# 16-820
# Computer Vision
# Fall 2024

Matthew O’Toole

# Welcome to the First Day!



# Am I in the right course?

- 16-820 is an introduction to computer vision for graduate students
  - Undergrad version is 16-385
- 16-820 and 16-720 are both introductory computer vision courses
  - 16-720 is for students without much prior exposure to computer vision
  - 16-820 is similar in content to 16-720 but moves a bit faster, covers a little bit extra material
- Not recommended to take both classes

# Outline

- **Introductions**
- What is this class about?
- Logistics for class (homeworks, grading)

# Instructor Introduction: Matthew O'Toole

computational imaging

**Computational Cameras**

optics

+


sensing

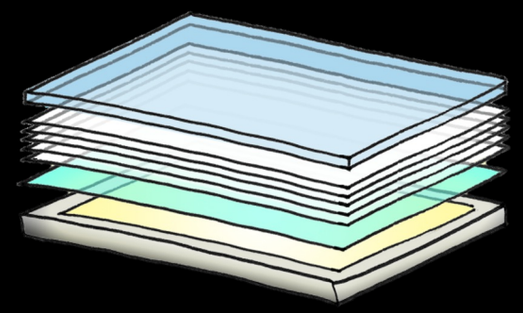
+

computation

**Computational Displays**

computation

+



optics & electronics

+

human visual system

# Instructor Introduction: Matthew O'Toole

## computational imaging

**Computational Cameras**

HDR Imaging [Debevec, Nayar, …]

Super-resolution [Baker, …]

Light Fields [Levoy, …]

**Computational Displays**

HDR Display [Seetzen, …]


Super-resolution [Hirsch, Heide, …]

Light Fields [Wetzstein, …]

# Nikhil Keetha

**Program: 2nd Year PhD in Robotics**

**Background:**

- **Undergrad in Physics at IIT (ISM) Dhanbad, India**
- **Masters in Robotics, CMU**

**Research Interests:**

- **Generalizable 3D Perception for Field Robotics**
- **SfM/SLAM, Localization & Scene Understanding**
- **Self-Supervised Learning**

**Hobbies: Cooking, Movies, Anime, Gaming & Hiking**

# Ayush Jain

- 2nd year PhD with Prof. Katerina Fragkiadaki
- Interested in scaling up 3D models with the help of 2D and video data
- Love playing badminton, cricket, (more recently) squash and reading books

https://ayushjain1144.github.io/

# Yuyao Shi

- 2nd year PhD with Prof. Jessica Hodgins
- Background:
  - Undergrad in mechanical engineering in SUTD (Singapore)
  - Master in Robotics at CMU
- Research interest: character animation and human motion

- Hobbies: video games, ping-pong, and anime

# About you!

Departments?

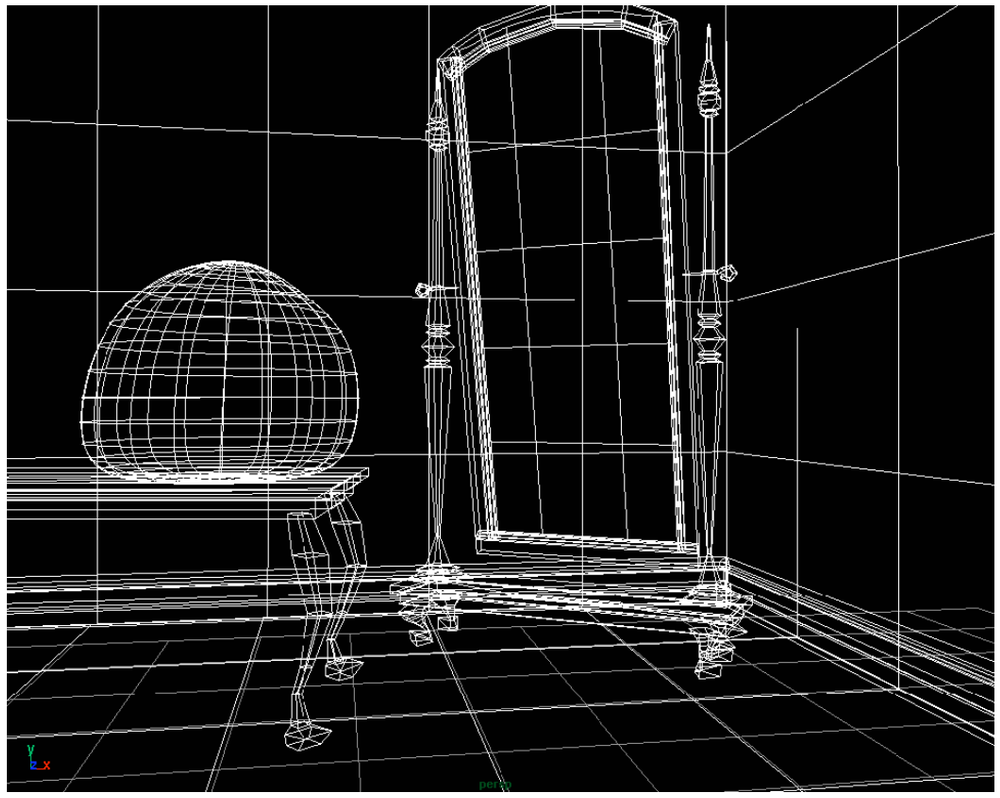
Years?

Related classes?
(computer graphics, machine learning,…)

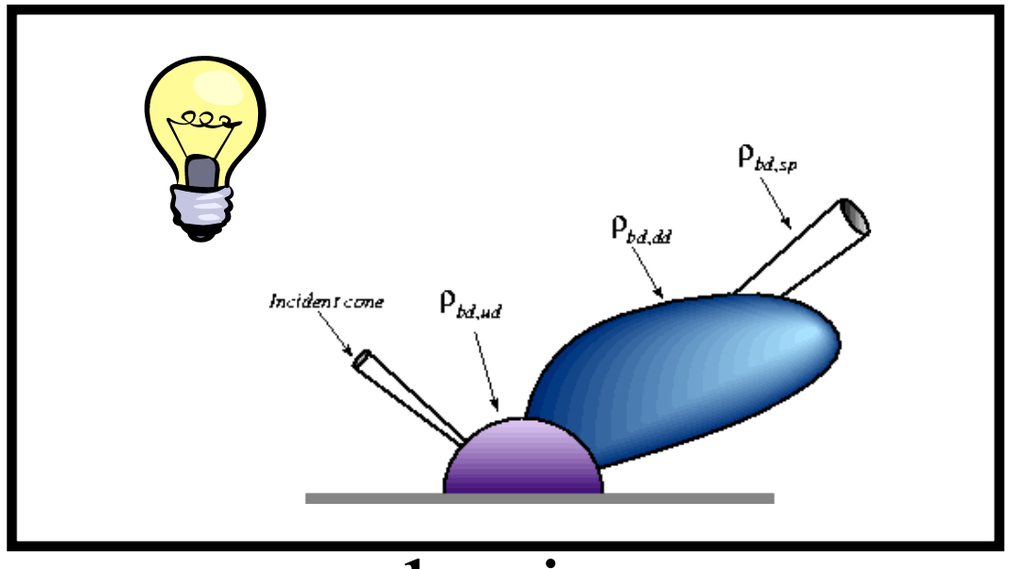
# Outline

- Introductions
- **What is this class about?**
- Logistics for class (homeworks, grading)

