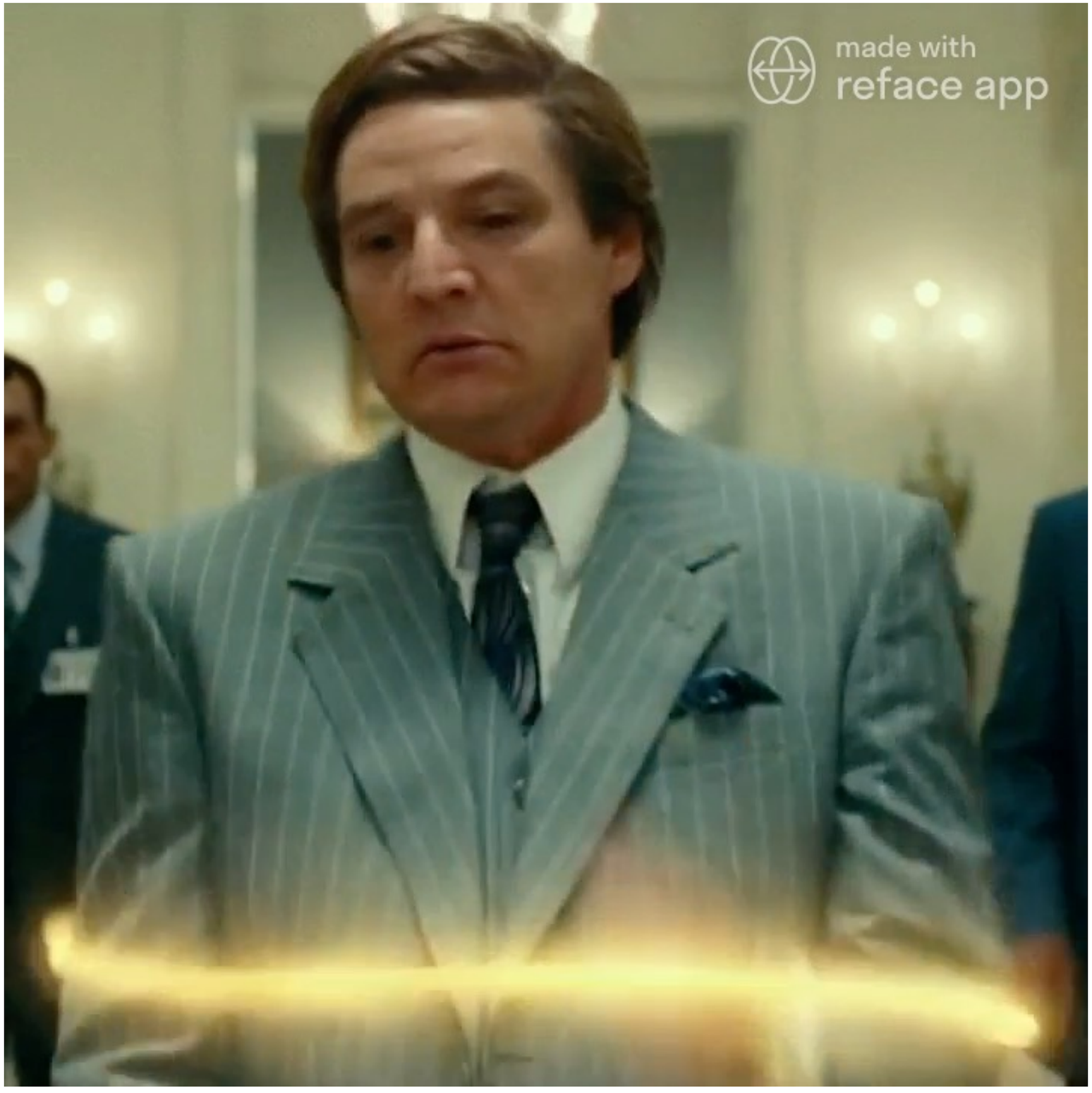
# 3D Scanning

# Style Transfer


Source: [Jamriška et al. 2019]

