Plan for today

- Topic overview
- Introductions
- Course overview:
  - Logistics and requirements

What is Computer Vision?
Computer Vision

*Enable machines to “see” the visual world as we do*


Computer Vision

- Automatic understanding of images and video
  1. Computing properties of the 3D world from visual data *(measurement)*


1. Vision for measurement

  - Real-time stereo
  - Structure from motion
  - Tracking

Slide credit: Kristen Grauman
Computer Vision

• Automatic understanding of images and video
  1. Computing properties of the 3D world from visual data (measurement)
  2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities (perception and interpretation)

2. Vision for perception, interpretation

3. Algorithms to mine, search, and interact with visual data (search and organization)
3. Visual search, organization

- Query
- Image or video archives
- Relevant content

Related disciplines

- Artificial intelligence
- Machine learning
- Cognitive science
- Computer vision
- Algorithms
- Graphics
- Image processing

Vision and graphics

- Images
- Vision
- Model
- Graphics
- Inverse problems: analysis and synthesis
Why is vision difficult?

What humans see

What computers see

Slide credit: Larry Zitnick
Why is vision difficult?

• Ill-posed problem: real world much more complex than what we can measure in images – 3D → 2D
• Impossible to literally “invert” image formation process

Challenges: ambiguity

• Many different 3D scenes could have given rise to a particular 2D picture

Challenges: many nuisance parameters

Illumination  
Object pose  
Clutter  
Occlusions  
Intra-class appearance  
Viewpoint

Slide credit: Kristen Grauman
Challenges: scale

Challenges: Motion

Challenges: occlusion, clutter
Challenges: object intra-class variation

Challenges: context and human experience

Fei Fei Li, Rob Fergus, Antonio Torralba
Challenges: context and human experience

Challenges: complexity

How many object categories are there?

-10,000 to 30,000

Slide credit: Fei-Fei, Fergus, Torralba

Challenges: complexity

<table>
<thead>
<tr>
<th>Service</th>
<th>Images</th>
<th>Uploads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flickr</td>
<td>10 billion images</td>
<td>100 hours</td>
</tr>
<tr>
<td>Facebook</td>
<td>250 billion images</td>
<td></td>
</tr>
<tr>
<td>Imgur</td>
<td>1 billion images</td>
<td></td>
</tr>
<tr>
<td>YouTube</td>
<td>1 billion images served daily</td>
<td></td>
</tr>
<tr>
<td>Photobucket</td>
<td>10 billion images</td>
<td></td>
</tr>
</tbody>
</table>

From Cisco: Almost 90% of web traffic is visual!
Challenges: complexity

- Thousands to millions of pixels in an image
- 30+ degrees of freedom in the pose of articulated objects (humans)
- About half of the cerebral cortex in primates is devoted to processing visual information [Felleman and van Essen 1991]

What works well today?

Optical character recognition (OCR)

- Digit recognition
- License plate readers
- Sudoku grabber

Source: S. Seitz, N. Snavely
Image classification and object detection

Biometrics

Face detection

Source: S. Seitz
Face detection for privacy protection

Technology gone wild...

Face recognition
Face anonymization

Zhongzheng Ren, Yong Jae Lee, and Michael Ryoo.

Interactive systems

Shotton et al.

Instance recognition

Google Goggles
Use pictures to search the web.

Slide credit: Devi Parikh
Pedestrian detection

Autonomous agents

3D reconstruction from photo collections

Slide credit: Devi Parikh

Slide credit: Svetlana Lazebnik

Q. Shan, R. Adams, B. Curless, Y. Furukawa, and S. Seitz, The Visual Turing Test for Scene Reconstruction, 3DV 2013
Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Source: S. Seitz

Special effects: motion capture

Pirates of the Caribbean, Industrial Light and Magic

Source: S. Seitz

Medical imaging

3D imaging
MRT, CT

Image guided surgery
Grimson et al., MIT

Source: S. Seitz
Visual data in 1963


Visual data today

Understand and organize and index all this data!!

Why vision?

- As image sources multiply, so do applications
  - Relieve humans of boring, easy tasks
  - Enhance human abilities
  - Advance human-computer interaction, visualization
  - Perception for robotics / autonomous agents
  - Organize and give access to visual content
Applications

- Law enforcement / Surveillance
- Robotics
- Autonomous driving
- Medical imaging
- Photo organization
- Image search
- E-commerce
- ... cell phone cameras, social media, Google Glass, etc.

Summary

- Computer Vision is useful, interesting, and difficult
- A growing and exciting field
- Lots of cool and important applications
- New teams in existing companies, startups, etc.

