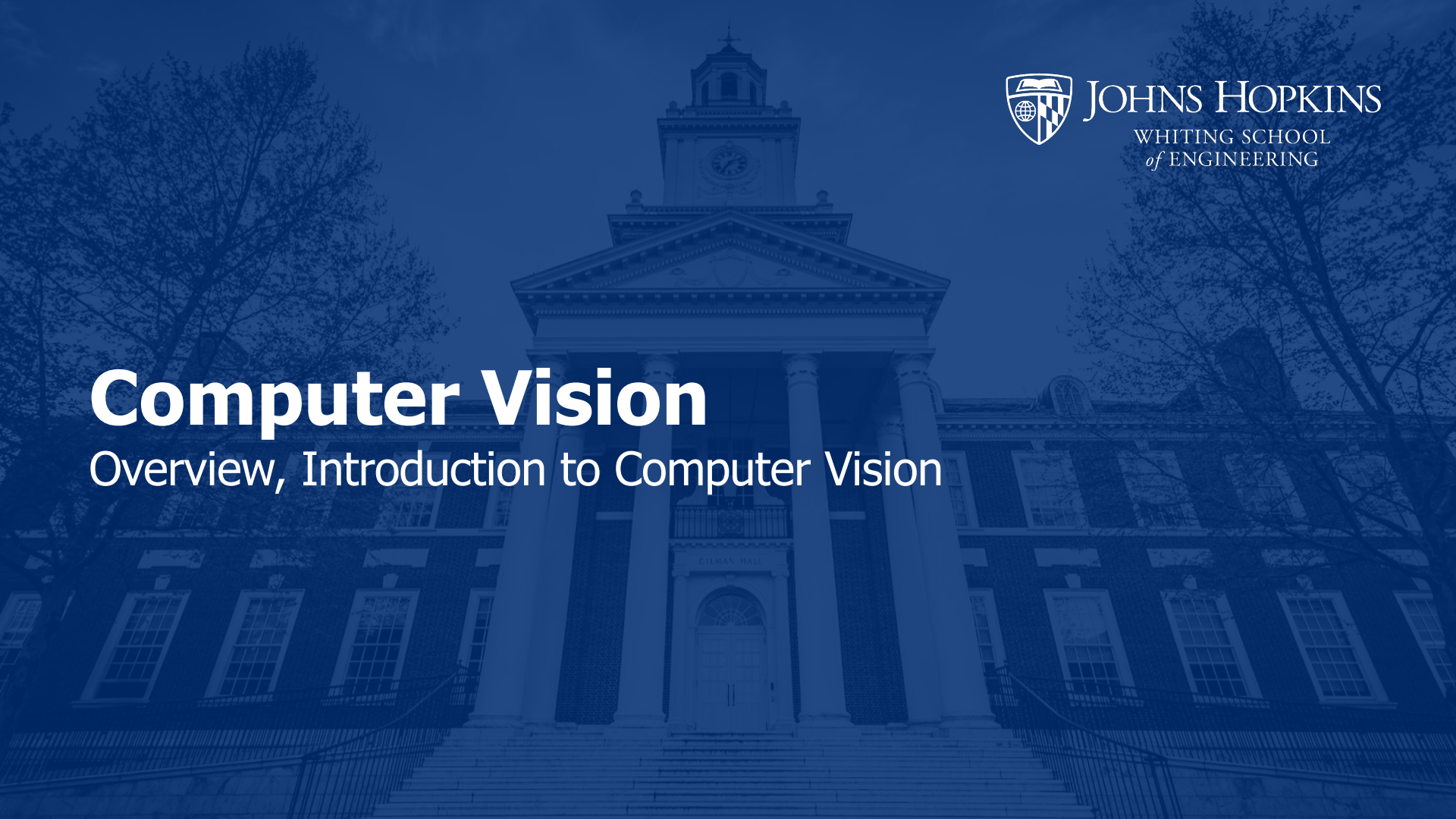


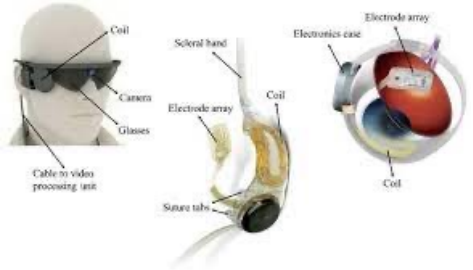
Computer Vision

Overview, Introduction to Computer Vision

Introduction

- Instructor: Kapil D. Katyal, kkatyal2@jhu.edu
  - Principal Staff Member at JHU/APL
  - Assistant Research Professor in CS department
  - PhD from JHU under Prof. Greg Hager
  - Background in Robot Perception

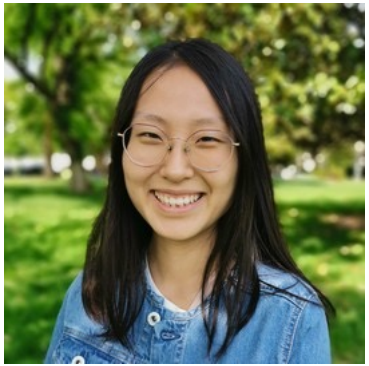
- Course Logistics
  - Class times: Tuesdays / Thursdays: 12:00 – 1:15 pm
  - Office hours: Fridays: 9:00 AM – 10:00 AM + by request
  - Course Location: 210 Hodson
  - Course Prerequisites: Intro to Programming (Python), Linear Algebra & Probability/Statistics
  - Course web page: canvas.jhu.edu



Teaching and Course Assistants

- TA: Zhuowan Li, zli110@jhu.edu
  - Office hours: Thursdays 10 am – 12 pm

- Course Assistants
  - Kaan Elgin, telgin1@jhu.edu
    - Office hours: Mondays 12 – 1 pm
  - Ju He, jhe47@jh.edu
    - Office hours: Fridays 9 – 10 am
  - Jay Paranjape, jparanj1@jhu.edu
    - Office hours: Fridays 5 – 6 pm


