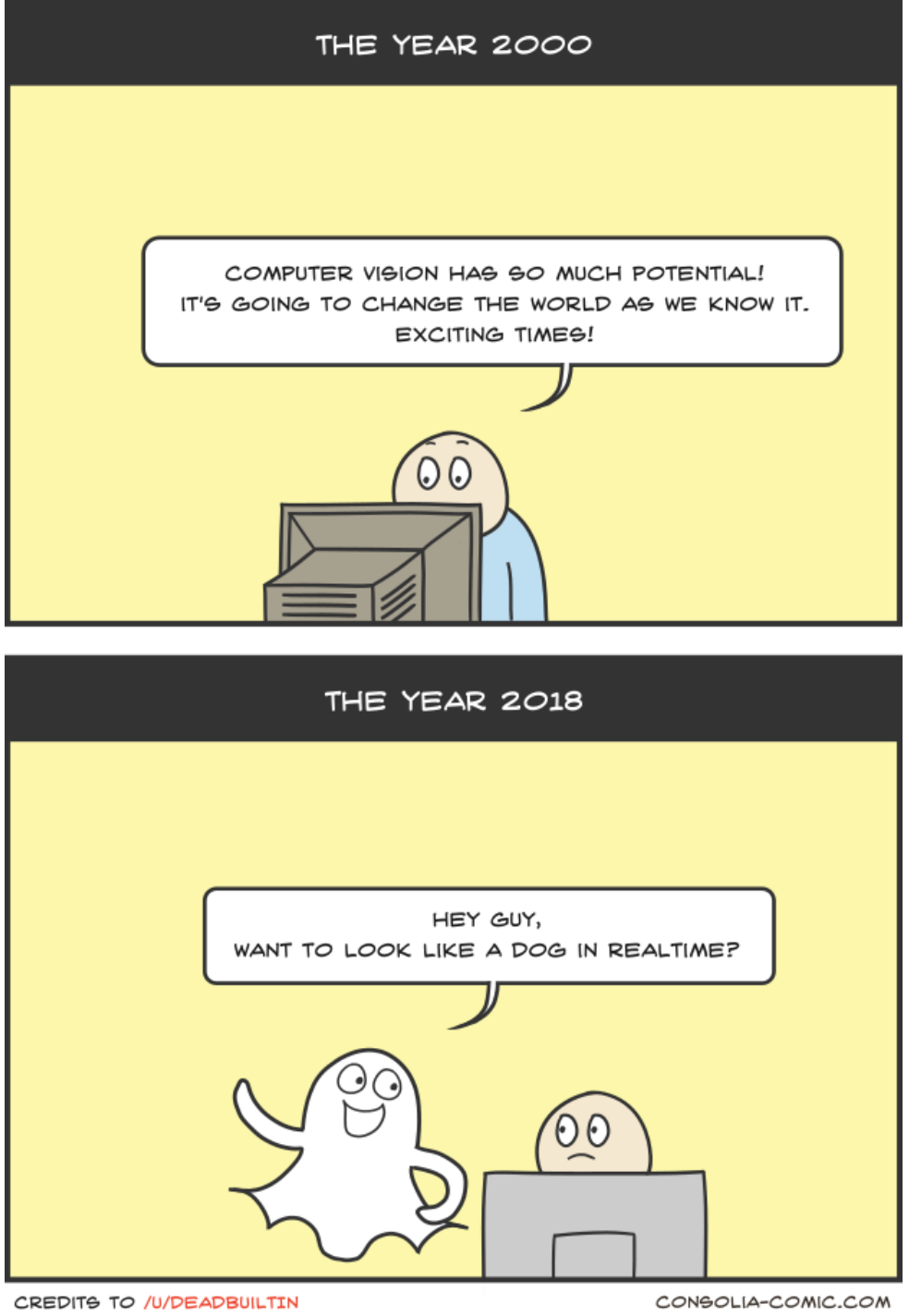
CMSC 491/691

Lecture 1

Computer Vision

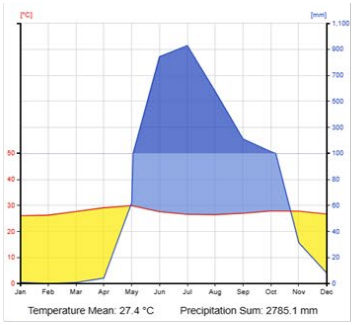
Some slides from Jayasuriya, Turaga, Szeliski



