Welcome!

Data Science 281 — Spring 2023

Introductions

Preferred name & pronouns

Geographic location

Professional background

Experience with computer vision

Favorite photo you’ve taken

Introductions

Preferred name & pronouns – *Rachel Brown (she/her)*

Geographic location – *Portland, Oregon*

Professional background – *Research, Human Perception & Computer Graphics*

Experience with computer vision – *Basically my day job* (NVIDIA)

Favorite photo I've taken:

Introductions

Preferred name & pronouns

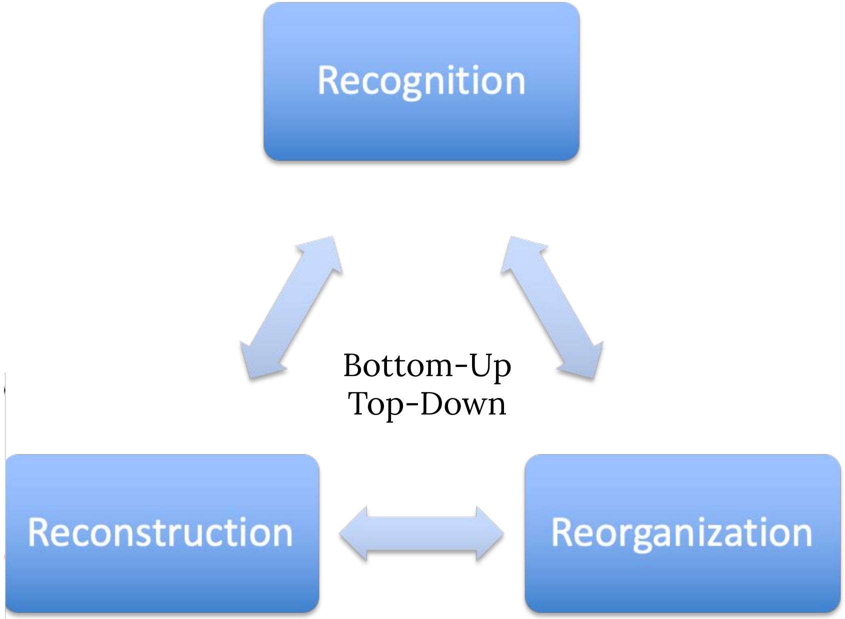
Geographic location

Professional background

Experience with computer vision

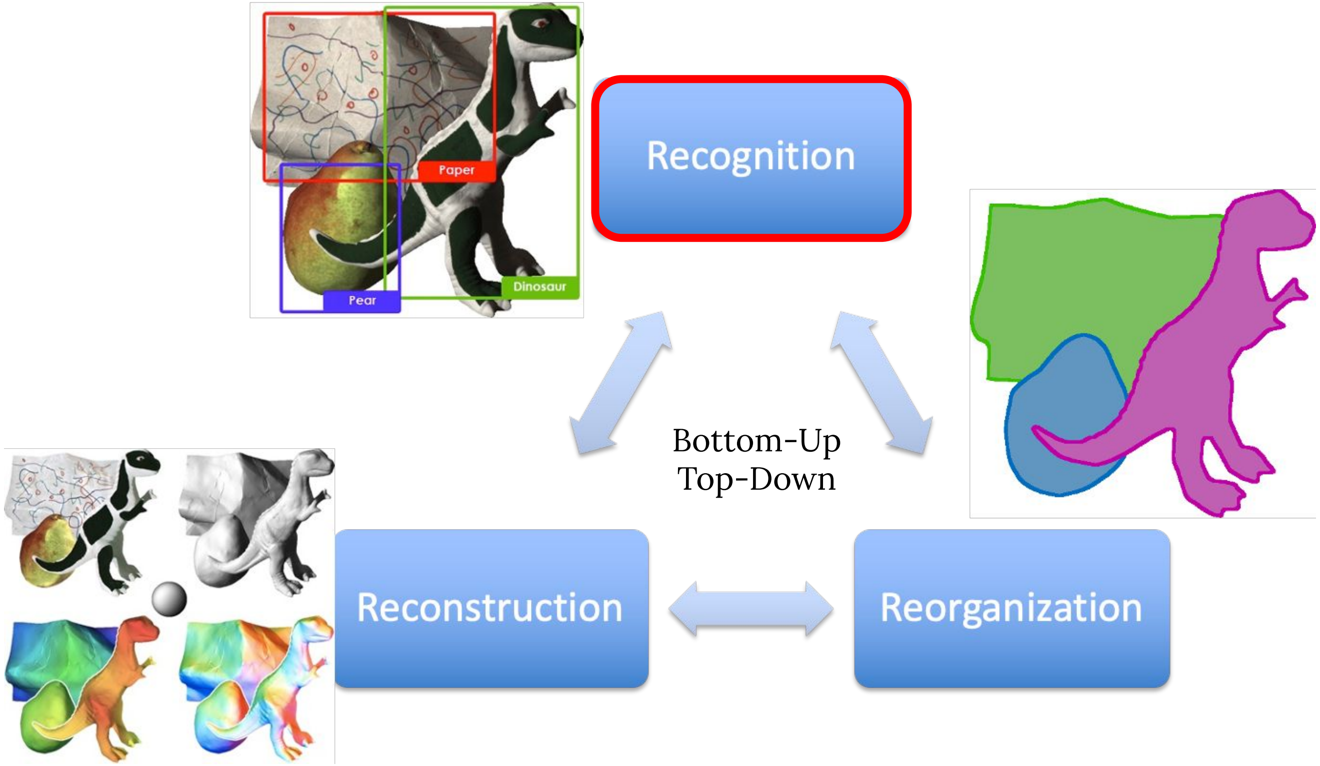
Favorite photo you’ve taken

What is computer vision?



Sources: https://inst.eecs.berkeley.edu/~cs280/sp18/, https://www.sciencedirect.com/science/article/abs/pii/S0167865516000313

What is computer vision?



Sources: https://inst.eecs.berkeley.edu/~cs280/sp18/, https://www.sciencedirect.com/science/article/abs/pii/S0167865516000313

About this course

Cameras
& Imaging

Image
Manipulation

Modeling
& Classification

About this course

- 1: Perspective Projection
- 2: Image Formation
- 3: Image Artifacts

Cameras & Imaging

What can we infer from an image?

What information is lost?

- 4: Convolution
- 5: Fourier
- 6: Pyramids, Edges, and Features
- 7: Image Analysis

Image Manipulation

How do we describe the content of an image?

What other ways can we represent an image? Why?

- 8: Least-Squares
- 9: Total and Iterative Least-Squares
- 10: Clustering
- 11: Dimensionality Reduction
- 12: Linear Classifiers
- 13: Nonlinear Classifiers

Modeling & Classification

Can we model content from groups of images?

What makes a good model?

How do we choose an error function?

How do we choose an optimization method?

What is an image classifier and how does it work?

Prerequisites

Linear Algebra – please review the MIDS 1A Linear Algebra course in preparation for this course

Programming – we will use Python for all examples, exercises, and assignments

DATASCI 207 – familiarity with machine learning techniques will be helpful but not required

Class Structure

Live sessions (90 min)