# Related field: computer graphics



3D geometry

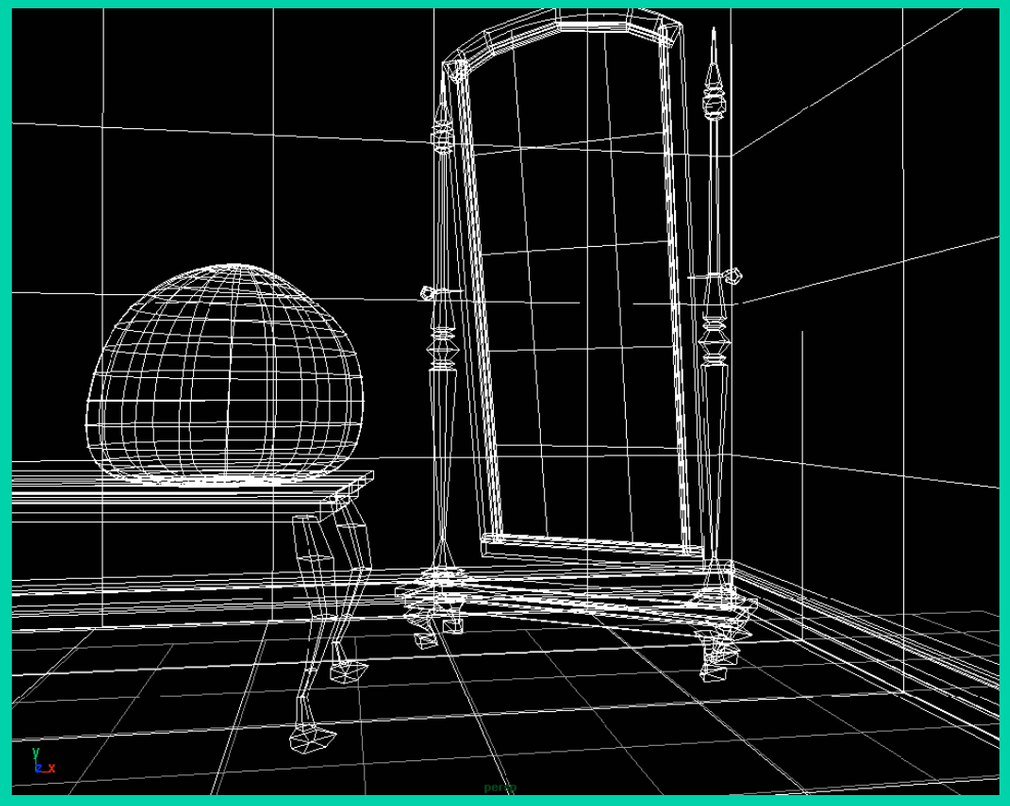



physics

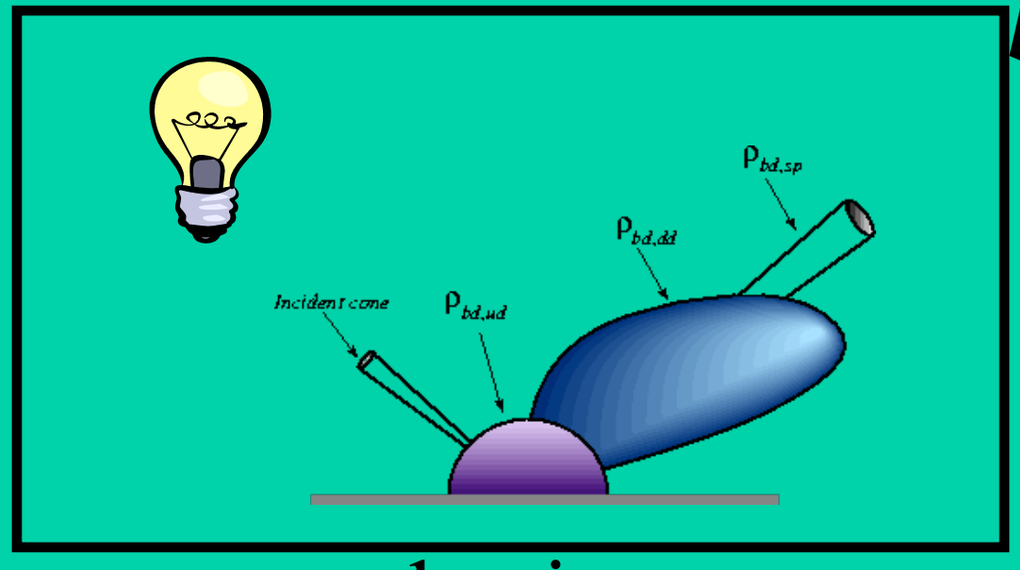
projection

Simulation

# What is computer vision?



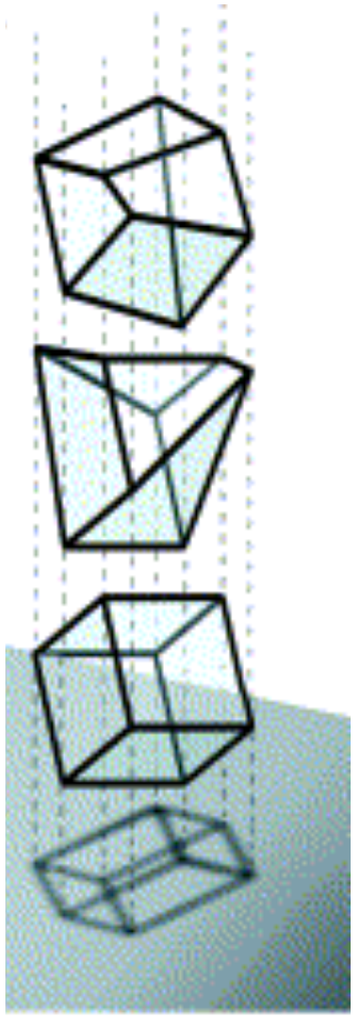
3D geometry




physics

Estimation

# Challenge: ill-posed problem







For any 2D image, there are an infinite number of 3D worlds that could have produced that image

Certain 3D interpretations are more *probable* than others

We want to build models that capture such prior knowledge about the word

Why do we want to understand the world from images (or videos)?

Why computer vision?

# Visual media is everywhere




8 years worth of video is uploaded to YouTube
250 million photos are uploaded to FaceBook

... each day

# Distributed sensors


Everyone is carrying a camera in their back pocket

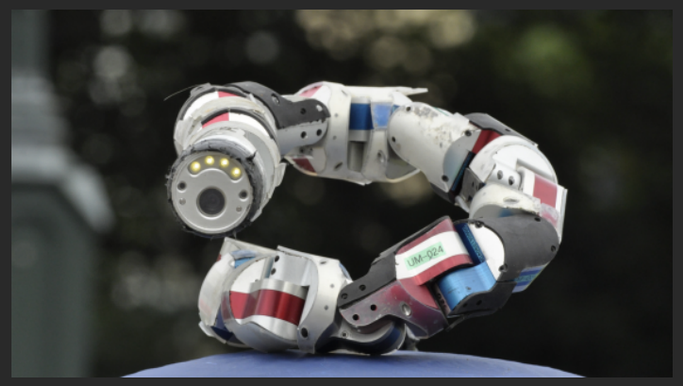
# Applications: Reconstructing the (4D) world

Photo Tourism (UWashington/Microsoft)