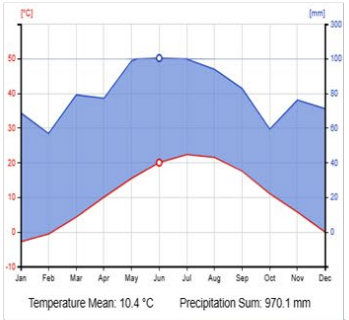
Tejas Gokhale~

Tejas: Tay + Juss
Gokhale: Go + Clay

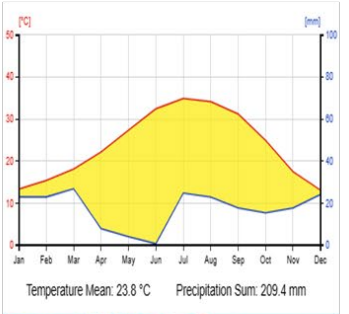
Assistant Professor
Computer Science
University of Maryland, Baltimore County



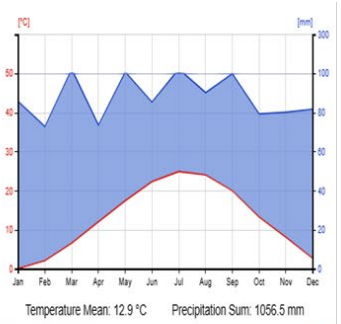
2011—2015
B.E. (Honours)
BITS Pilani
(Goa)



2016-2018:
M.S.
Carnegie Mellon
University



2018—2023:
Ph.D.
Arizona State
University



2023—present
Assistant Professor
University of Maryland
Baltimore County

https://www.tejasgokhale.com/

Tejas Gokhale, Assistant Professor

www.tejasgokhale.com Lab: ITE 368

Current Projects

- Domain Adaptation/Generalization
- Quantifying Visual Typicality
- Semantics and Pragmatics of Vision
- Multimodal Continual Learning
- Novel Concept Discovery
- Generative AI evaluation

Recent Activities

- AAAI 2024: New Faculty Highlights Invited Talk
- WACV 2024: Tutorial on Reliability of Generative Models
- Area Chair: NAACL 2024
- UMBC PPR Seminar

Research Areas

- Computer Vision
- Vision & Language
- Visual Reasoning
- Active Perception
- Robustness & Reliability

UMBC




Course Staff

Instructor: Tejas Gokhale
Assistant Professor, CSEE

OH: Wednesday 2 – 3:30 PM ITE 214

gokhale@umbc.edu

TA: Sourajit Saha
Ph.D. student, CSEE

OH:
Monday 1:30 -- 3:30 PM & Tuesday 2:30 -- 4:30 PM

ssaha2@umbc.edu

Class Website

https://redirect.cs.umbc.edu/courses/graduate/691cv/

Quick Round of Introductions

(1) Name

(2) Major (e.g. CS)

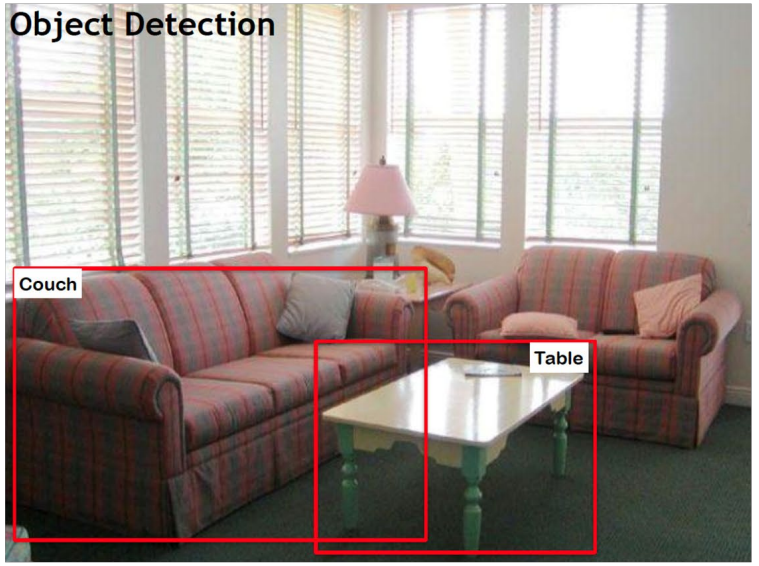
(3) Level (BS / MS / PhD)

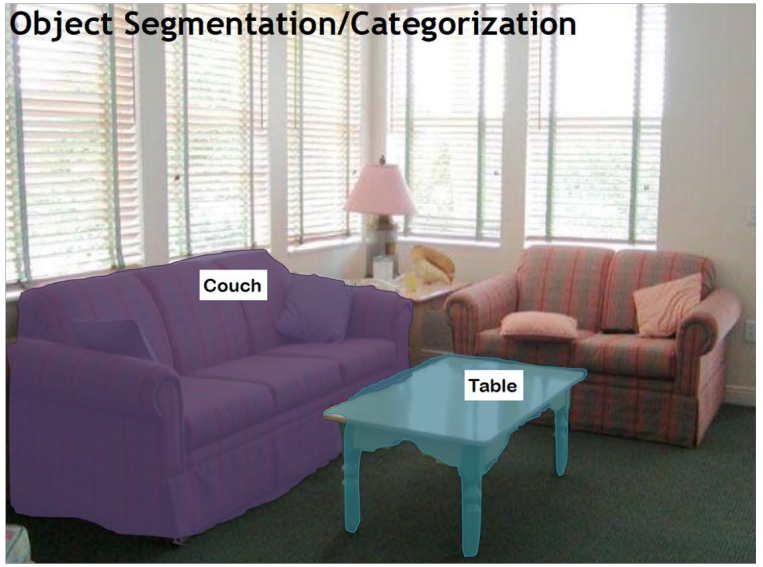
(4) Why are you taking this class?