Zhuowan Li


Kaan Elgin


Ju He

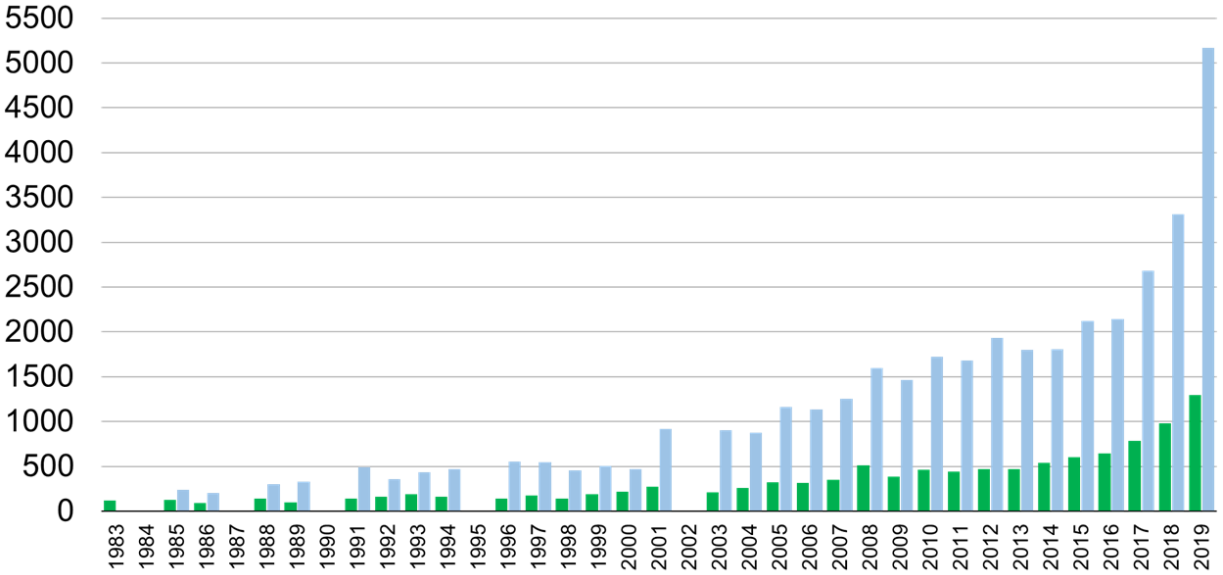

Jay Paranjape

Course Objectives

- Understand the principles of computer vision including cameras, image representations, image processing, feature extraction/detection, stereo camera/3D geometry, camera calibration, object recognition, motion, segmentation and deep learning principles.
- Apply the course material in a variety of vision applications including object detection, semantic segmentation, stereo camera calibration and pose estimation.
- Understand papers presented at leading computer vision conferences and journals.
- Develop the foundation to pursue research in computer vision.

Why Study Computer Vision?

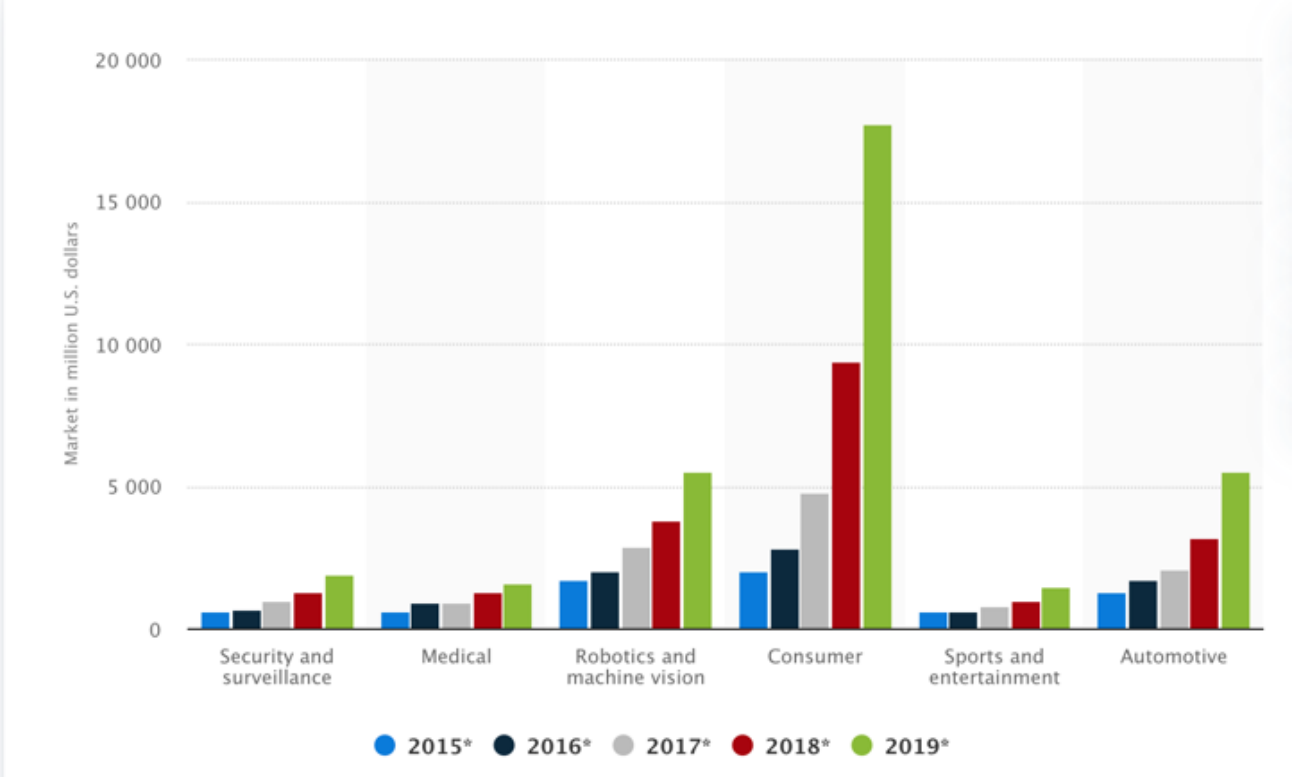
- Tremendous research advancements in the past decade
- Significantly more opportunities available in research