# Deep Fake

# Deep Fake



# 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

IBM

MIT

MIT

Toronto

UW

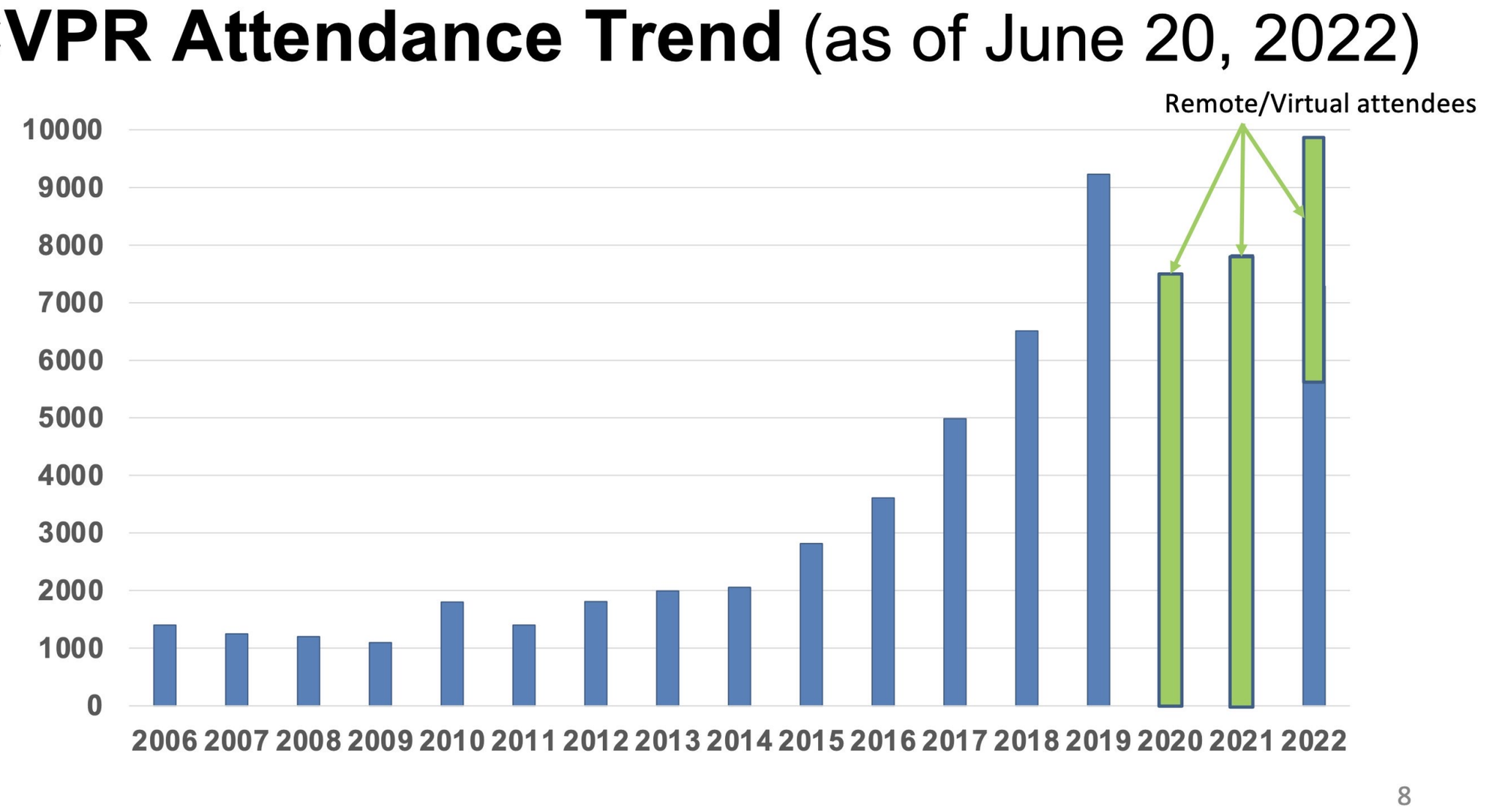

# Industry aggressively hiring CV graduates, or even students!

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

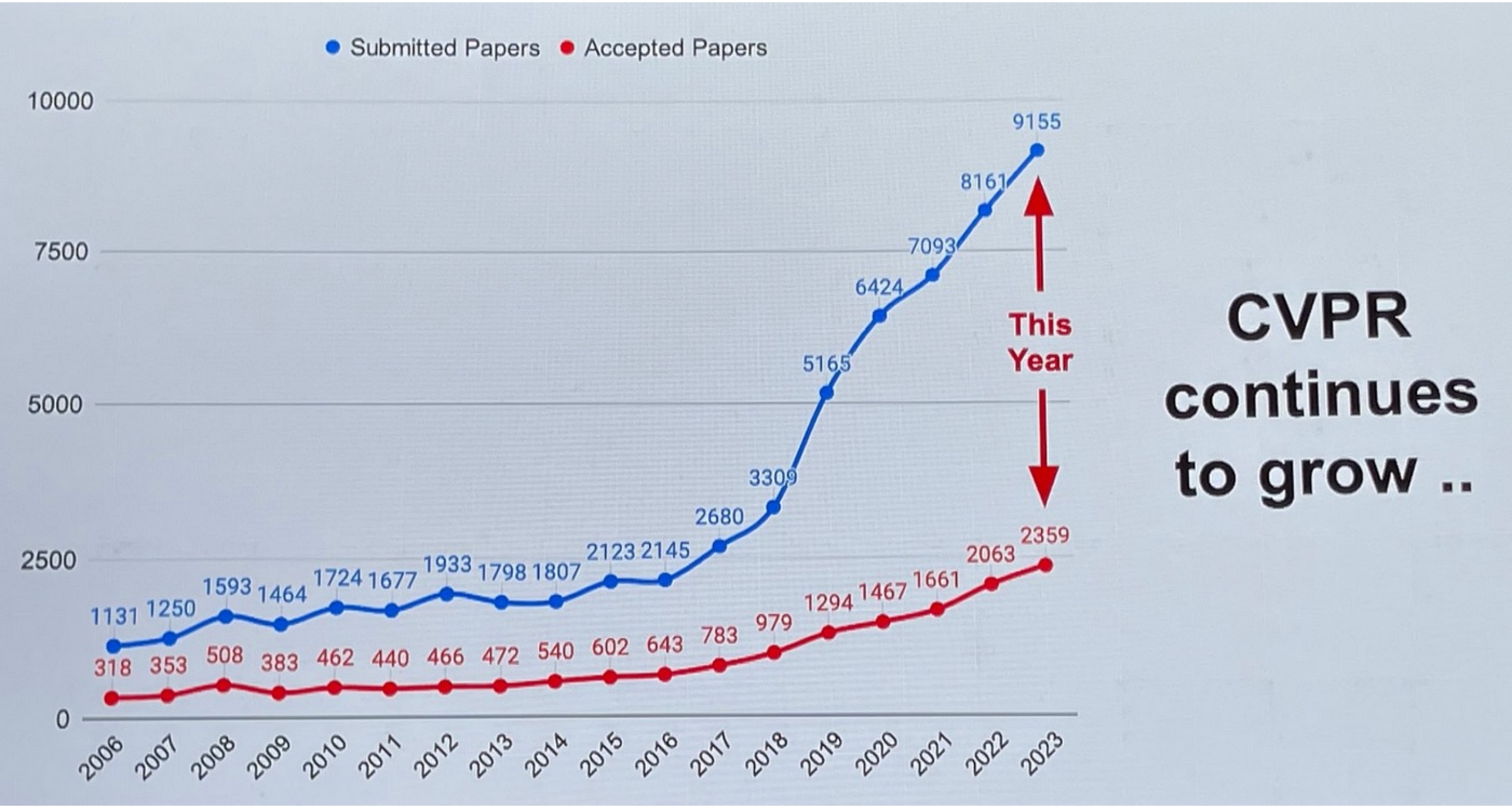
# Leading Conferences for Computer Science (https://research.com/conference-rankings/computer-science)

| Rank | Conference Details | Impact Score |
|---|---|---|
| 1 | **Computer Vision and Pattern Recognition**<br>18-06-2023 - 22-06-2023 - **Vancouver** | 60.70 |
| 2 | **European Conference on Computer Vision**<br>24-10-2022 - 28-10-2022 - **Tel Aviv** | 38.70 |
| 3 | **Neural Information Processing Systems**<br>12-12-2023 - 14-12-2023 - **New Orleans** | 38.50 |
| 4 | **International Conference on Learning Representations**<br>01-05-2023 - 05-05-2023 - **Kigali** | 35.70 |
| 5 | **International Conference on Computer Vision**<br>11-10-2021 - 11-10-2021 - **Montreal** | 31.80 |

# Stats for CVPR (Computer Vision and Pattern Recognition)



# Stats for CVPR (Computer Vision and Pattern Recognition)



| | | | |
|---|---|---|---|
| 1 | 3D from multi-view and sensors | 1,090 | 246 |
| 2 | Image and video synthesis and generation | 889 | 185 |
| 3 | Humans: Face, body, pose, gesture, movement | 813 | 166 |
| 4 | Transfer, meta, low-shot, continual, or long-tail learning | 688 | 153 |
| 5 | Recognition: Categorization, detection, retrieval | 673 | 139 |
| 6 | Vision, language, and reasoning | 631 | 118 |