- Big picture overview before each unit
- Bridge between theory and application (exercises)
- Broader real-world applications

Async lectures (90 min)

- Mathematical derivations
- Intuition about CV concepts and methods

Assignments (7 + Final Project)

- Demonstration of understanding

Office Hours (60 min / Section)

Participation

Goal is 100% participation every class
There are no stupid questions and no stupid answers (really truly)
Public discussion on Piazza/Slack is strongly encouraged

Grading

10% Per Assignment (70% total)
30% Final Project

Assignment Rubric

100%	Good notebook and matched outputs
90%	Good effort and conceptual understanding but key problem(s) with the deliverable
70%	Significant effort but lacks conceptual understanding and/or code has significant problems
0%	Incomplete

Absolute scale, no curve

Each section of multi-part assignments is graded separately

5 Total late days per semester, prior approval required

Online Platforms

Piazza

Sign-up Link – http://piazza.com/berkeley/spring2023/datasci281

Access Code – DATASCI_CV_281

TODO: Post your github ID to get access to assignments

Github

Assignments – github classroom links posted on Piazza

Resources – https://github.com/W281/mids-281

Slack

#datasci-281-2023-spring

Week	Topics	Additional Reading	Assignment Due
1	Perspective Projection	Camera Obscura	
2	Image Formation	Pointwise Operation	
3	Image Artifacts	JPEG	
4	Convolution	Linear Systems, Linear Time Invariant Systems	Assignment 1
5	Fourier	Fourier, Sampling	Assignment 2

Office Hours

Section 1

Rachel Brown — Tuesdays 3 pm Pacific

Sections 1 & 3

Allen Yang — By Appointment

TA

Albert Jiang — Sunday morning TBD

Administrative details

Collaboration Policy & Academic Integrity

We encourage studying in groups of two to four people. This applies to working on homework, discussing labs and projects, and studying for the exam. However, students must always adhere to the UC Berkeley Code of Conduct (http://sa.berkeley.edu/code-of-conduct) and the UC Berkeley Honor Code (https://teaching.berkeley.edu/berkeley-honor-code).

All materials that are turned in for credit or evaluation must be written solely by the submitting student or group.