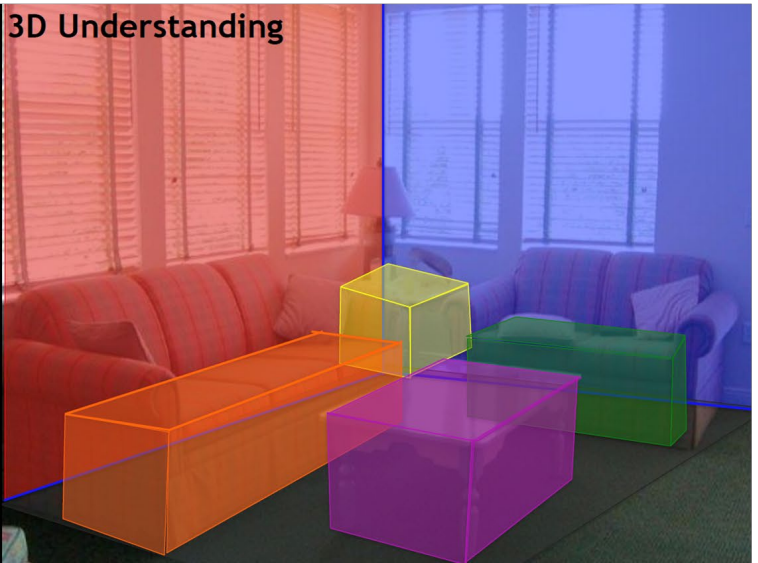
What is this course about ?

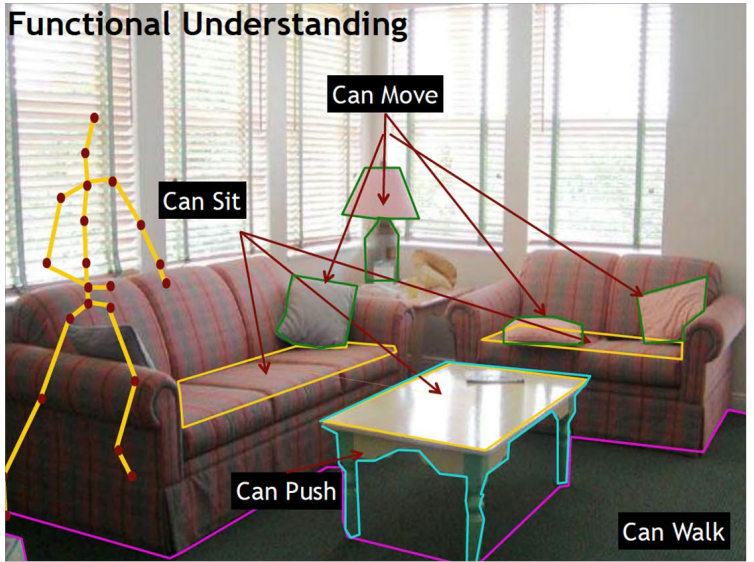
"Understanding Images"

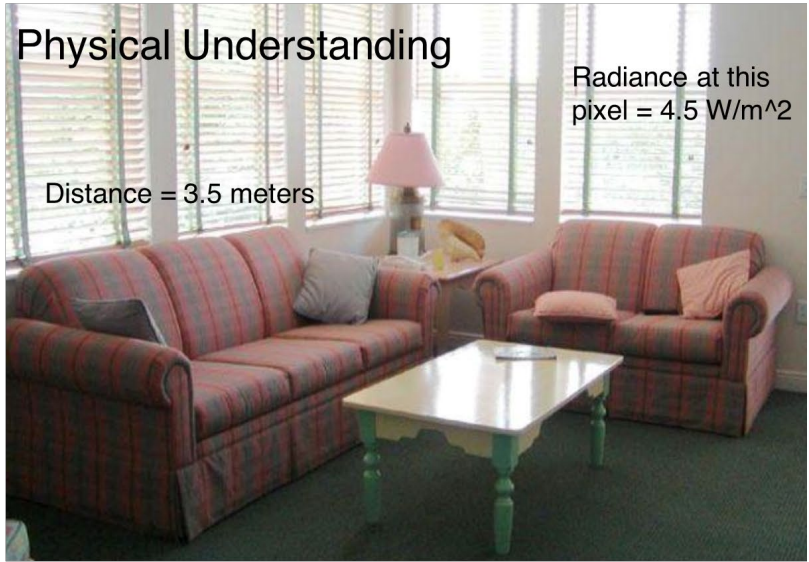
Computer Vision: Understanding Images











Why is this difficult?