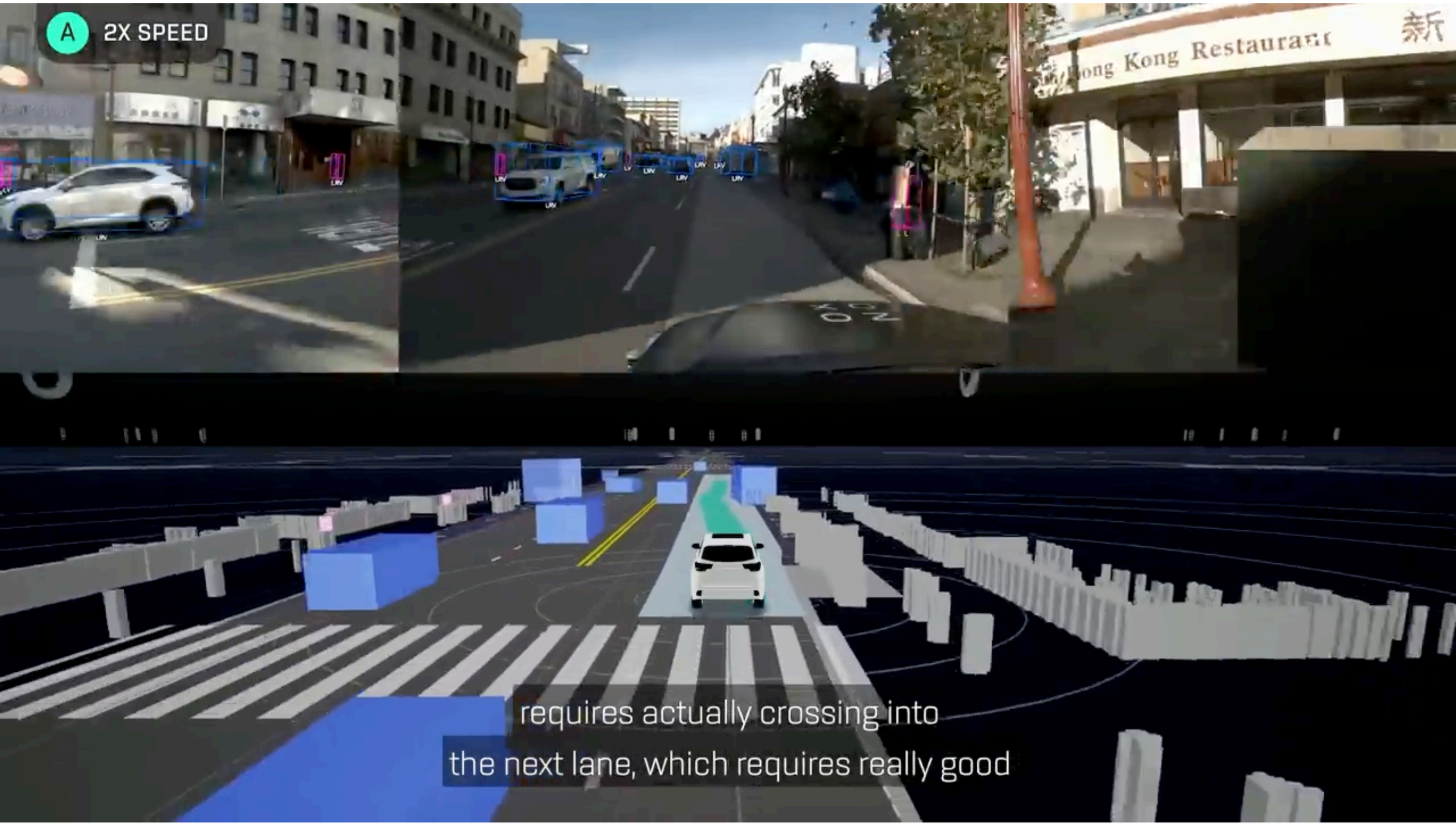
# Applications: assistive technology


CMU Quality of Life Center

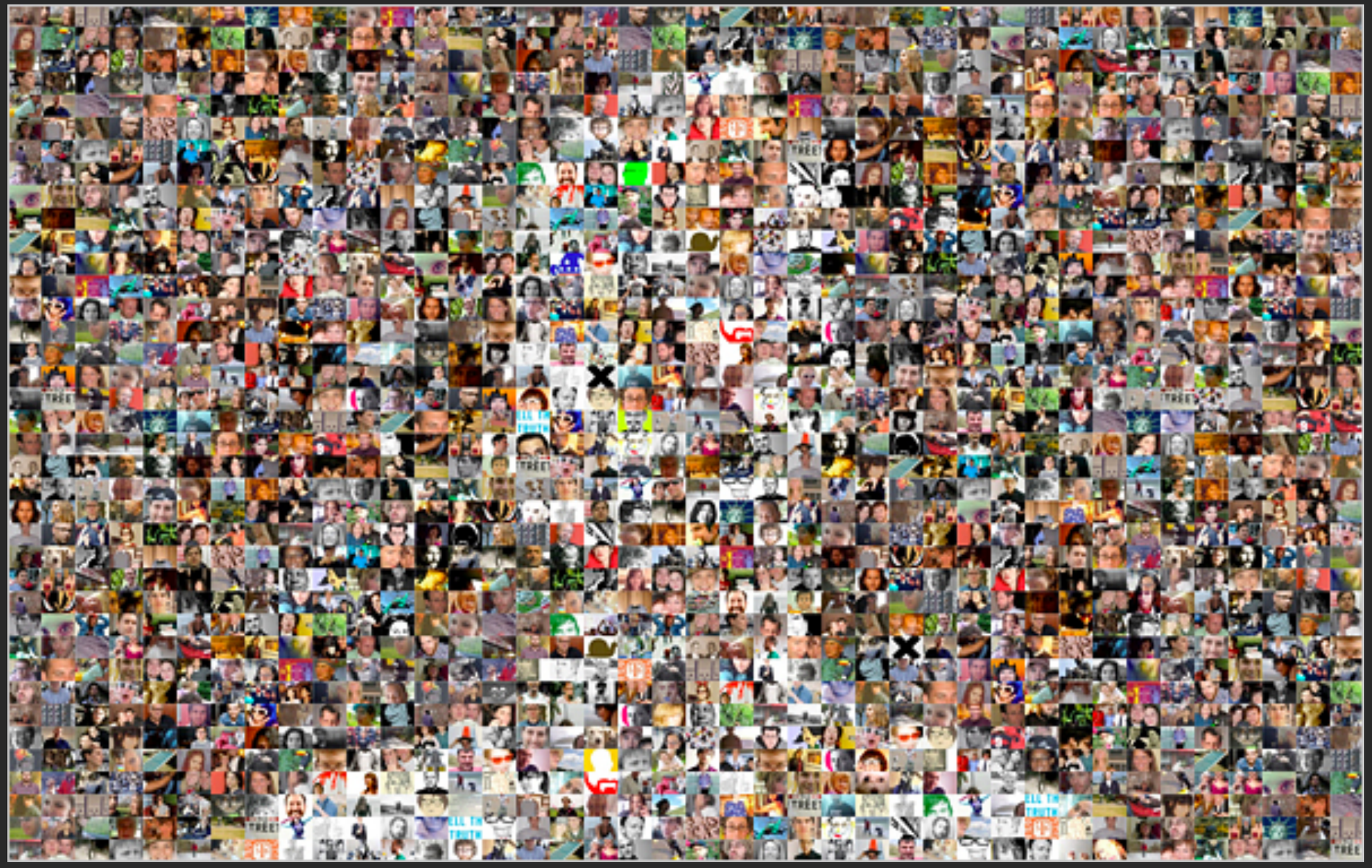
# Applications: Robotics



# Applications: Self-driving



# Applications: image search



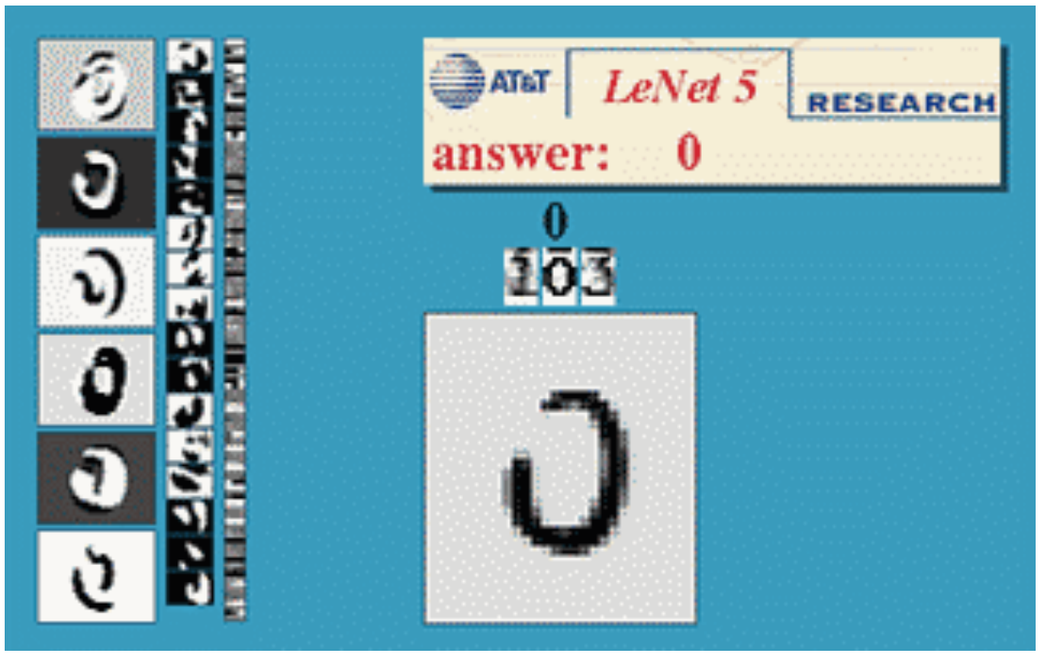
# Applications: surveillance

“The work was painstaking and mind-numbing: One agent watched the same segment of video 400 times.The goal was to construct a timeline of images, following possible suspects as they moved along the sidewalks, building a narrative out of a random jumble of pictures from thousands of different phones and cameras. It took a couple of days, but analysts began to focus on two men in baseball caps who had brought heavy black bags into the crowd near the marathon’s finish line but left without those bags.”