| 7 | Low-level vision | 553 | 126 |
| 8 | Segmentation, grouping and shape analysis | 524 | 113 |
| 9 | Deep learning architectures and techniques | 485 | 92 |
| 10 | Multi-modal learning | 450 | 89 |
| 11 | 3D from single images | 431 | 91 |
| 12 | Medical and biological vision, cell microscopy | 420 | 53 |
| 13 | Video: Action and event understanding | 373 | 83 |
| 14 | Autonomous driving | 359 | 69 |
| 15 | Self-supervised or unsupervised representation learning | 349 | 71 |
| 16 | Datasets and evaluation | 344 | 54 |
| 17 | Scene analysis and understanding | 276 | 54 |
| 18 | Adversarial attack and defense | 274 | 61 |
| 19 | Efficient and scalable vision | 252 | 48 |
| 20 | Computational imaging | 226 | 53 |
| 21 | Video: Low-level analysis, motion, and tracking | 215 | 46 |
| 22 | Vision applications and systems | 171 | 35 |
| 23 | Vision + graphics | 155 | 32 |
| 24 | Robotics | 141 | 23 |
| 25 | Transparency, fairness, accountability, privacy, ethics in vision | 129 | 30 |
| 26 | Explainable computer vision | 107 | 24 |
| 27 | Embodied vision: Active agents, simulation | 80 | 14 |
| 28 | Document analysis and understanding | 72 | 12 |
| 29 | Machine learning (other than deep learning) | 65 | 14 |
| 30 | Physics-based vision and shape-from-X | 55 | 12 |
| 31 | Biometrics | 51 | 11 |
| 32 | Others | 47 | 12 |
| 33 | Optimization methods (other than deep learning) | 46 | 12 |
| 34 | Photogrammetry and remote sensing | 38 | 8 |
| 35 | Computer vision theory | 33 | 5 |
| 36 | Computer vision for social good | 25 | 5 |