What I see

Why is this difficult?

What I see

What a computer sees

```
08 02 22 97 38 15 00 40 00 75 04 05 07 78 52 12 50 77 91 08
49 49 99 40 17 81 18 57 60 87 17 40 98 43 69 48 04 56 62 00
81 49 31 73 55 79 14 29 93 71 40 67 53 88 30 03 49 13 36 65
52 70 95 23 04 60 11 42 69 24 68 56 01 32 56 71 37 02 36 91
22 31 16 71 51 67 63 89 41 92 36 54 22 40 40 28 66 33 13 80
24 47 32 60 99 03 45 02 44 75 33 53 78 36 84 20 35 17 12 50
32 98 81 28 64 23 67 10 26 38 40 67 59 54 70 66 18 38 64 70
67 26 20 68 02 62 12 20 95 63 94 39 63 08 40 91 66 49 94 21
24 55 58 05 66 73 99 26 97 17 78 78 96 83 14 88 34 89 63 72
21 36 23 09 75 00 76 44 20 45 35 14 00 61 33 97 34 31 33 95
78 17 53 28 22 75 31 67 15 94 03 80 04 62 16 14 09 53 56 92
16 39 05 42 96 35 31 47 55 58 88 24 00 17 54 24 36 29 85 57
86 56 00 48 35 71 89 07 05 44 44 37 44 60 21 58 51 54 17 58
19 80 81 68 05 94 47 69 28 73 92 13 86 52 17 77 04 89 55 40
04 52 08 83 97 35 99 16 07 97 57 32 16 26 26 79 33 27 98 66
88 36 68 87 57 62 20 72 03 46 33 67 46 55 12 32 63 93 53 69
04 42 16 73 38 25 39 11 24 94 72 18 08 46 29 32 40 62 76 36
20 69 36 41 72 30 23 88 34 62 99 69 82 67 59 85 74 04 36 16
20 73 35 29 78 31 90 01 74 31 49 71 48 86 81 16 23 57 05 54
01 70 54 71 83 51 54 69 16 92 33 48 61 43 52 01 89 19 67 48
```

Humans are good at it …

Are Humans good at it ?

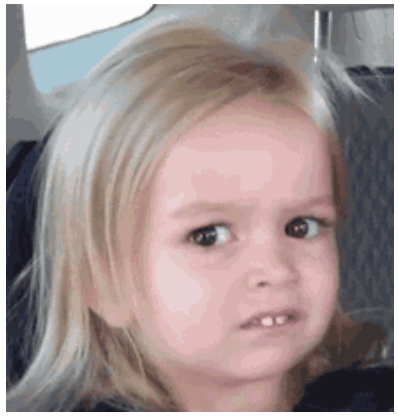
Challenges: object intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: Motion