Washington Post

# Applications: Optical character

Technology to convert scanned docs to text

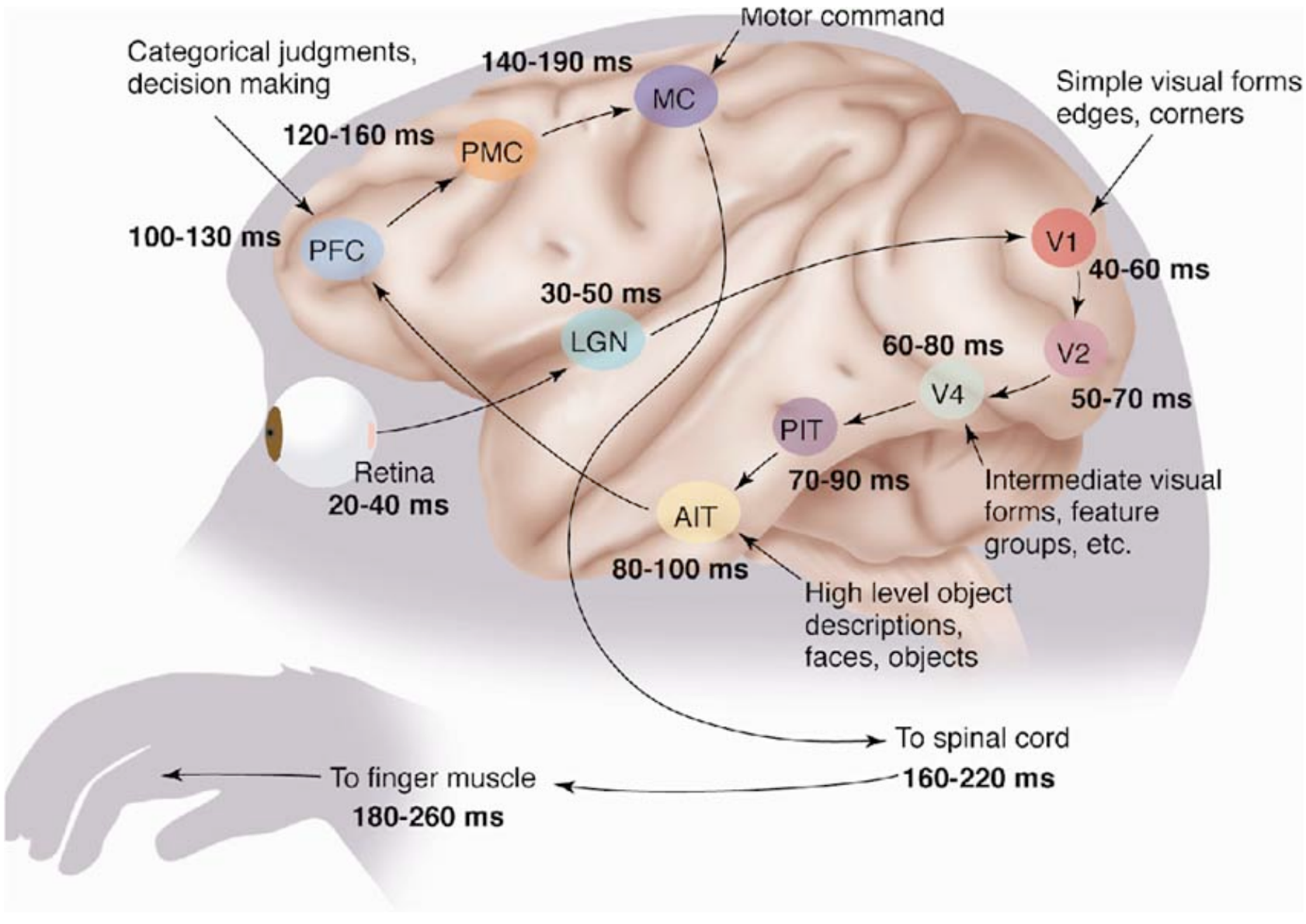



Digit recognition, AT&T labs
http://www.research.att.com/~yann/


License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

# Science: Computer vision vs Human Vision

About 1/3 of brain is devoted to visual processing




Computational understanding of human perception

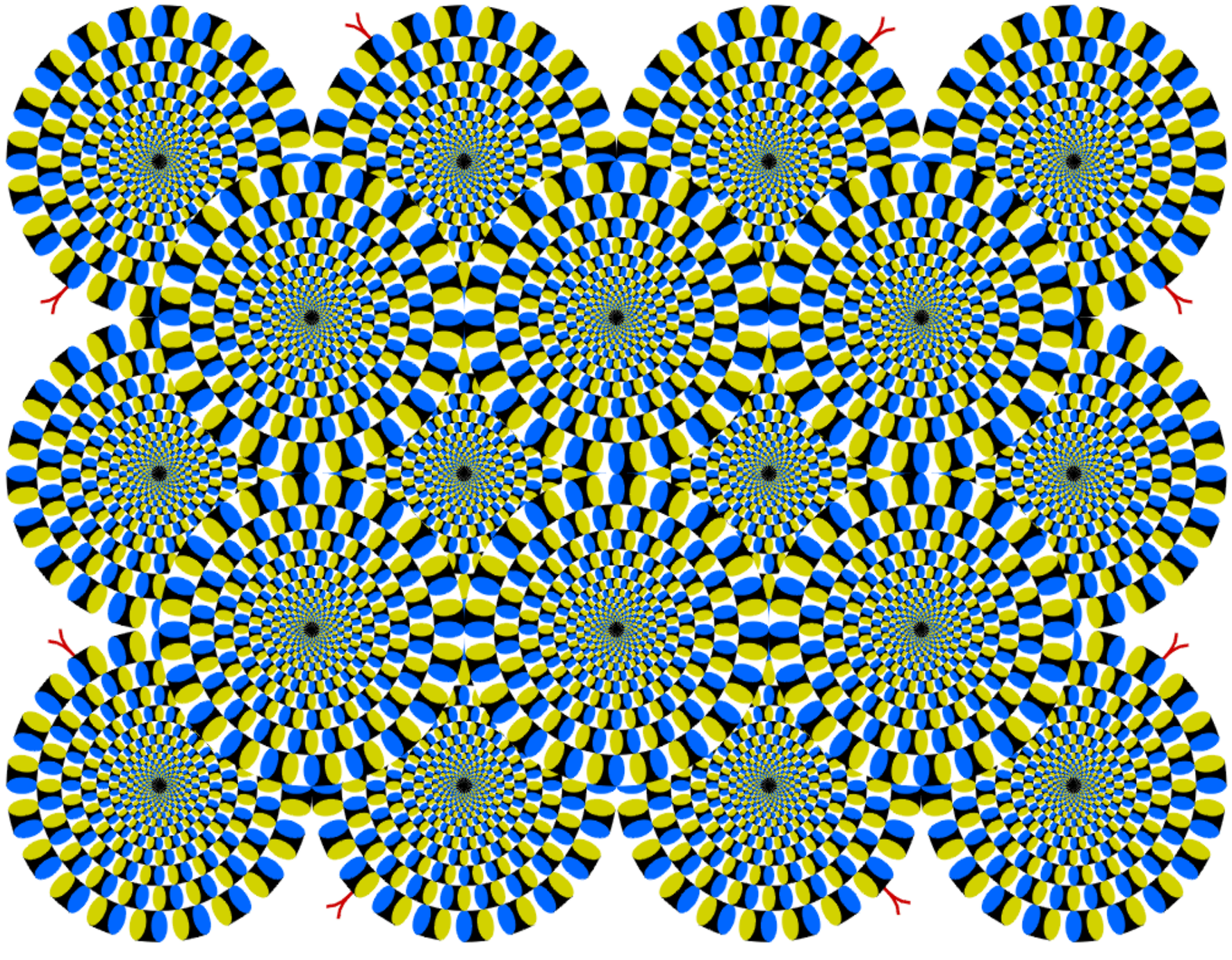
# Humans make strong use of contextual cues