Number of submitted (blue) and accepted (green) papers in CVPR by year.

https://medium.com/reconstruct-inc/the-golden-age-of-computer-vision-338da3e471d1

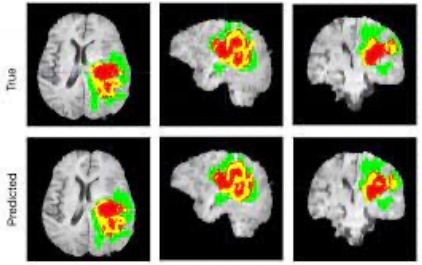
Significant Industry Demand



https://indatalabs.com/blog/applications-computer-vision-across-industries

Applications

Autonomous Driving



Tumor Detection

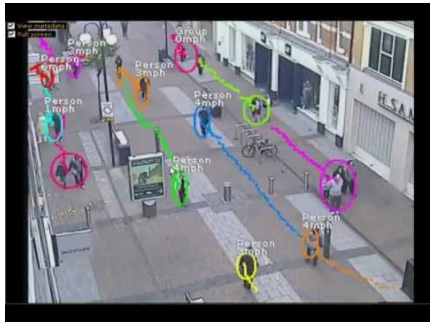
Industrial Automation

Optical Character Recognition

Object Recognition - Grocery

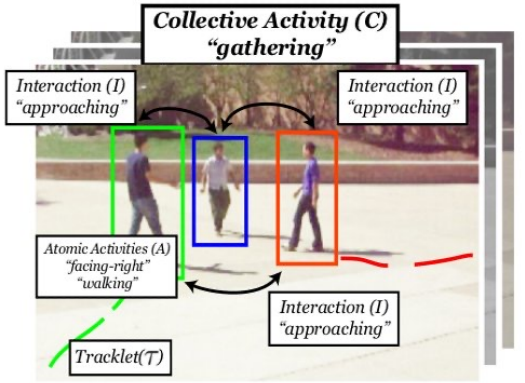
Style Transfer

More Applications

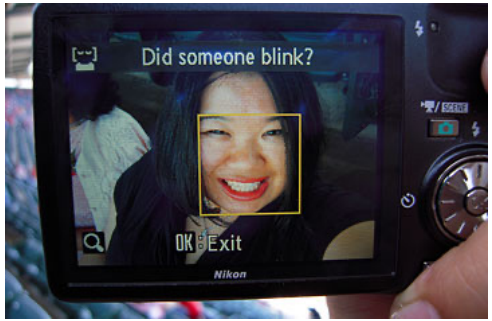


People Detection and Tracking

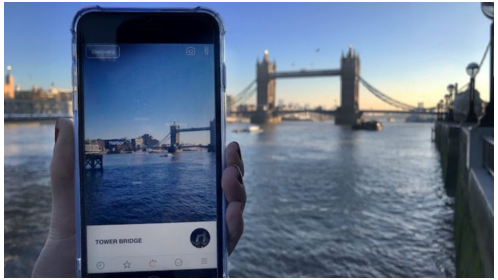
Gaming



Activity Recognition



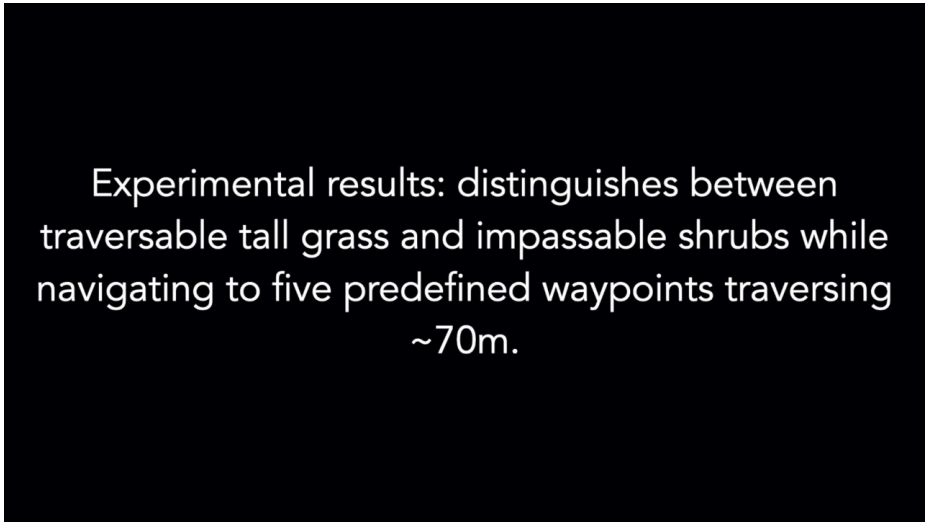
Face/Smile/Blink Detection



Landmark Recognition

Biometrics

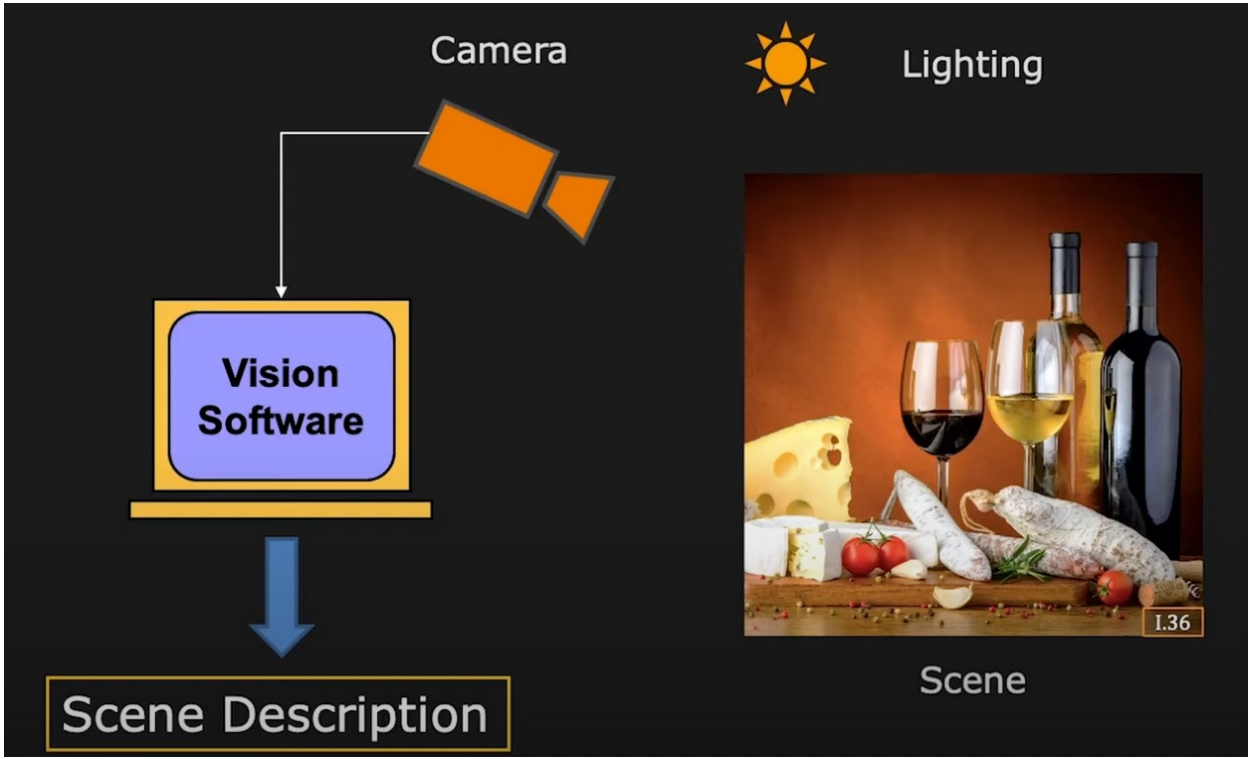
Robotic Navigation Applications



Traversability Estimation

High Speed Obstacle Avoidance

What is Computer Vision?



What is Computer Vision?

- Allowing computers to understand a scene the way humans do (emulate human vision)
- Being able to say things about the objects in the scene and the relationship to the environment
- Very difficult problem – no one really knows how to the human visual cortex works
- Inherently multi-disciplinary
  - optics, signal processing, computer science, neuroscience

What Types of Problems Can Computer Vision Solve?

- Object detection
- Semantic segmentation
- Activity recognition
- Motion and Tracking
- 6DoF pose estimation
- Traversability estimation
- …