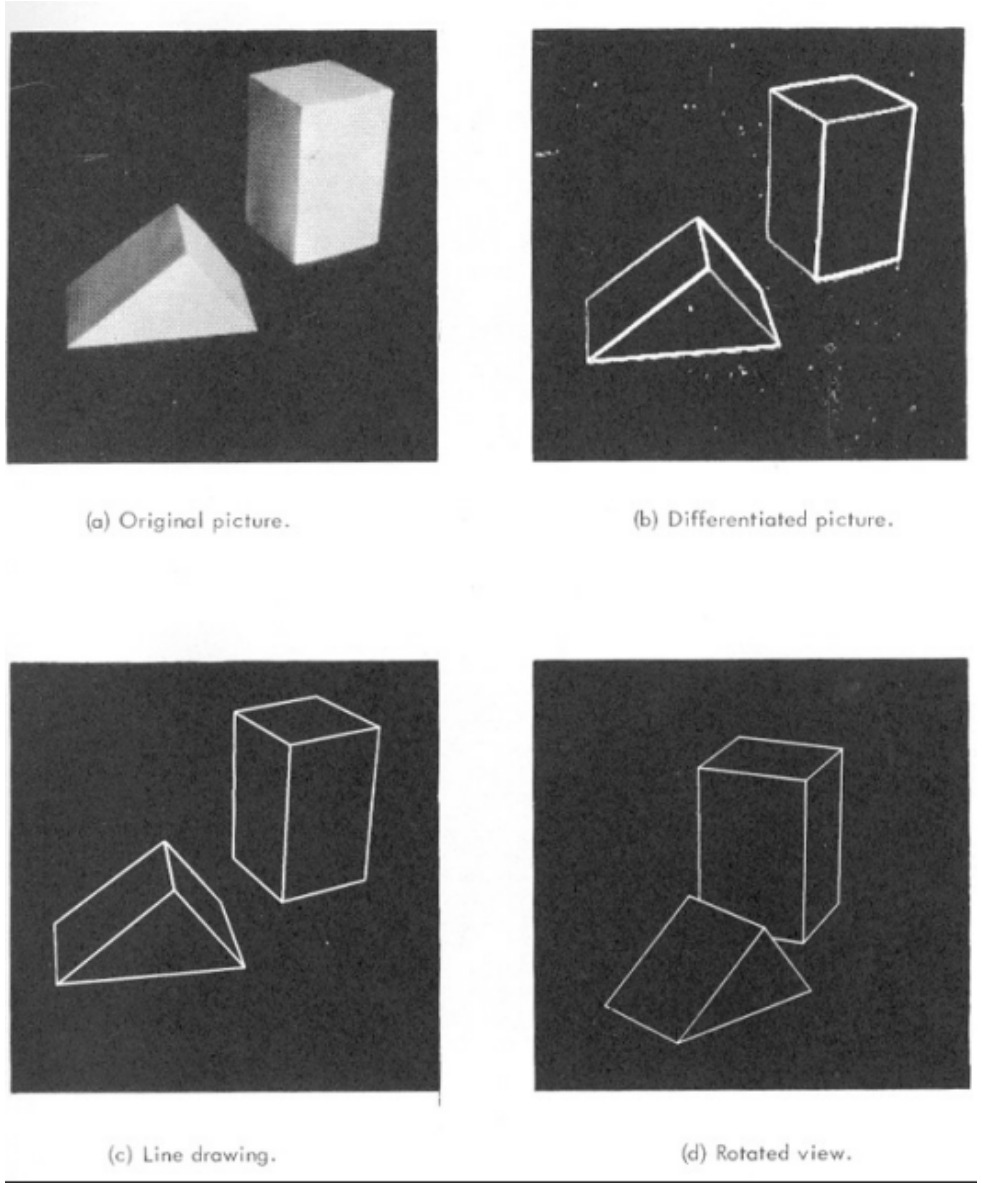
Challenges: background clutter

Humans have *different* ways of understanding



Images convey human emotions

Some history ...



(a) Original picture.

(b) Differentiated picture.

(c) Line drawing.

(d) Rotated view.

L. G. Roberts,

Machine Perception of Three Dimensional Solids

Ph.D. thesis,

MIT Dept of Electrical Engineering

1963

Some history ...

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group — July 7, 1966
Vision Memo. No. 100.

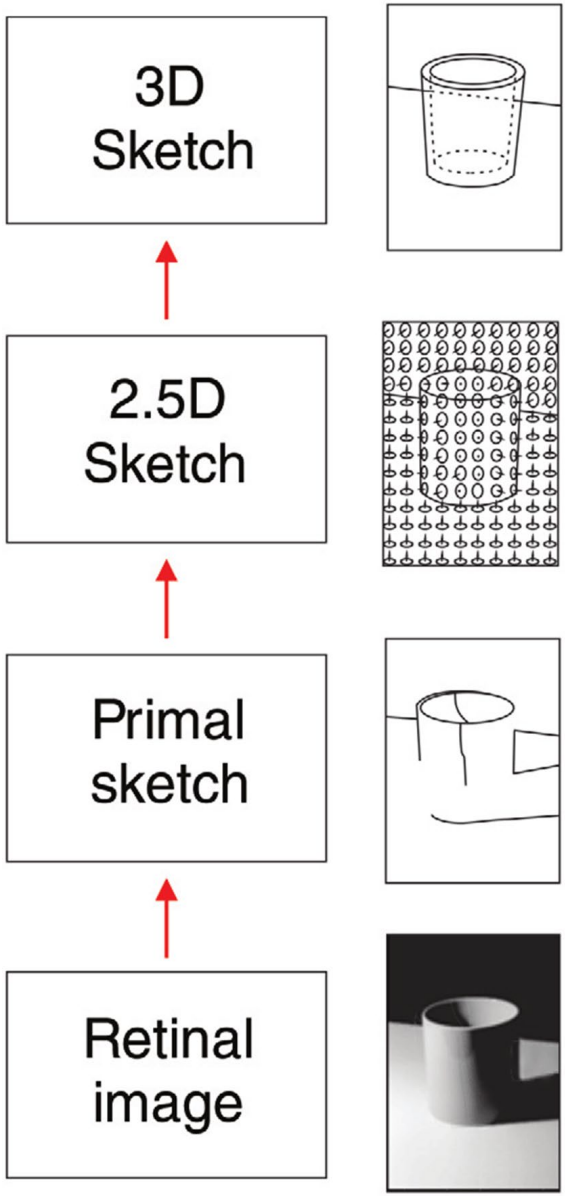
THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Marr's Vision (1982)

- British neuroscientist David Marr was very influential with his computational theory of vision
- Marr believed that visual understanding could be progressively grown from primitive lines/edges to 2.5D information and textures to finally 3D shapes
- It is still an open question of whether the brain *actually* works in this way, but many engineers and scientists found inspiration in making artificial vision systems based on Marr's model of vision.