# Computer vision at CMU

Dedicated courses for each subject we cover in this class:

- Physics-based Methods in Vision
- Geometry-based Methods in Computer Vision
- Computational Photography
- Visual Learning and Recognition
- Statistical Techniques in Robotics
- Sensors and sensing

… plus an entire department's worth of ML courses.

# Master in Computer Vision at CMU

# Course logistics

# Website

http://16385.courses.cs.cmu.edu/

(includes links to Canvas and Piazza)

# Assignments Canvas

https://canvas.cmu.edu/courses/40004

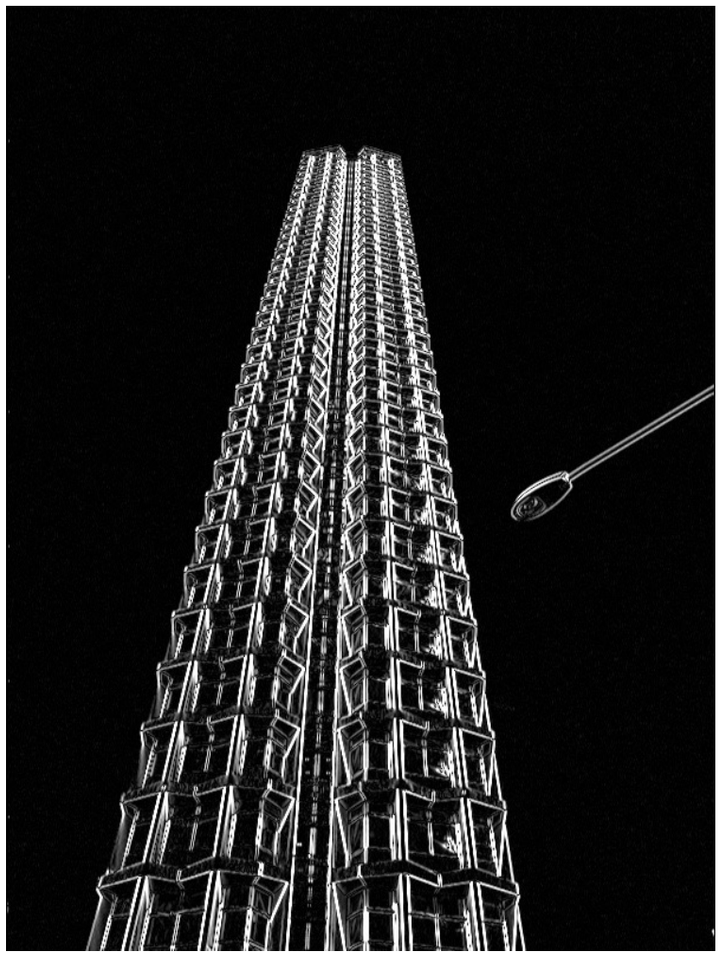
# Discussion & Notes piazza

https://piazza.com/cmu/spring2024/16385a/home

# Topics to be covered

Image processing:

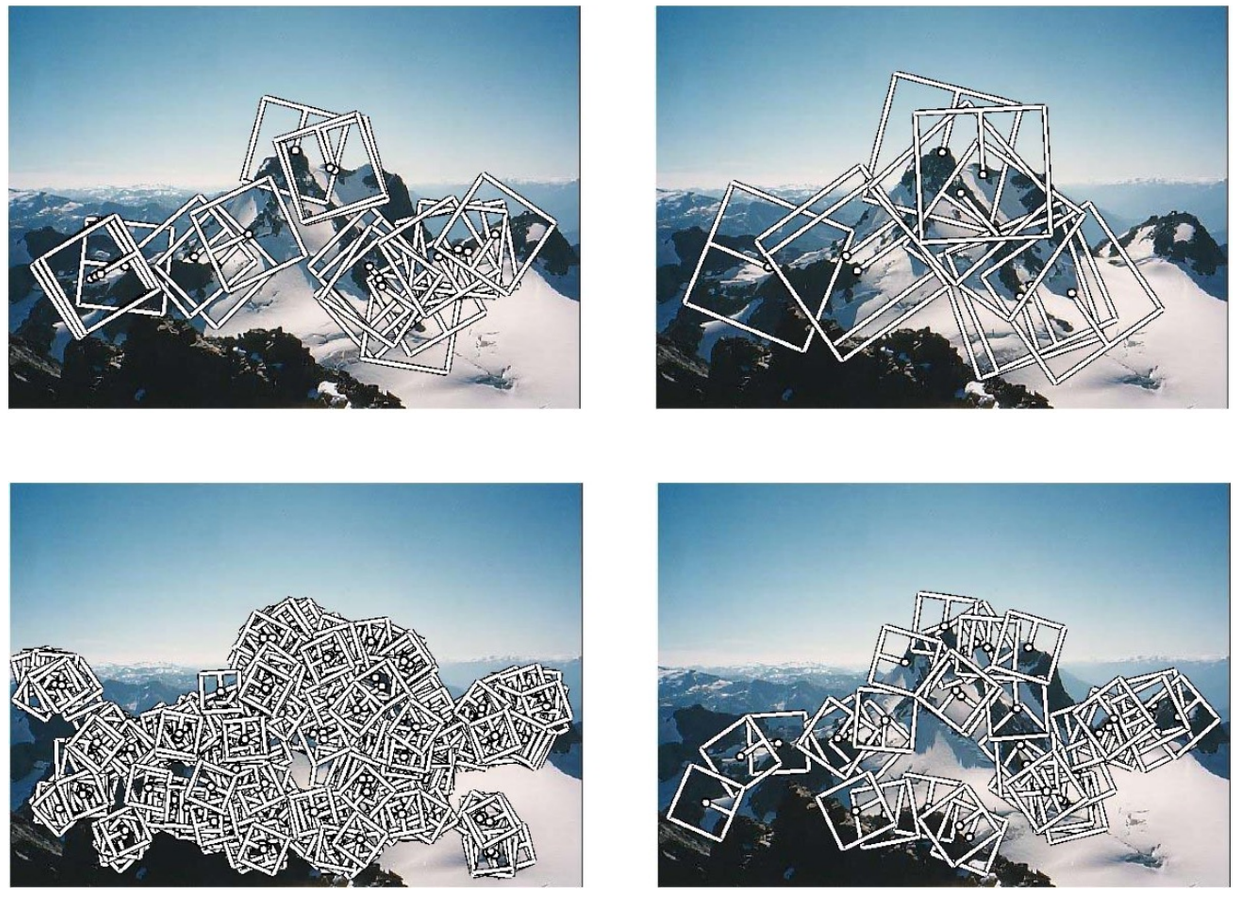
- Basics of filtering.
- Image pyramids.
- Gradients and lines.
- Hough transforms.



# Topics to be covered

Feature detection and correspondences:

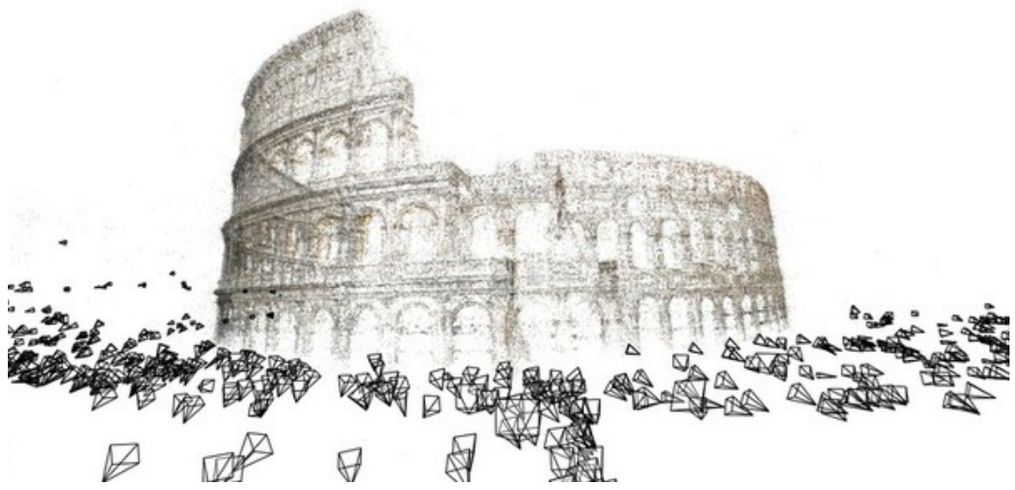
- Corner detection.
- SIFT et al.
- Feature descriptors.
- RANSAC.