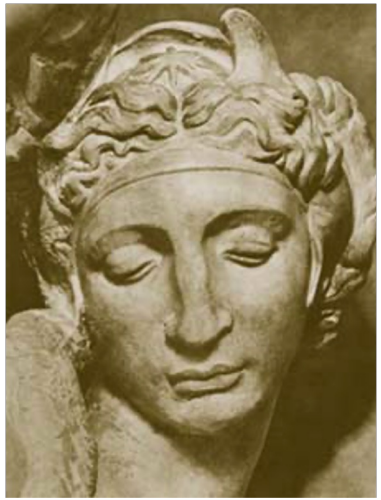
Challenges

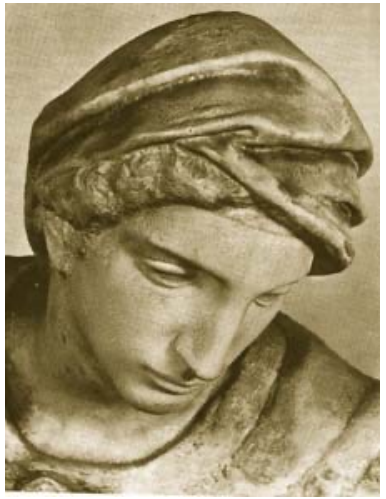
What we see

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

What a computer sees

Challenges: Viewpoint Variation



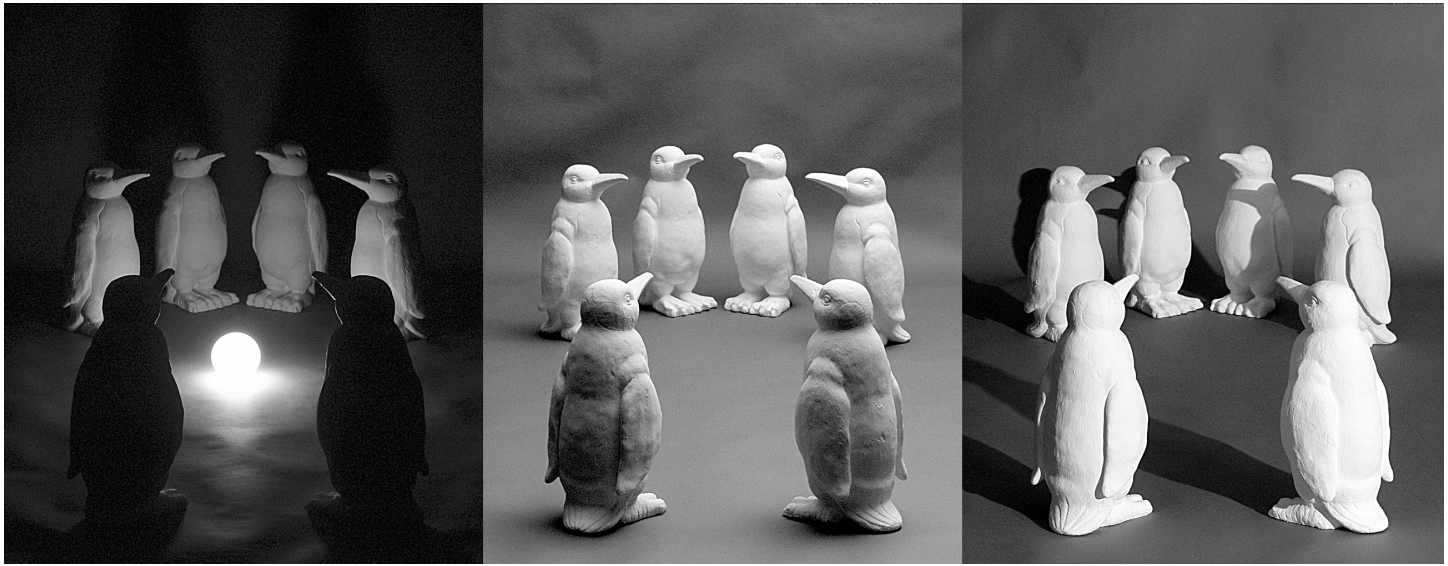


Robust to different viewpoints

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Challenges: Illumination



Robust to different lighting conditions

image credit: J. Koenderink