Administrative details

Disability Services & Accommodations

If you need disability-related accommodations in this class, if you have emergency medical information you wish to share with me, or if you need special arrangements, please inform me as soon as possible.

Administrative details

Publishing your work

You are highly encouraged to use your program coursework to build an academic/professional portfolio

- Blog about your coursework (and other ideas) and share on the I School Medium
- Publish projects to your I School project portfolio gallery (more than just for capstone)
- Publish your work on LinkedIn and tag the @UC Berkeley School of Information.
- Publish in academic journals -- contact your professors for assistance.
- Publish your news (e.g., conference talks, award, scholarships) to the I School internal newsletter

About this course

1: Perspective Projection
2: Image Formation
3: Image Artifacts
4: Convolution
5: Fourier
6: Pyramids, Edges, and Features
7: Image Analysis
8: Least-Squares
9: Total and Iterative Least-Squares
10: Clustering
11: Dimensionality Reduction
12: Linear Classifiers
13: Nonlinear Classifiers

Cameras & Imaging

What can we infer from an image?
What information is lost?

Image Manipulation

How do we describe the content of an image?
What other ways can we represent an image? Why?

Modeling & Classification

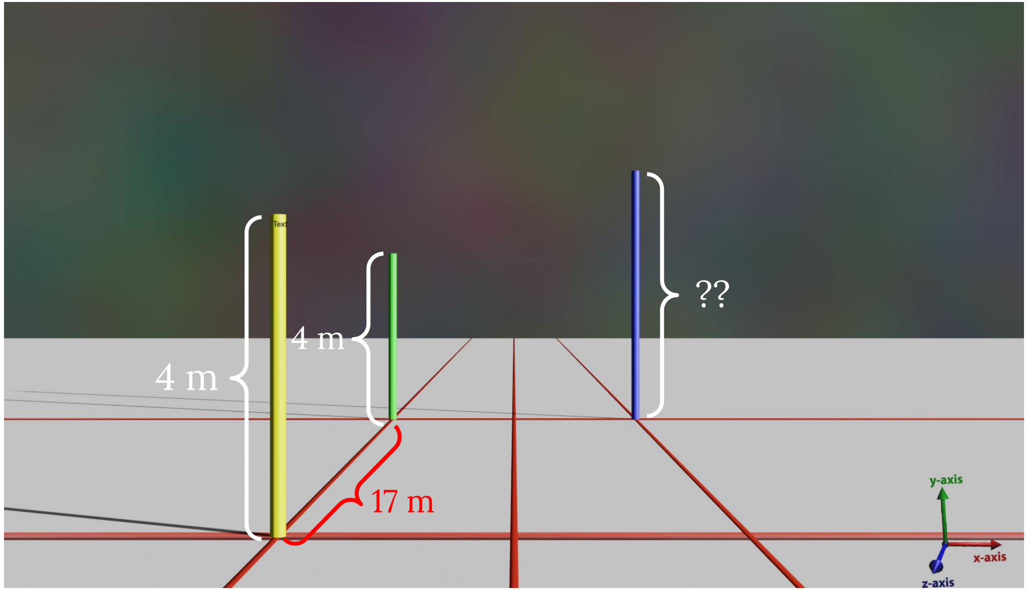
Can we model content from groups of images?
What makes a good model?
How do we choose an error function?
How do we choose an optimization method?
What is an image classifier and how does it work?

Assignment 1 — Perspective Projection

https://classroom.github.com/a/2hRaqODV

Assignment 1 — Perspective Projection

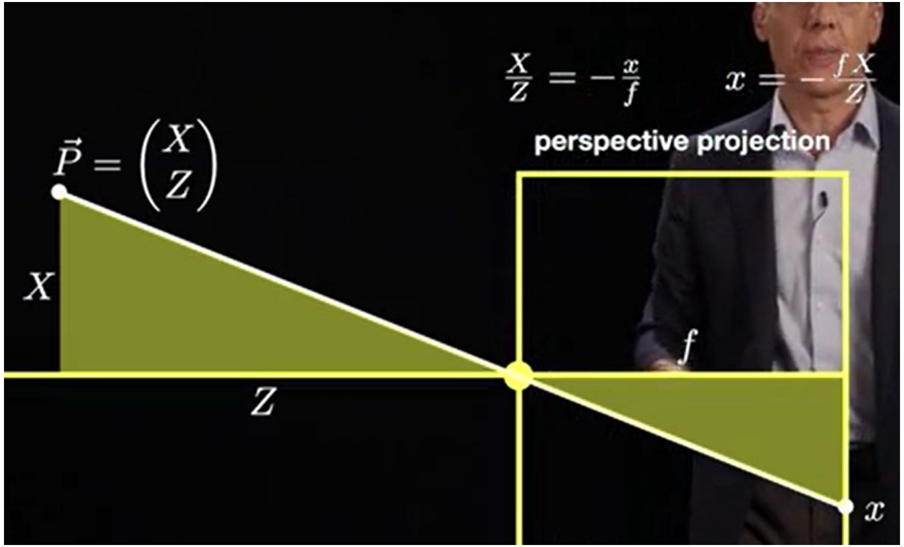
Part 1 — Given objects of known size in an image, infer the size of an unknown object based on its relative position in the scene