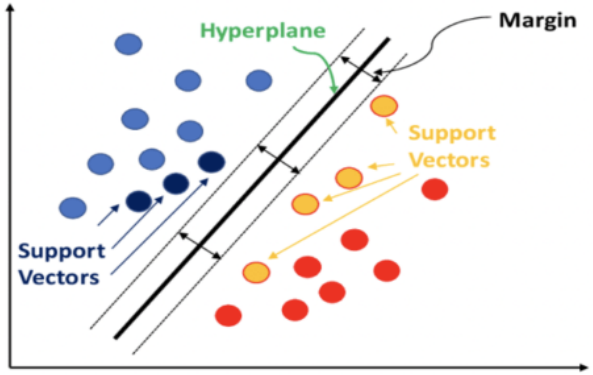
slide credit: Suren Jayasuriya

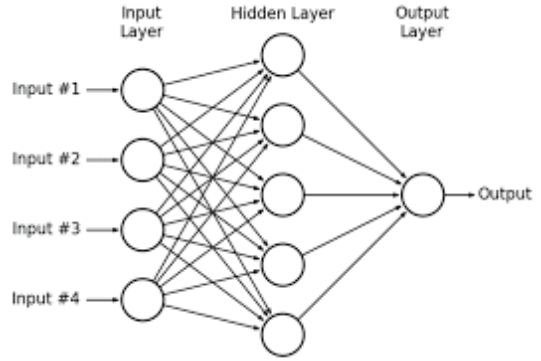
Counterpoint: Machine Learning & Statistical Methods

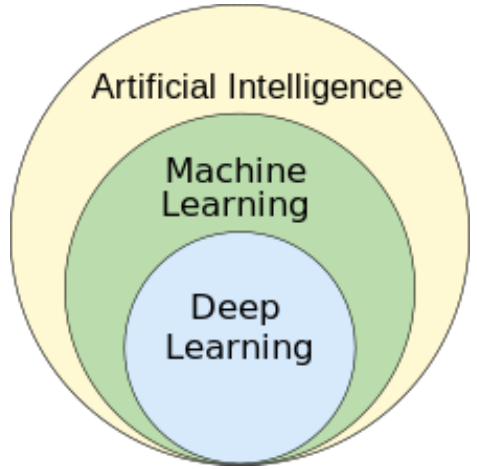
- To counter the computational theories of Marr, statistics and data-driven methods were being invented in the 1980s

- Such methods ***had*** limited application due to lack of available data and compute power (more on this later)

- Deep learning (by 2012 onward) has become a backbone of most computer vision systems