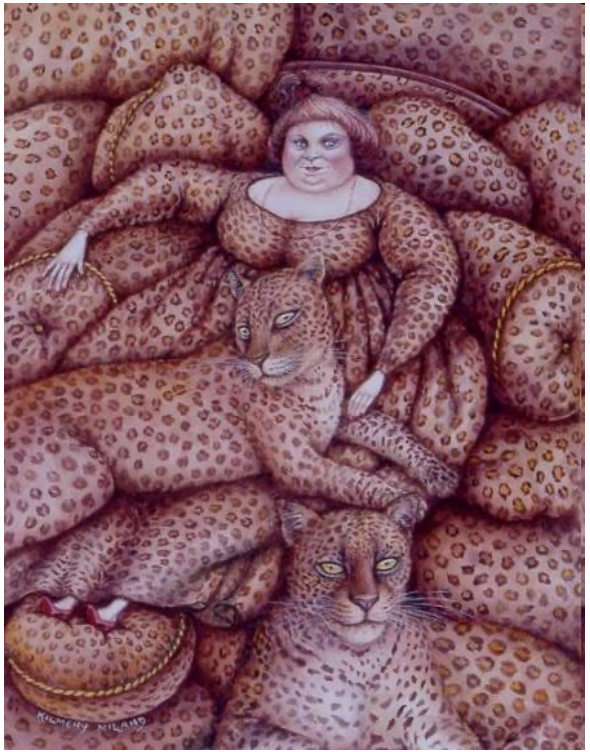
Challenges: Scale

slide credit: Fei-Fei, Fergus & Torralba

Challenges: Deformation

Challenges: Occlusion

Challenges: Background Clutter



Kilmeny Niland. 1995

Challenges: Object Similarity

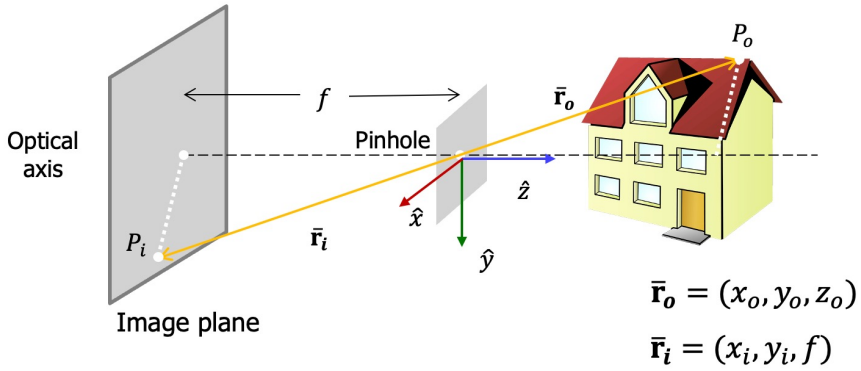
J. Tompkin, Brown

Challenges: Intra-Class Variation

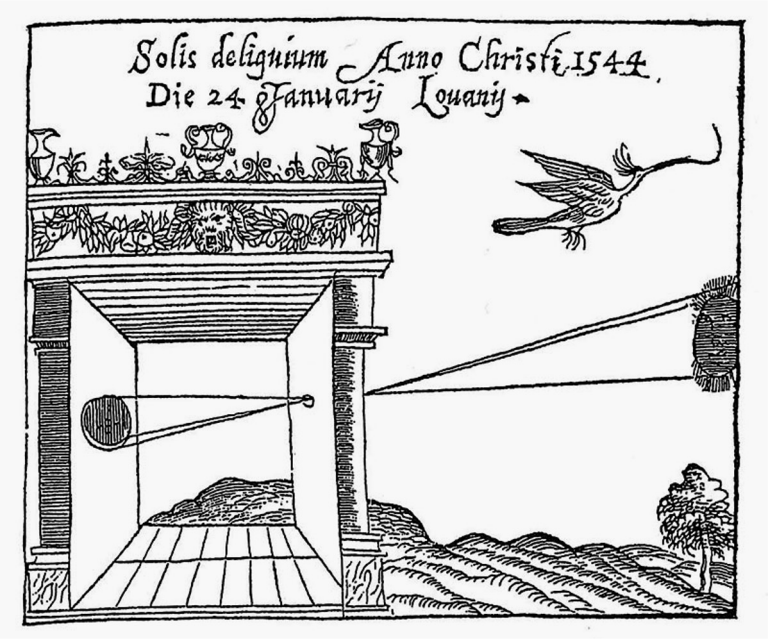
Topics covered

- Part 1: Image Formation
- Part 2: Image Geometry