The human visual system exploits the ecological regularities of the world

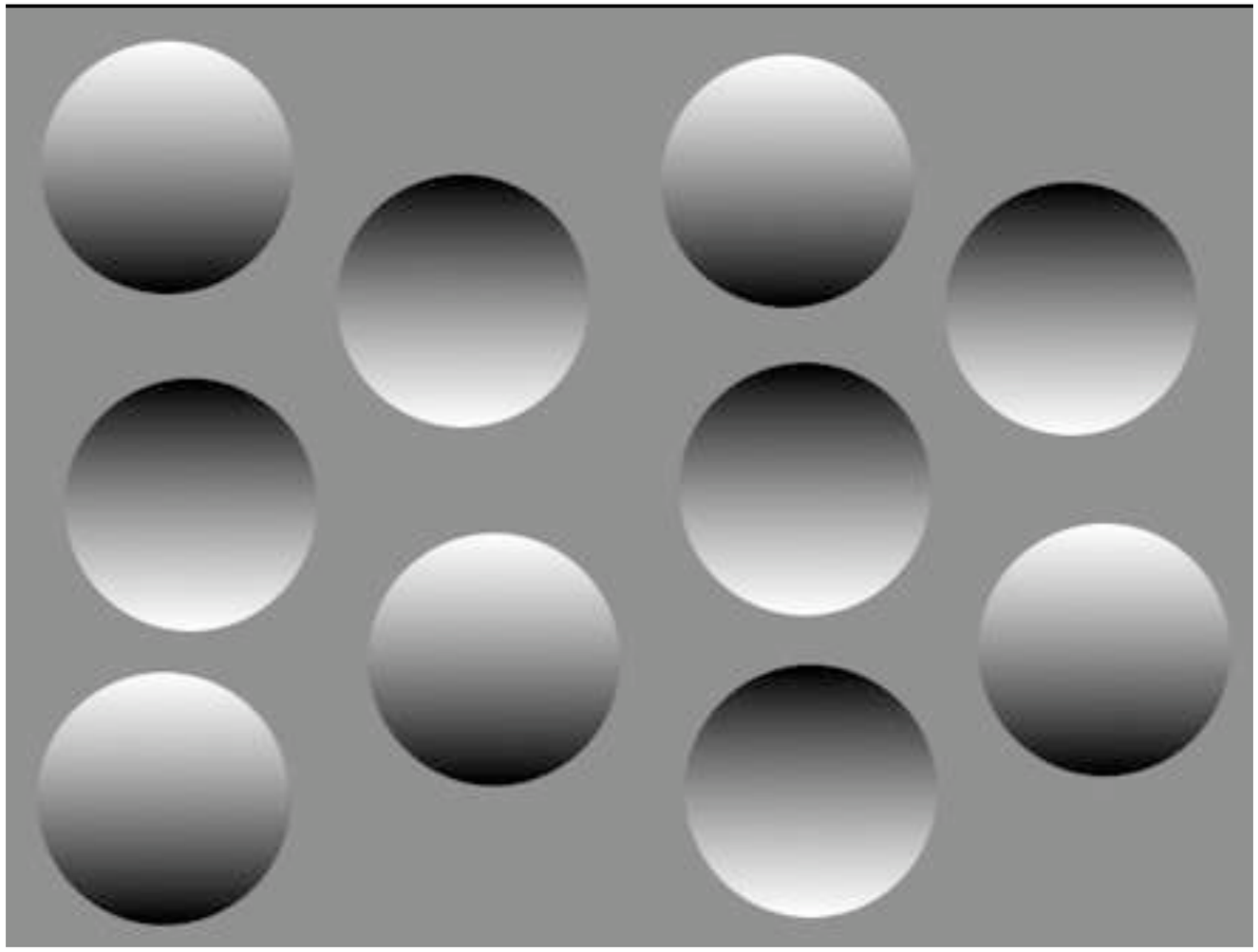
Optical illusions are violations of this regularity