# Topics to be covered

Transformations and geometry:

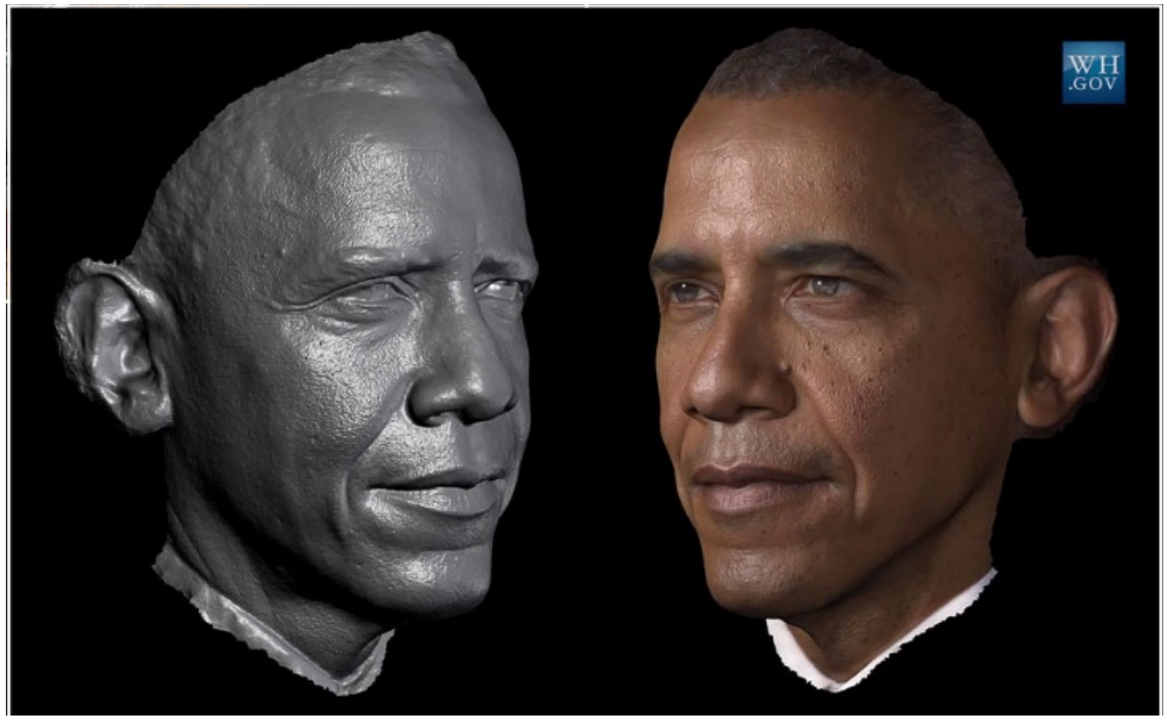
- Homographies and image alignment.
- Camera models.
- Fundamental matrix.
- Epipolar geometry and stereo.
- Structure from motion.



# Topics to be covered

Physics-based vision:

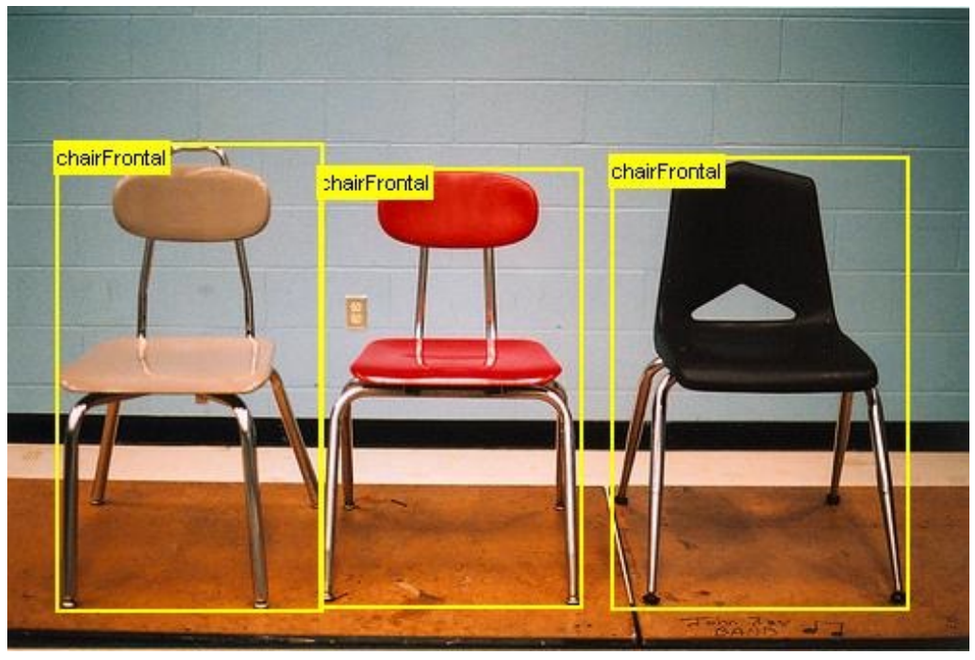
- Reflectance and image formation.
- Radiometry.
- Shape from shading.
- Photometric stereo.
- Color.