- Part 3: Machine Learning

Image Formation

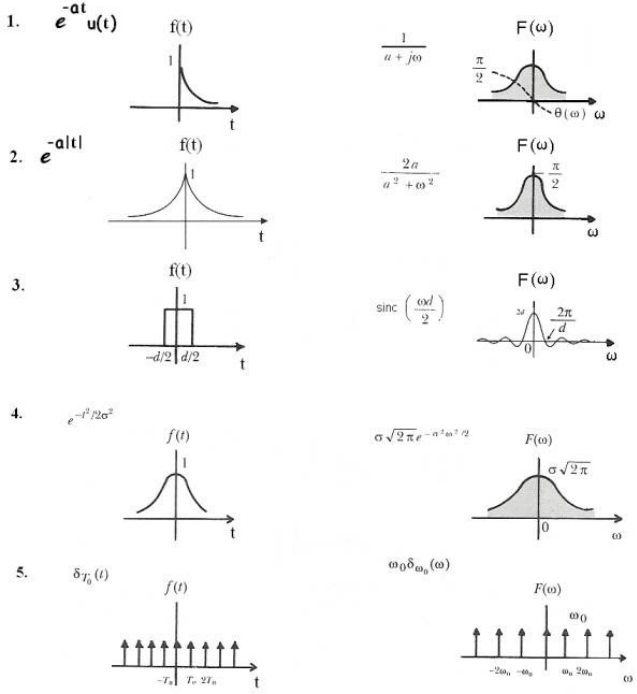


Pinhole Camera Model



The *Camera Obscura*

Image Processing



Fourier Transform
Sampling, Convolution

Image enhancement
Feature detection

Edge and Boundary Detection

Original Image

Edges/Boundary

Feature Detection

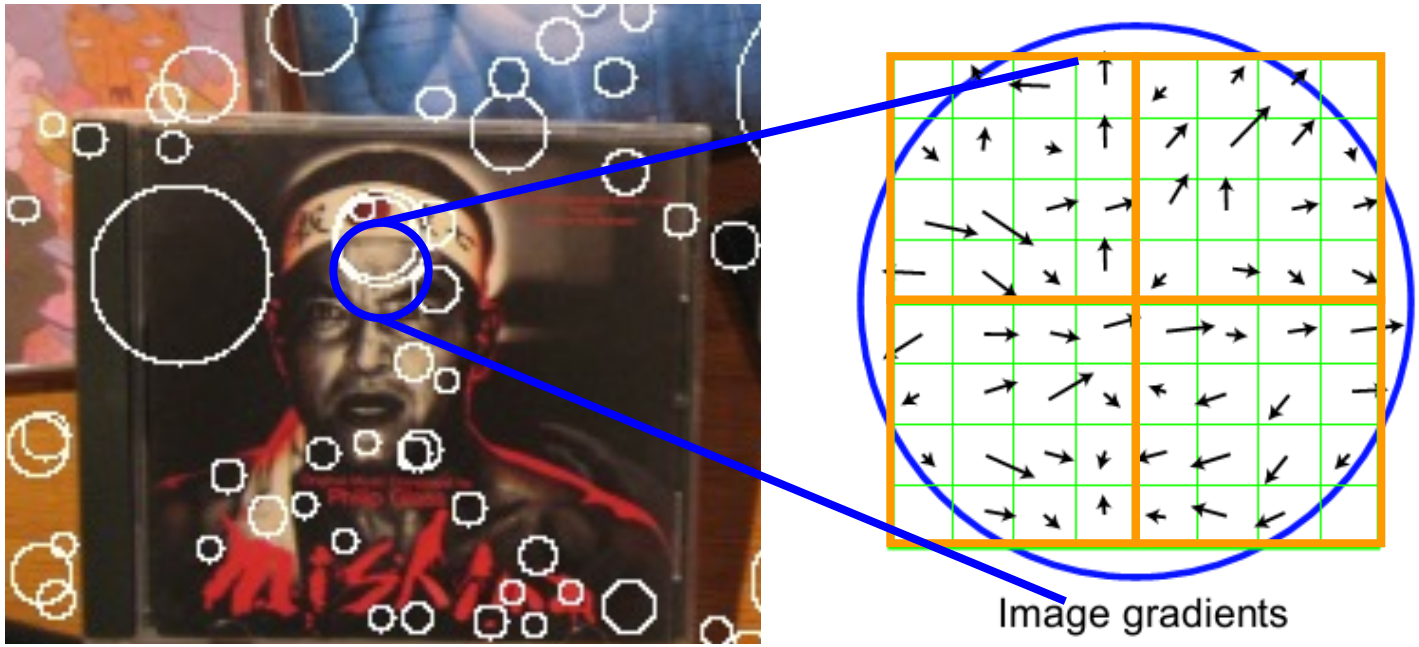
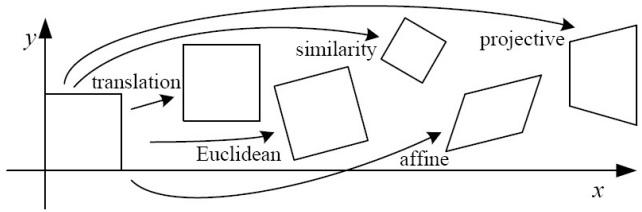
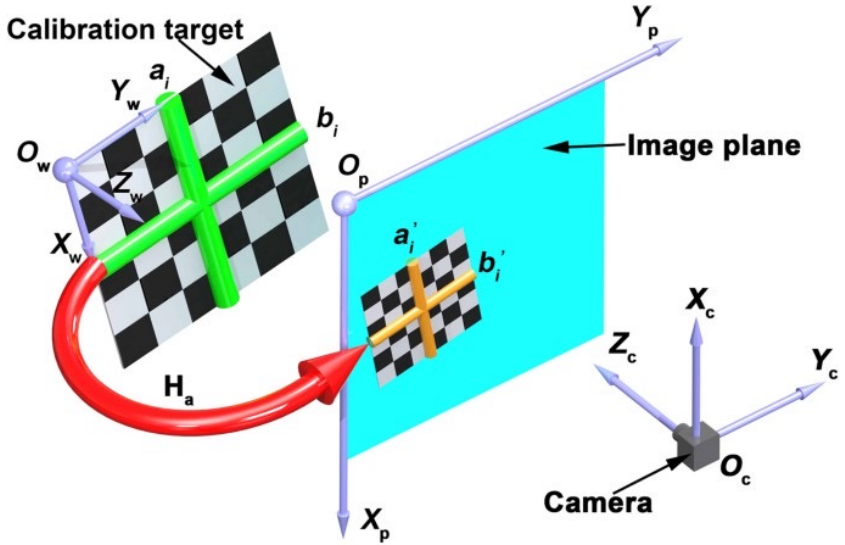