# Illusory motion

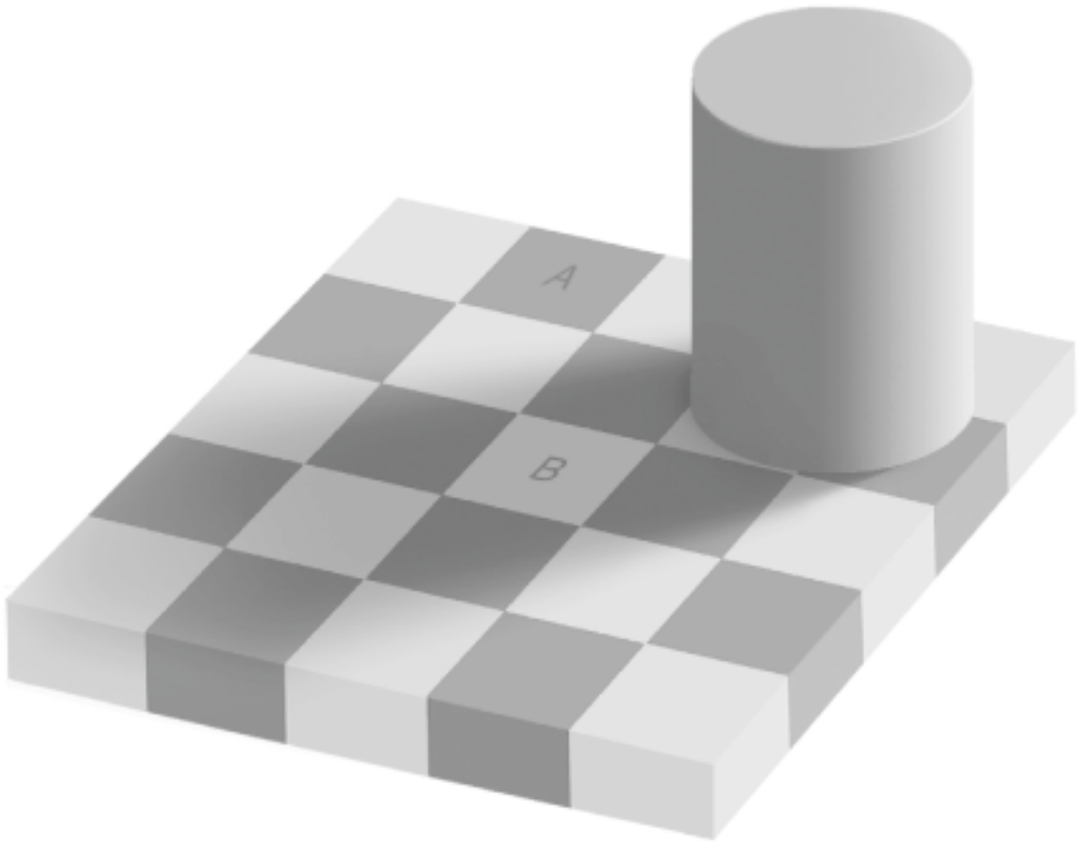




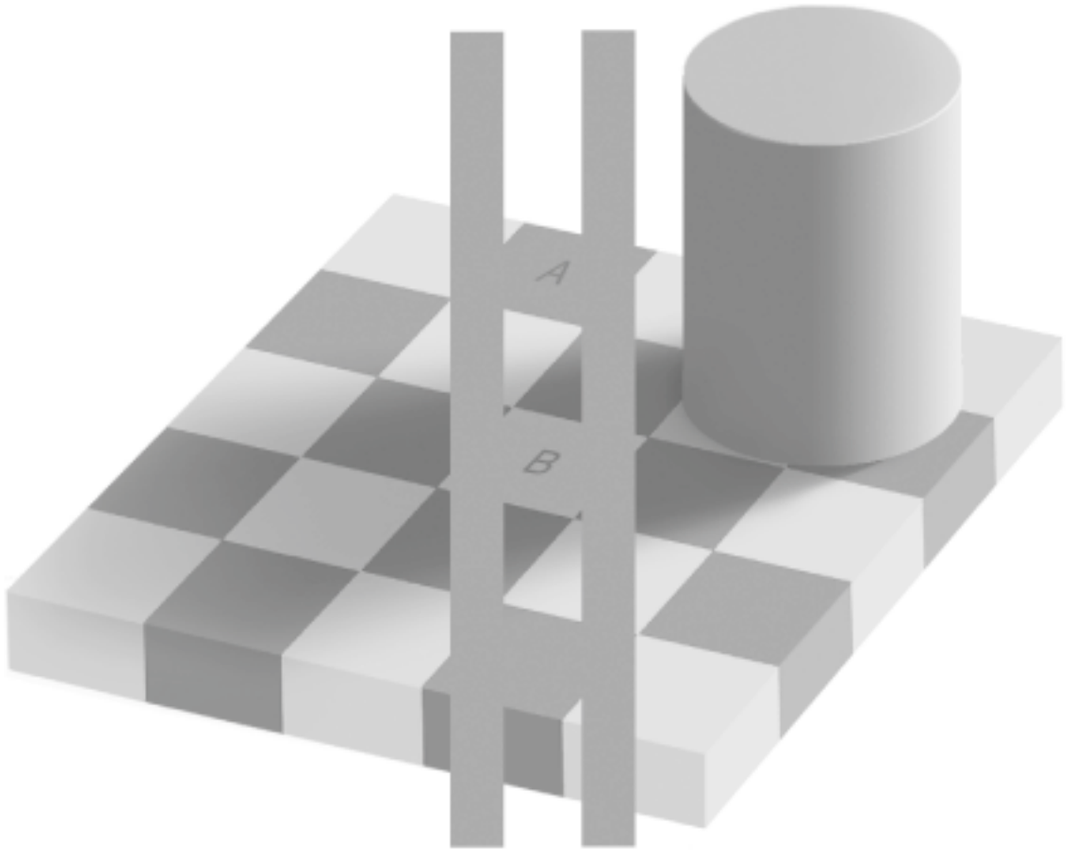
# Shading



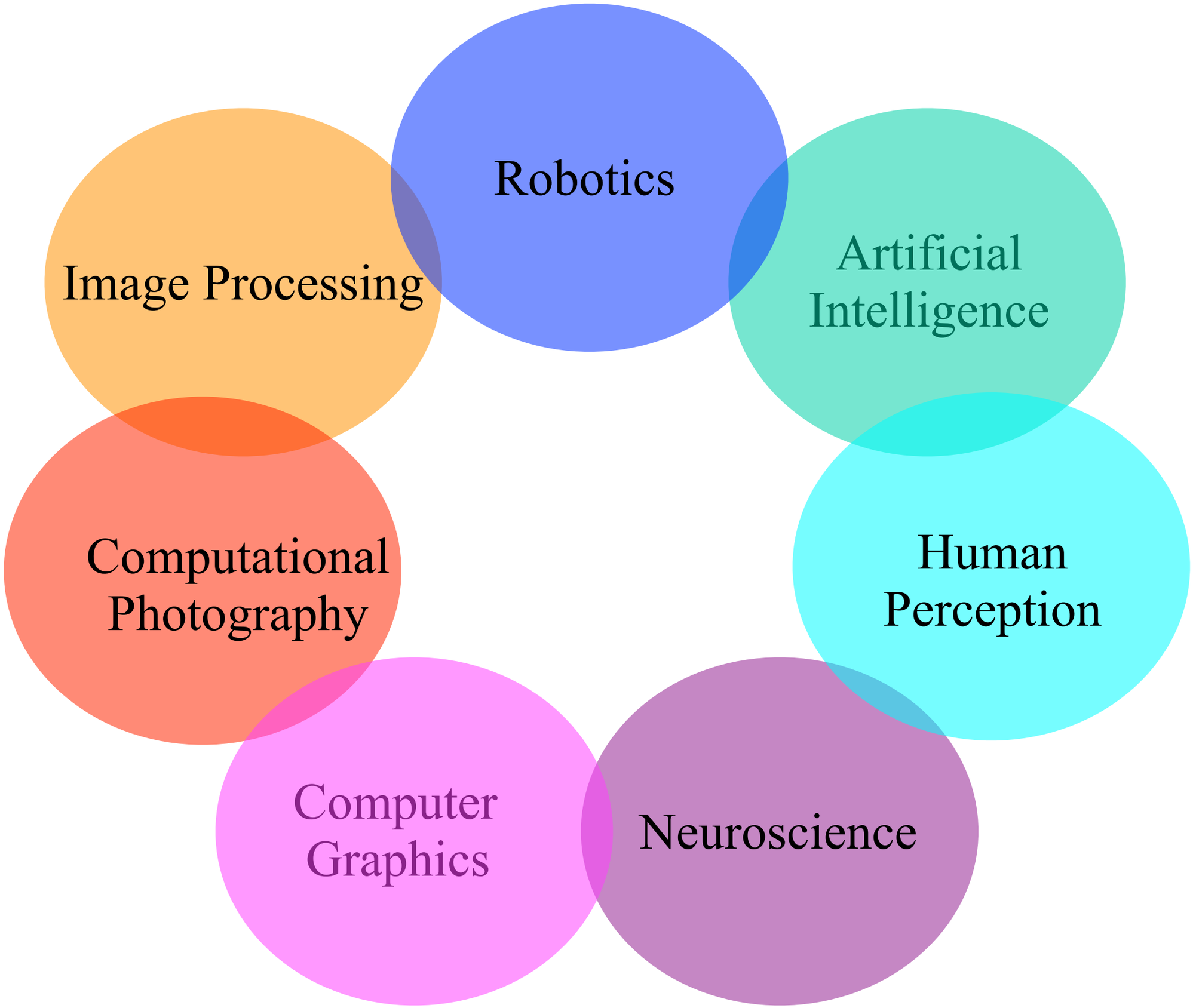
# Understanding shadows



# Understanding shadows



# Relation to Other fields



# Outline

- Introductions
- What is this class about?
- **Logistics for class (homeworks, grading)**

# Canvas

https://canvas.cmu.edu/courses/43480/

Class webpage: https://16820advancedcv.github.io/index.html

**16820 Course Description**

This course introduces the fundamental techniques used in computer vision, which is the analysis of patterns in visual images to understand the objects and scenes that generated them. Topics covered include image formation and representation, camera geometry, and calibration, computational imaging, multi-view geometry, stereo, 3D reconstruction from images, motion analysis, physics-based vision, image segmentation and object recognition.  Homeworks involve Python programming exercises.

This course is modeled off of 16-720, but moving at a bit faster pace.  We will also have a number of guest lectures on cutting-edge research in CV.

**Location and Class Times**

- Time: Monday and Wednesday 12:30 PM - 1:50 PM
- Location: DH [Doherty Hall] 1212

# Waitlist

We have plenty of room in the class at the moment

No one should be on the waitlist

Undergrads should instead take the undergrad version of the course (16-385)

If you are on the waitlist but are not an undergrad, please post on Piazza
to let me know your status and I will look into it.

# Pre-reqs

Knowledge of linear algebra, vector calculus (e.g. gradients), and basic probability are required.

Python programming experience and previous exposure to image processing are desirable, but not required.

…. but your ability to code in Python will be a crucial factor in your success
If you are not familiar with Python, you will need to put in extra effort at the beginning to learning it quickly

# Grading

Grade based on 6 homeworks (with considerable Python implementation)

- HW 1-5 are worth 18% each
- HW 6 (last homework) is worth 10% (it's a bit smaller).

Extra credit (worth up to 3% of your final grade):

- Class participation (Piazza / lecture)
- Organizing study groups

**Extra credit works differently in this course (pay attention!).**

- The formula used for each homework is :
  - Total score obtained / (Max score without extra credit) * 100
  - There is no penalty for not doing the extra credit

- Letter grade cut-offs (approximate, may be adjusted later at the instructor's discretion):
  - Much above 100% - A+
  - Above 95% - A
  - Above 90% - A-
  - Above 80% - B-, B, or B+
  - Above 70% - C-, C, or C+
  - Above 65% - D-, D, D+
  - Less than 65% - Fail

# Tools used by class

**Canvas**: lecture schedule, homework links

**Gradescope**: submission of PDF writeup, code zipfile, and returning of grades

**Piazza**: asking and answering questions about course and homeworks

# HW submission

**Submission Guidelines**

- Homework is submitted through the Gradescope Assignments page.
- Homeworks are due by 11:59pm EST on the given due date.
- You are required to typeset all answers (no handwritten responses) and submit as a PDF
- You must assign individual pages of your PDF to individual questions in gradescope. Not doing so will dock your point total by 3%.
- Submissions must be in the form of
  - a PDF containing the write up including screen grabs of image outputs and copies of code snippets when requested
  - a zip file containing the code.