Introductions

- Instructor
  - Yong Jae Lee
  - yongjaelee@ucdavis.edu
  - Assistant Professor in CS, UC Davis since July 2014

  - Research areas: Computer Vision and machine learning
    - Visual Recognition
    - Graphics Applications
Introductions

- TAs:
  - Chongruo Wu
  - crwu@ucdavis.edu
  - PhD student in CS
  - Maheen Rashid
  - mhnrashid@ucdavis.edu
  - PhD student in CS

This course

- ECS 174 (4-units)
- Lecture: Tues & Thurs 12:10-1:30 pm, Giedt Hall 1002
- Discussion section: Mon 5:10-6pm, Chem 179 (location may change)
- Office hours: Academic Surge 1044/2075
  - Maheen: Tues 10 am-noon (AS 1044)
  - Chongruo: Thurs 3-5 pm (AS 1044)
  - Yong Jae: Fri 3-5 pm (AS 2075)

This course

- Course webpage
  https://sites.google.com/a/ucdavis.edu/ecs-174-computer-vision---spring-2018/
- Canvas (assignment submission, grades)
  https://canvas.ucdavis.edu/courses/225624
- Piazza
  piazza.com/uc_davis/spring2018/ecs174/
Goals of this course

• Introduction to primary topics in Computer Vision
• Basics and fundamentals
• Practical experience through assignments
• Views of computer vision as a research area

Prerequisites

• Upper-division undergrad course

• Basic knowledge of probability and linear algebra
• Data structures, algorithms
• Programming experience

• Experience with image processing or Matlab will help but is not necessary

Topics overview

• Features and filters
• Grouping and fitting
• Recognition and learning

Focus is on algorithms, rather than specific systems
**Features and filters**

Transforming and describing images; textures, edges

Slide credit: Kristen Grauman

**Grouping and fitting**

Clustering, segmentation, fitting; what parts belong together?

Slide credit: Kristen Grauman

**Recognition and learning**

Recognizing objects and categories, learning techniques

Slide credit: Kristen Grauman
Recognition and learning

Deep learning

Not covered: Multiple views and motion

Multi-view geometry, stereo vision

Not covered: Video processing

Tracking objects, video analysis, low level motion, optical flow
Textbooks

By Rick Szeliski
http://szeliski.org/Book/

By Kristen Grauman, Bastian Leibe
Visual Object Recognition

Requirements / Grading

• Problem sets (60%)

• Final exam (35%)
  – comprehensive (cover all topics learned in class)

• Class and Piazza participation, including attendance (5%)
  – Piazza: participation points for posting (sensible) questions and answers

Problem sets

• Some short answer concept questions
• Matlab programming problems
  – Implementation
  – Explanation, results
• Follow instructions; points will be deducted if we can’t run your code out of the box
• Ask questions on Piazza first
• Submit to Canvas
• The assignments will take significant time to do
• Start early

• TA will go over problem set during discussion sections after release (others will be used as extra office hours)
Matlab

- Built-in toolboxes for low-level image processing, visualization
- Compact programs
- Intuitive interactive debugging
- Widely used in engineering

Matlab

- CSIF labs 67, 71, 75 (pc1-pc60)
- Academic Surge 1044 and 1116
- Lab schedule (reservations) and remote access info found on class website

- Matlab available for free from campus software site

Problem Set 0

- Matlab warmup
- Basic image manipulation
- Out Thursday, due 4/13
Problem set deadlines

- Problem sets due 11:59 PM
  - Follow submission instructions given in assignment
  - Submit to Canvas; no hard copy submissions
  - Deadlines are firm. We’ll use Canvas timestamp. Even 1 minute late is late.

- Late submissions: 1 point deduction for every hour after the deadline up to 72 hours; after 72 hours, you will receive a 0

- If your program doesn’t work, clean up the code, comment it well, explain what you have, and still submit. Draw our attention to this in your answer sheet.

Collaboration policy

- All responses and code must be written individually or in pairs (a group of 2)

- Students submitting answers or code found to be identical or substantially similar (due to inappropriate collaboration) risk failing the course
  - We will be using MOSS to check for cheating!
  - Copying online solutions also counts as cheating!
  - Please don’t cheat… you are going to get caught!

- Read and follow UC Davis code of conduct

MOSS
Miscellaneous

• Check class website regularly for assignment files, notes, announcements, etc.
• Come to lecture on time
• No laptops, phones, tablets, etc. in class please
• Please interrupt with questions at any time

Coming up

• Read the class webpage carefully
• Next class (Thurs): lecture on linear filters
• PS0 out Thursday, due 4/13

Questions?

See you Thursday!