Hints for Assignment 1

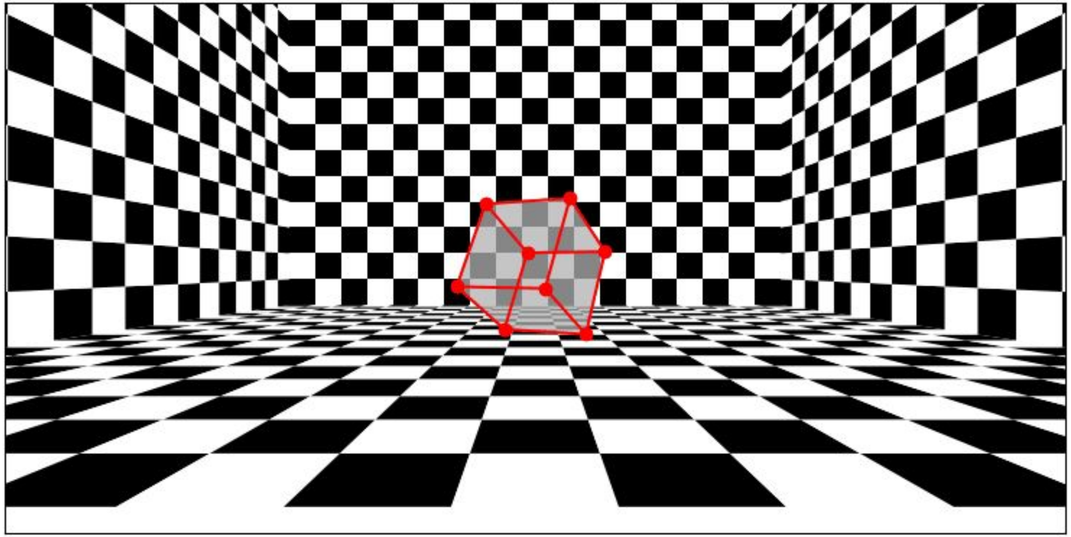
Draw this side view diagram of the problem

Label the knowns and unknowns



Assignment 1 — Perspective Projection

Part 2 — Learn how to transform point locations between world coordinates, object coordinates, and camera coordinates; use this to re-create the Dolly Zoom effect



Hints for Assignment 1

Homogeneous coordinates allow us to do rotation & translation in one step

‘s’ is for ‘sensor’

‘f’ is for ‘focal length’

‘c’ is for ‘camera’

‘r’ is for ‘rotation’

‘t’ is for ‘translation’

‘w’ is for ‘world’

$$\begin{pmatrix} x_s \\ y_s \\ s \end{pmatrix} = \cancel{\lambda} \overset{\text{intrinsic matrix (camera)}}{\begin{pmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{pmatrix}} \overset{\text{extrinsic matrix (world)}}{\begin{pmatrix} r_{11} & r_{21} & r_{31} & t_X \\ r_{12} & r_{22} & r_{32} & t_Y \\ r_{13} & r_{23} & r_{33} & t_Z \end{pmatrix}} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

homogeneous coordinates

Error in Unit 1 Async Lecture

Second exercise (3D projection), the rotation matrix in the problem description has rotation applied as Rz first, then Ry, and finally Rx. Whereas in the solution the rotation is performed in the reverse order Rx first, then Ry, and last Rz. Rotation in 3D is not commutative i.e. RxRyRz != RzRyRx. Please take this into consideration.

In the generalized projection equation, the lambda term (conversion from real-world units to pixels) is kept outside the intrinsic matrix. However, this term is multiplied only with the focal length of the intrinsic matrix. Therefore, the correct equation is:

$$\begin{pmatrix} x_s \\ s \end{pmatrix} = \begin{pmatrix} \lambda f & c_x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(\theta) & -\sin(\theta) & t_X \\ \sin(\theta) & \cos(\theta) & t_Z \end{pmatrix} \begin{pmatrix} X_w \\ Z_w \\ 1 \end{pmatrix}$$

Upcoming ToDo's

Sign up for Piazza

Post Github ID on Piazza

Watch async lectures for Unit 1

Accept Assignment 1 on GitHub
(due Monday, Jan 30th)