# Homework schedule

https://canvas.cmu.edu/courses/35931/pages/schedule

- First homework due Sept 16
- Will be released no later than Monday, Sept 2 (2 weeks prior)

**Lecture slides are available here.**

| Date | Topic | Reference Material | HW Due |
|---|---|---|---|
| Aug 26 | Introductions, Policies, Grading, Applications of CV | Ballard & Brown Chap 1 | |
| Aug 28 | Filtering, Convolution, Edge Detection | FP 4.1,4.2,4.5,4.6, 4.7,5.1,5.2, Gabor Filters, Links to an external site. SVD | |
| Sept 2 | NO CLASS - LABOR DAY | | |
| Sept 4 | Correspondence, Descriptors, RANSAC | | |
| Sept 9 | Pinhole Camera Model, Camera Calibration, Homographies | FP 1.1, 1.2 | |
| Sept 11 | Image Formation, Other Camera Models, Transformations | Szeliski 3.6 | |
| Sept 16 | Alignment, Lucas-Kanade | | HW1 - Augmented Reality with Planar Homographies |

# Late policy

Each student has a total of 6 late-day points for the course.

You can extend an assignment deadline by one day using one late-day point.

Rules for the late-day points are as follows:

- A maximum of 3 late-day points are allowed for the same assignment. If the submission is late by more than 3 days, it will be graded as 0%.
- There are no partial late days (5 minutes late = 1 late day)
- Late-day points CANNOT be used for HW6. HW6 has a hard deadline of its assigned due date.
- If all late-day points are used up, late homework submissions will be graded with 20% deduction for each additional late day (maximum score of 80% after one day, 60% after two days, 40% after 3 late days, etc). If the submission is late by more than 3 days, it will be graded as 0%.

# Academic Honesty

- Homeworks can be discussed, but each student must independently write up their own solutions.
- In particular, no sharing of code with other students.
- It is fine to use reference materials found online, but do not search for homework solutions.
- If you are stuck, you are encouraged to ask questions at both office hours, in Piazza, and to discuss the homework concepts with other students.

# What counts as cheating?

- General rule: If you do not understand what you are handing in, you are probably cheating.
- Copying (program or assignment) files from another person or source, including retyping their files, changing variable names, etc
- Allowing someone else to copy your code or written assignment
- Sharing your code in a public GitHub repository (private repositories that are not shared with others is fine)
- Receiving answers from students who have taken the course in previous years.
- Reviewing any homework solutions or code from previous years.

# Academic Honesty

- We've seen attempts at cheating in the past
- If we spot any cases of cheating, we will prosecute to the fullest extent
- In my opinion, if you take this route, you're missing the point of grad school.
- You are **encouraged** to discuss homeworks with each other
- Discussion with others can often enhance learning.

*To protect yourself, please mark your collaborators and website references in your homework submission*

# Study Groups

- Students are encouraged to host study groups for extra credit.
- To receive extra credit, study groups must be listed on this Google spreadsheet
- To receive extra credit, at least 4 people must attend the study group.
- Hosts will receive extra credit (up to 3%).

Open Online Study Groups 16-720A Spring 2022

File Edit View Insert Format Data Tools Extensions Help Last edit was made 4 days ago by Guillermo Gutierrez

Share

A1

| | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| 1 | | | | | | | | |
| 2 | **Date** | **Start Time** | **End Time** | **Student 1 (Full Name, Andrew ID)** | **Student 2 (Full Name, Andrew ID)** | **Group Description (brief background and experience with computer vision)** | **Zoom Link** | **Peak Attendence (fill out after session with andrew names of attendees)** |
| 3 | Saturday January 1 | 9:30 AM | 11:30 AM | Susan Tartan, susan@andrew | Scotty Mellon, scotty@andrew | Susan is MSCV student in RI, and Scott is PhD student in ECE. Both have prior experience in computer vision. We will work on question 5,6,7 mostly. | https://cmu.zoom.us/j/dummy_zoom | 4 (Susan, Scotty, Dave, Srinivas) |
| 4 | | | | | | | | |
| 5 | | | | | | | | |
| 6 | | | | | | | | |

# Piazza discussions

- Questions are encouraged as a way to get help from other students when you are stuck, or feel like you're going down the wrong path.
- Questions are not meant as a way to solve a problem before you've struggled with it.
- **Most learning comes from a little bit of struggling, and the assignments are meant to make you think about how to implement things.**
- Before posting a question, ask yourself whether you're truly stuck or are just avoiding spending the time to figure it out. Struggling and debugging is a big part of learning in this class!

# Piazza discussions

While posting questions:

- Please avoid private questions in general (if you have a question, it is likely many others do!).
    - Exceptions are private questions for office hours, described later.
- Check to see if the same question has been posted before!
- Tag the appropriate HW folder to help organize your questions

# Piazza discussions

Good Posts:

- Asking about unclear questions on the homework
- Clarification or questions with the slides or course material
- Questions about general topics in computer vision

Don’t be afraid to ask questions if something is unclear to you!

Bad Posts:

- Asking for help with debugging
- Public complaints about the course; please bring up concerns directly with instructor/TAs directly or in a private Piazza post. We’d love any feedback on how to improve!

# Office hours

- Mix of virtual + in person
- Office hours schedule and location can be found on Canvas.

# Office hours

Office hours tend to involve large queues of students. In order to keep things fair and efficient, we will use the following policy:

- Look at the OH spreadsheet to see which TAs will be present and locate their personal Zoom links / room.
- Please prepare a private post on Piazza before the office hours with your question. The more thought-out the question, the more likely you'll be able to get an immediate and helpful answer. **If you have a coding or debugging question, please include screenshots, detailed information about your issue, and what you have tried.** This will make it much easier for course staff to understand and help resolve your issue.
- Add yourself to the office hours queue for 16-820. Link to your private Piazza post.
- TAs will spend 10 minutes per person, and bump them to the bottom of the queue if they need additional help.

# Student Wellness

- Take care of yourself.  Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

- All of us benefit from support during times of struggle. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is almost always helpful.

- If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at http://www.cmu.edu/counseling/ . Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

# Statement of Diversity

We are diverse in many ways, and this diversity is fundamental to building and maintaining an equitable and inclusive campus community.

Diversity can refer to multiple ways that we identify ourselves, including but not limited to race, color, national origin, language, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information.

Our diverse identities shape the perspectives our students, faculty, and staff bring to our campus.

We must treat every individual with respect.

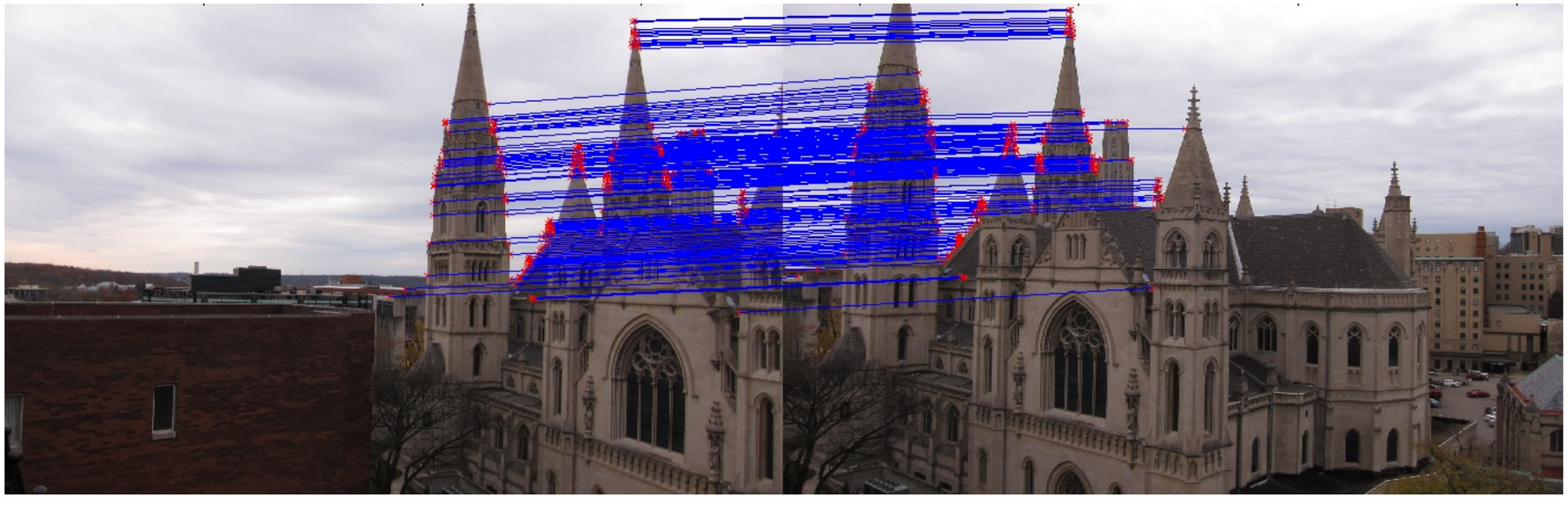
Each of us is responsible for creating a safer, more inclusive environment.