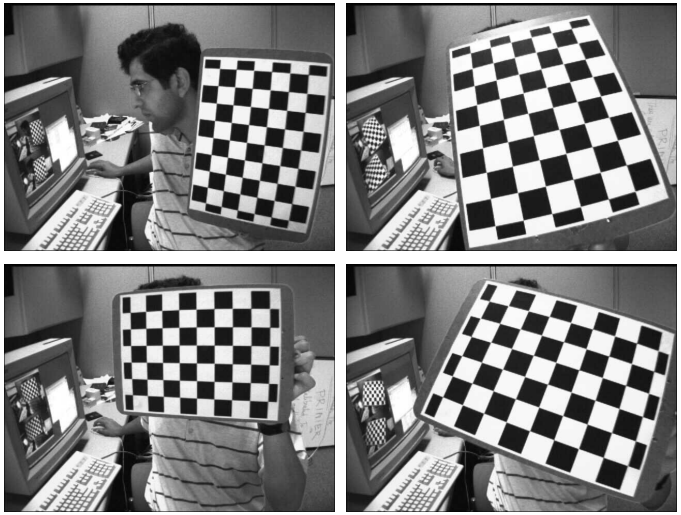
Image Geometry / Alignment



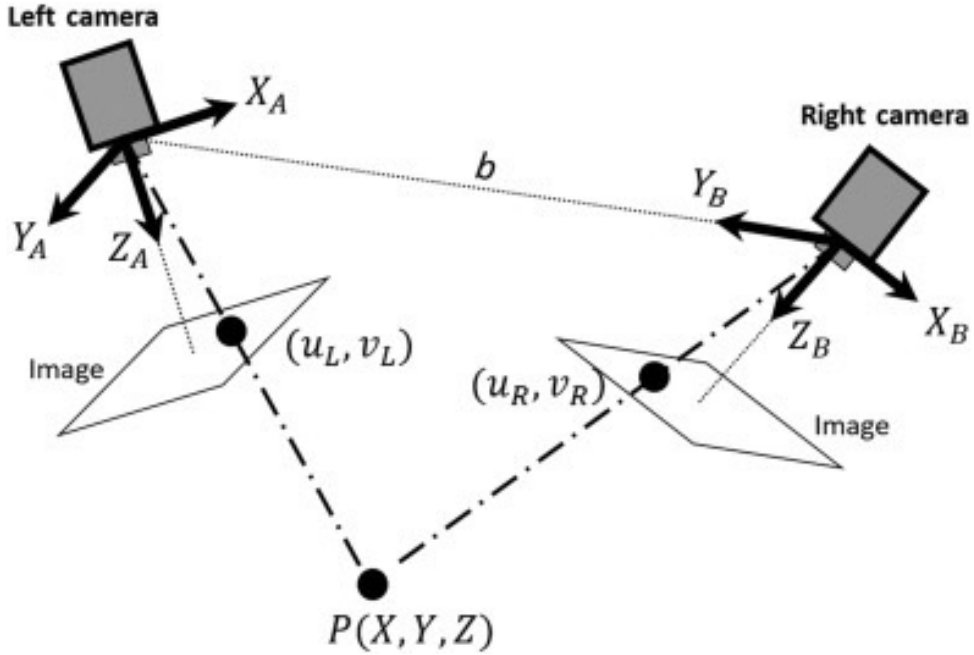
Name	Matrix	# D.O.F.	Preserves:	Icon
translation	$\begin{bmatrix} \boldsymbol{I} \mid \boldsymbol{t} \end{bmatrix}_{2\times3}$	2	orientation $+\cdots$	
rigid (Euclidean)	$\begin{bmatrix} \boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times3}$	3	lengths $+\cdots$	
similarity	$\begin{bmatrix} s\boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times3}$	4	angles $+\cdots$	
affine	$\begin{bmatrix} \boldsymbol{A} \end{bmatrix}_{2\times3}$	6	parallelism $+\cdots$	
projective	$\begin{bmatrix} \tilde{\boldsymbol{H}} \end{bmatrix}_{3\times3}$	8	straight lines	