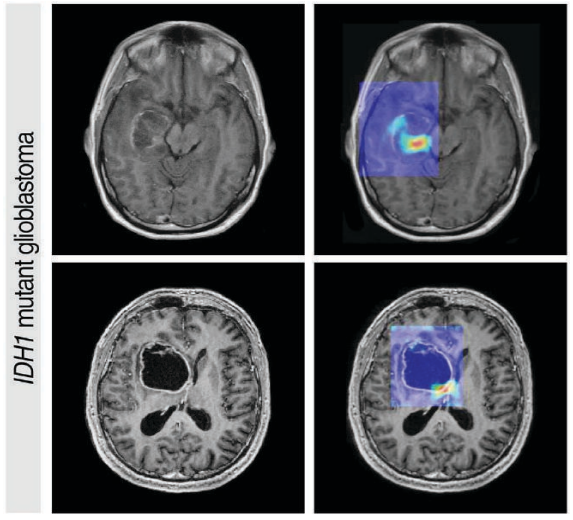




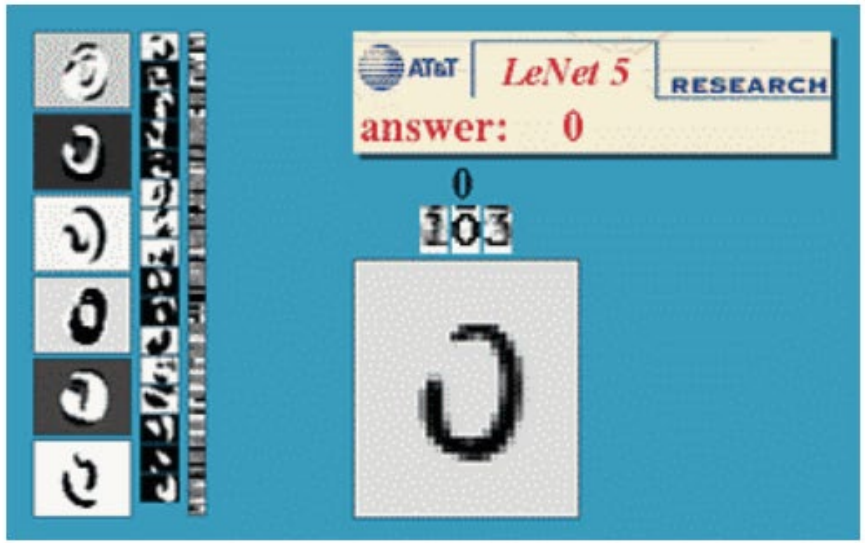


slide credit: Suren Jayasuriya

Fast Forward to Today



Optical character recognition (OCR)



Digit recognition
yann.lecun.com

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

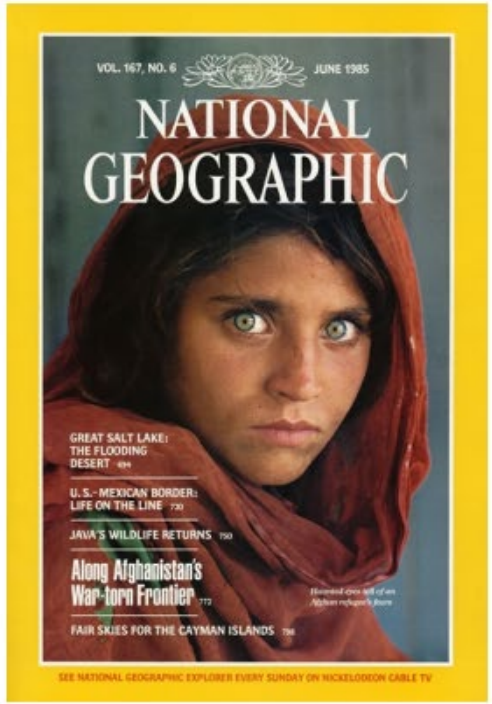
Automatic check processing

Sudoku grabber

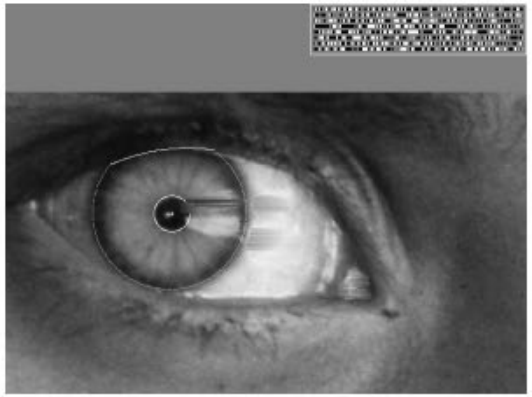
http://sudokugrab.blogspot.com/

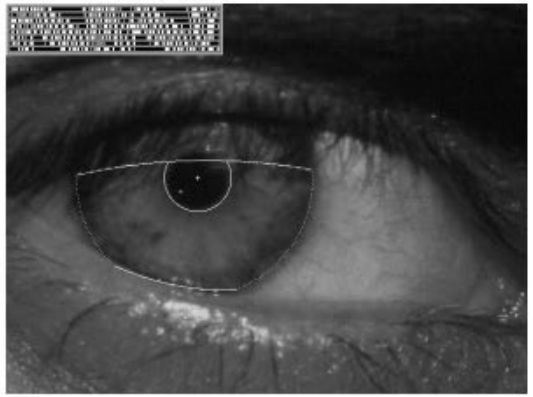
Source: S. Seitz, N. Snavely

Biometrics



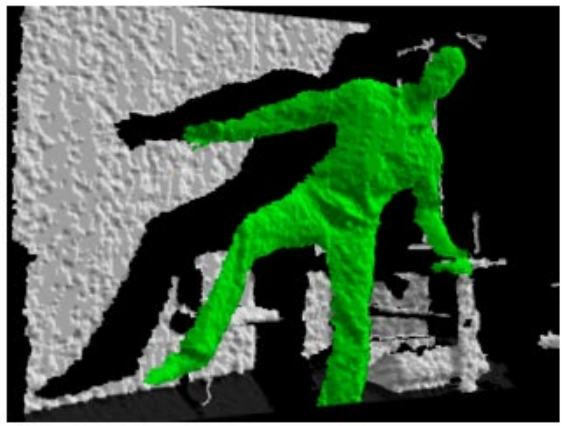
[How the Afghan Girl was Identified by Her Iris Patterns]