# Reporting Bias or Discrimination

Unfortunately, incidents of bias or discrimination do occur, whether intentional or unintentional. They contribute to creating an unwelcoming environment for individuals and groups at the university. Therefore, the university encourages anyone who experiences or observes unfair or hostile treatment on the basis of identity to speak out for justice and support, within the moment of the incident or after the incident has passed. Anyone can share these experiences using the following resources:

- **Center for Student Diversity and Inclusion:** csdi@andrew.cmu.edu, (412) 268-2150
- **Report-It online anonymous reporting platform:** reportit.net username**:** *tartans* password**:** *plaid*

# HW1: Panoramic Mosaics



Concepts: interest points, descriptors, RANSAC, homographies

# HW2: Tracking


Concepts: Templates, Lucas-Kanade alignment

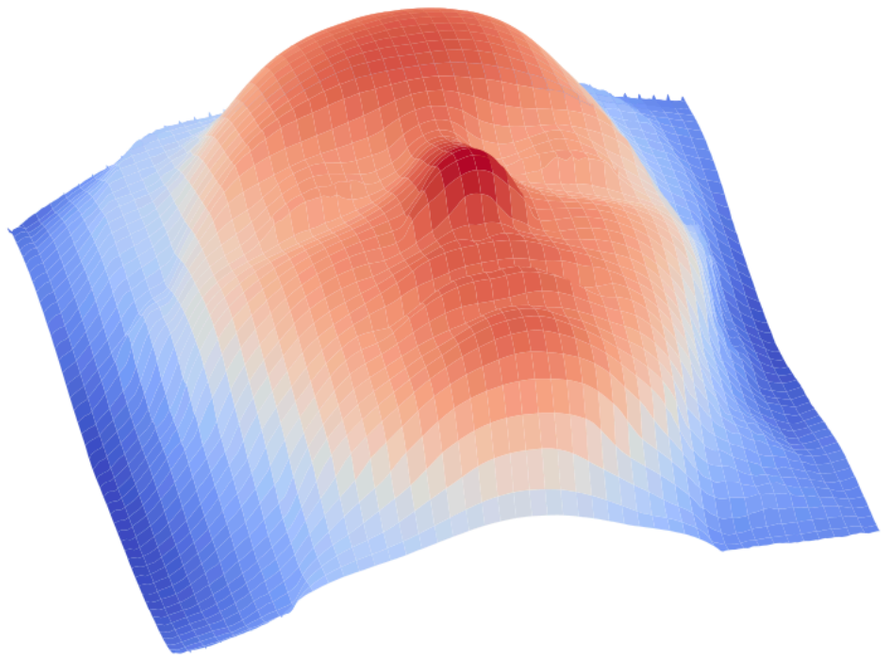
# HW3: 3D Reconstruction

Concepts: Epipolar geometry, triangulation

# HW4: Deep learning

Concepts: Pattern classification

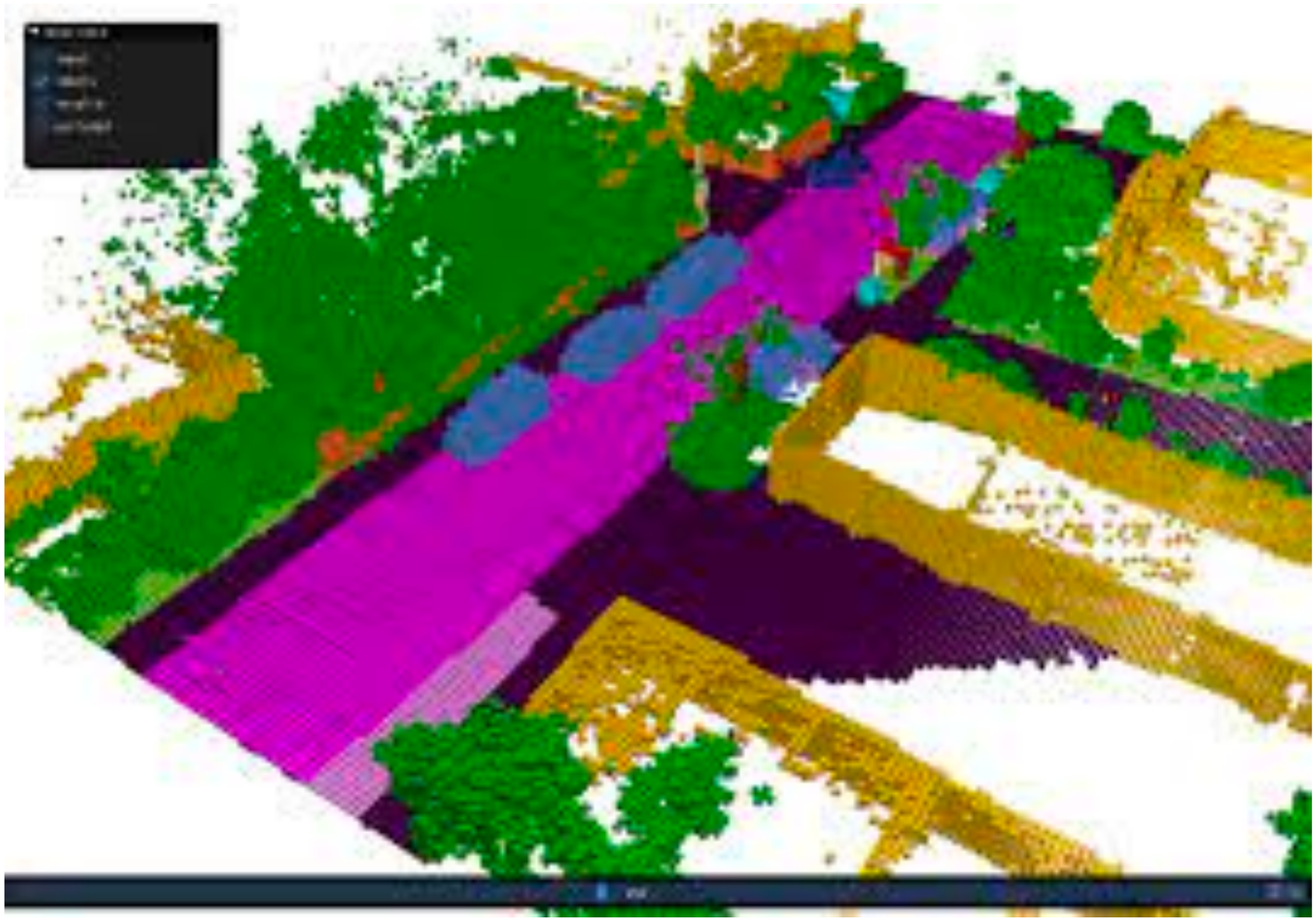
# HW5: Photometric Stereo



Concepts: physics-based vision



# HW6: 3D Segmentation



# Suggestions for this class

*Its all about the assignments*

A big factor in success will be your comfort and debugging capability in Python

(I probably spent 80 percent of my time as a grad student “debugging”)

**Previous evaluations of this class**

Difficulty: 7/10 (10 = impossible to get through)

“This is a demanding course, but is well worth it.”

# Some final class philosophies

- Diverse background of class implies folks will find some of you will find it too slow / some too fast

- To differentiate from 16-720, I will err on the side of going a bit fast

- If you find the class too fast, you can switch to 16-720

- One-way lectures are very hard for teaching. Discussions are way more fun! I encourage you to partake in the lecture and interrupt with questions.  I also highly encourage you to leave your video on.