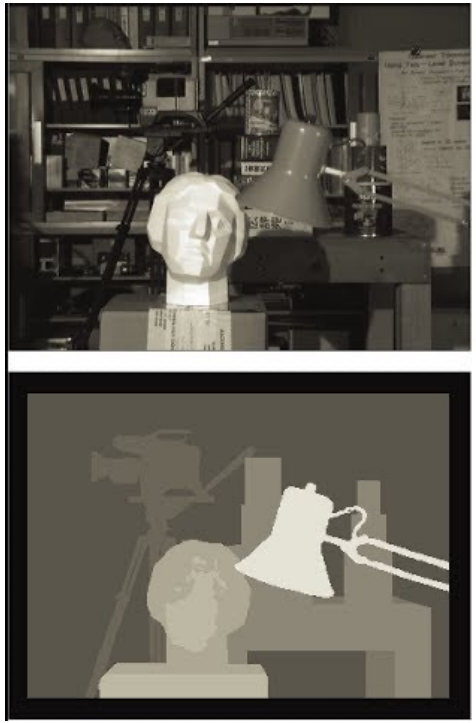
Camera Calibration



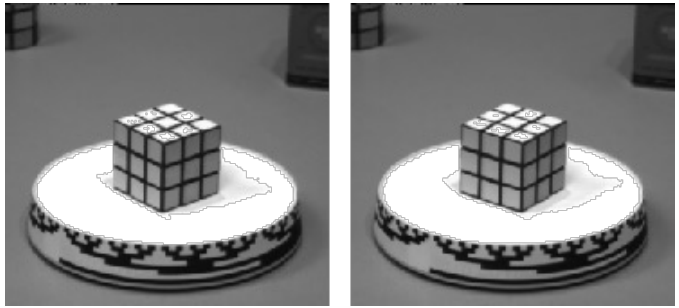


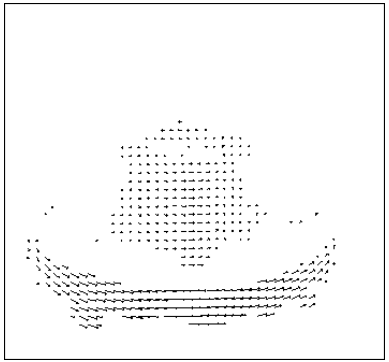
Stereo Vision

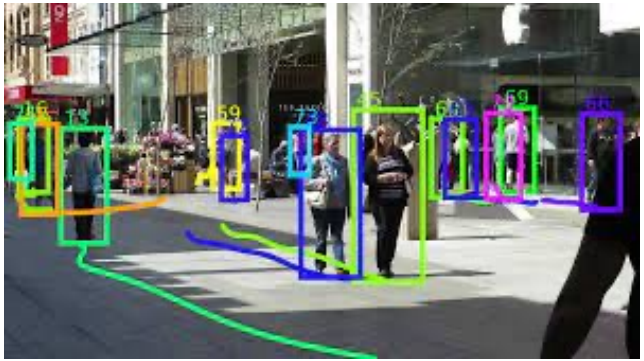




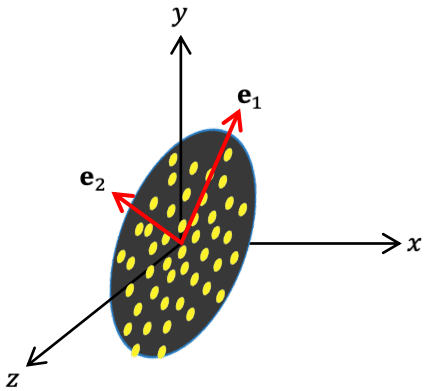
Optical Flow, Tracking and Prediction



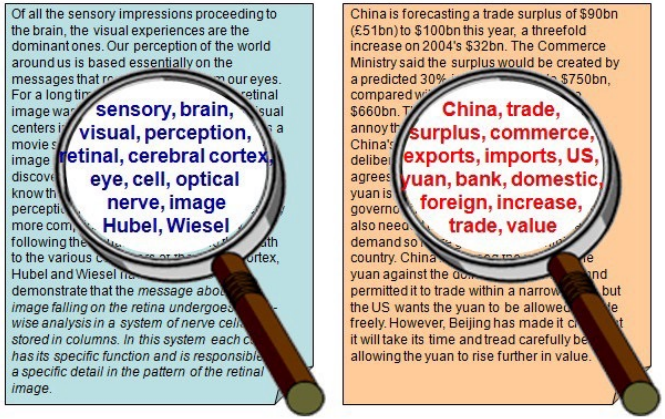




Classical Machine Learning Techniques

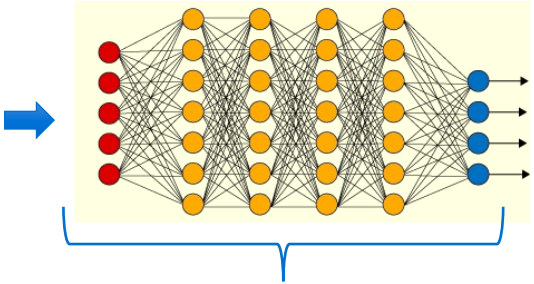


Dimensionality Reduction



Object Recognition Using Visual Bag of Words

Introduction to Deep Neural Networks



"Brown Bear"

F( ;W)

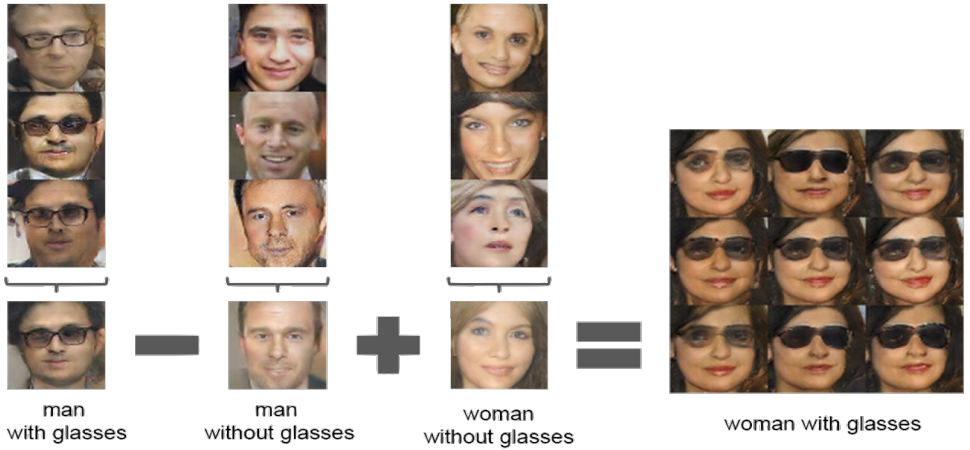
A few 10's of millions of parameters that are fit to a very large **labelled** data set including lots of brown bears

Classification vs Detection

Semantic Segmentation

Generative Neural Networks

Designed to Deceive: Do These People Look Real to You?



Evaluation and Grading

- Discussion board and general participation (10%)
- Weekly Python notebooks to reinforce course lecture notes (10%)
- Weekly-ish Quizzes (10%)
- 3 Homework assignments: (25%)
  - The first will be on image processing, template matching and Hough transform.
  - The second will be on binocular image matching and/or image stitching.
  - The third will be on deep learning basics.
- A Midterm exam covering Lectures 1 through 16 (20%)
- Course project: (25%)
  - This will be a self-directed course project that allows the student to dive deeper on a particular area of computer vision. The student will submit a project proposal that will be reviewed by the instructor, a mid-point check-in update and a final report and presentation.

Course Projects

- Ideally groups of 2-4 people. Please see me if you want to form a group with more or less students.
- Opportunity to deep dive into an area of interest!
- Potentially submit results to conference or workshop
- **Who already has an idea of what they want to work on?**
- **Who already knows they will work on deep learning?**

Questions?