# Topics to be covered

Objects, faces, and learning:

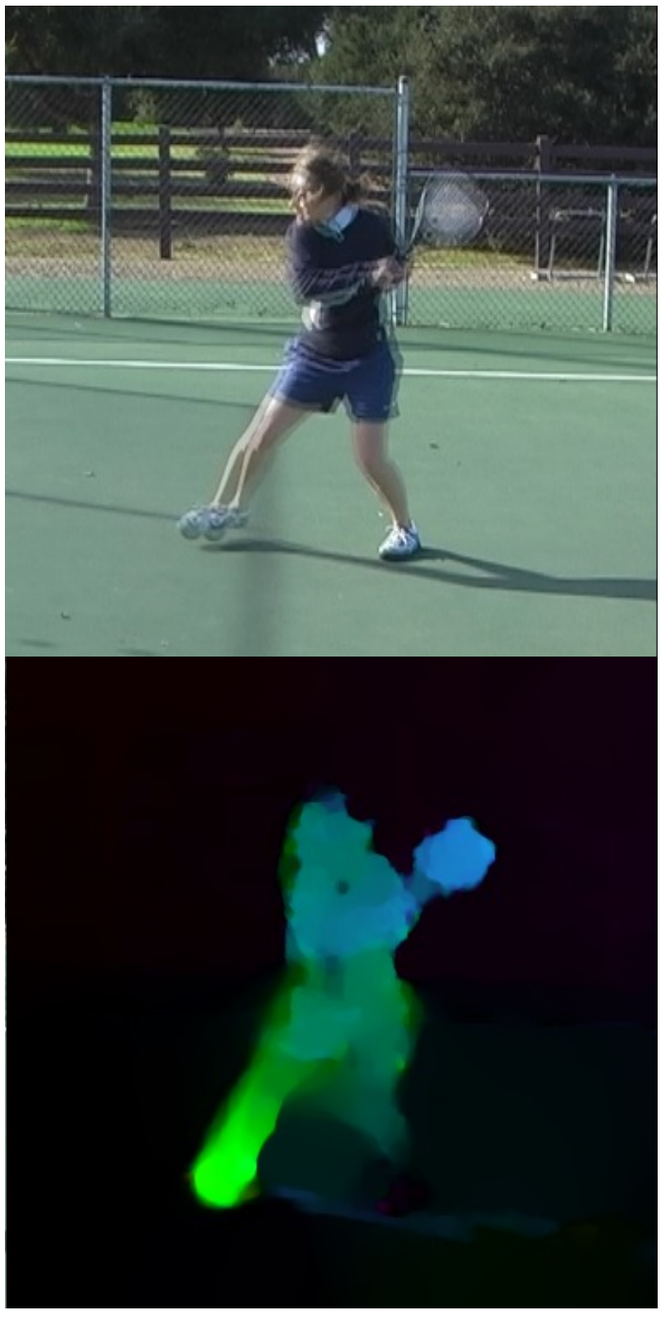
- Basics of probability.
- K-means, KNN, PCA, SVM.
- Bag of words.
- Viola-Jones face detection.
- Perceptron, backpropagation.
- Convolutional neural networks.



# Topics to be covered

Dealing with motion:

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



# Grading

- Six two-week programming assignments: 98%
- Class, Website, and Piazza participation: 2%

**Participation:**

- Be active!  Ask questions.
- Post on Piazza and course website.

# Programming Assignments

- a lot of programming in Python
- hours and hours of programming
- days and days of debugging
- generous grading policy
- take advantage of extra credit

Assignment 1 Hough Transform
Assignment 2 Homography
Assignment 3 Stereo
Assignment 4 Bag of Words
Assignment 5 Convolutional Neural Nets
Assignment 6 Image Alignment

# Programming Assignments

- a lot of programming in Python
- hours and hours of programming
- days and days of debugging
- generous grading policy
- take advantage of extra credit

Assignment 1 Hough Transform
Assignment 2 Homography
Assignment 3 Stereo
Assignment 4 Bag of Words
Assignment 5 Convolutional Neural Nets
Assignment 6 Image Alignment

**Seriously.. a lot of programming, so start early!**

# Leniency

Late days for <u>programming assignments</u>:

- 10% reduction of points per late day
- 6 free late days <u>total</u>
- Intended to cover sick days, interviews, family emergencies, etc. **<u>No additional late days will be provided!!</u>**
- Use them wisely… save for later (harder) assignments and for emergencies!

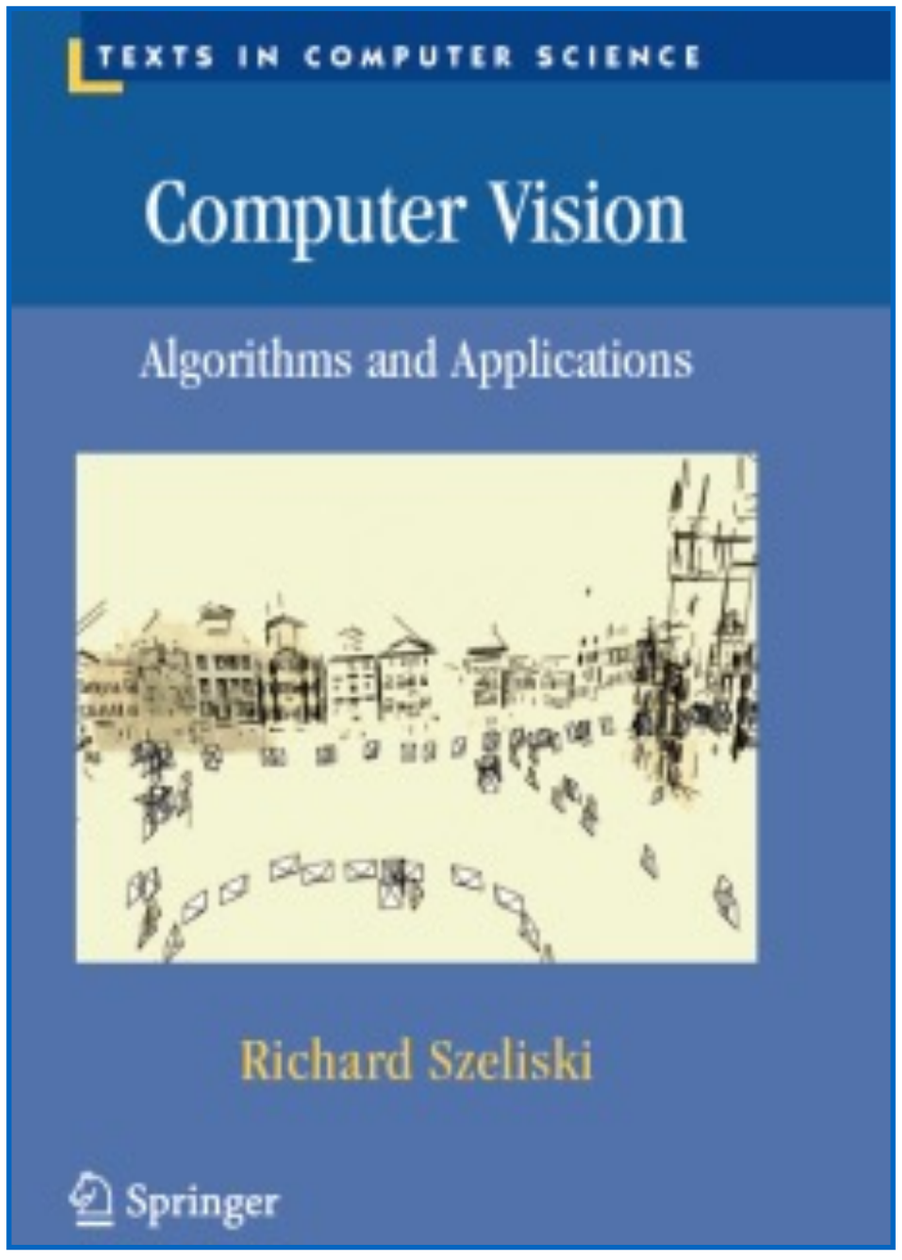
# Collaboration Policy

(see **Course Info** page on website)

"Students in 16-385 are absolutely encouraged to talk to each other, to the TAs, to the instructors, or to anyone else about course assignments. Any assistance, though, must be limited to discussion of the problems and sketching general approaches to a solution. Each student must write their own code and produce their own writeup. Consulting another student's solution, or solutions from the internet, is prohibited on assignments. These and any other form of collaboration constitute cheating. If you have any question about whether some activity would constitute cheating, just be cautious and ask the instructor before proceeding!

You may not supply code or assignment writeups you complete during 16-385 to other students in future instances of this course or make these items available (e.g., on the web) for use in future instances of this course (just as you may not use work completed by students who've taken the course previously). We encourage you to use public source control hosts like Github for your assignments, however please be sure to make your programming assignment repositories private."

# Book




PDF online

http://szeliski.org/Book/

# Contact information

- Feel free to email us about administrative questions.
    - please use [16385] in email title!
- Lecture questions should be asked on course website (or in lecture), and assignment/logistic questions should be asked on Piazza.
    - we won't answer technical questions through email.
    - you can post anonymously if you prefer.
- Office hours will be determined by poll.
    - feel free to email me about additional office hours.