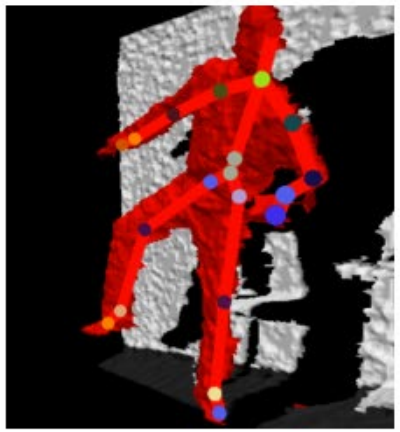




Source: S. Seitz

Vision-based interaction: Xbox Kinect





http://blogs.howstuffworks.com/2010/11/05/how-microsoft-kinect-works-an-amazing-use-of-infrared-light/

http://electronics.howstuffworks.com/microsoft-kinect.htm

http://www.xbox.com/en-US/Live/EngineeringBlog/122910-HowYouBecometheController

http://www.ismashphone.com/2010/12/kinect-hacks-more-interesting-than-the-devices-original-intention.html

Earth viewers (3D modeling)

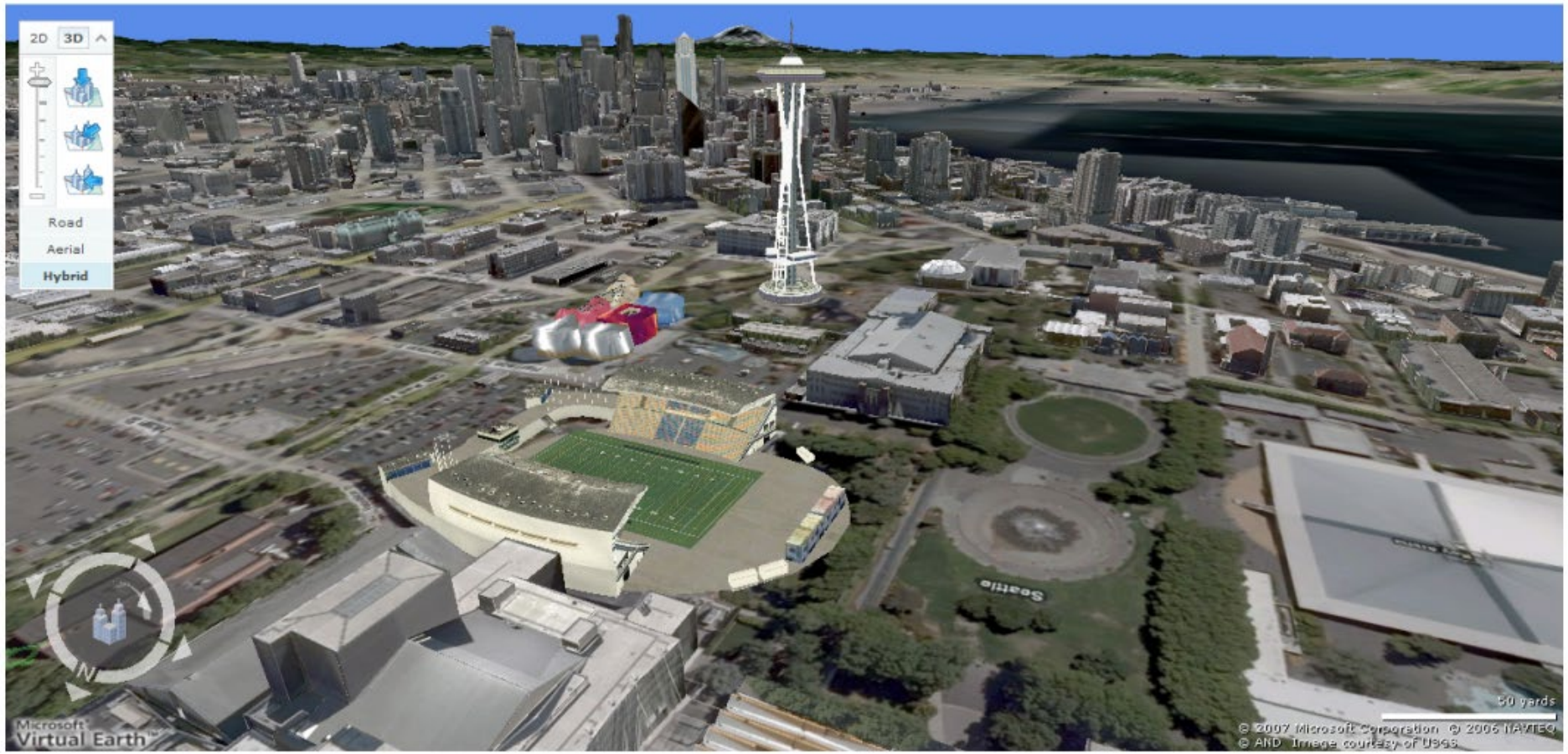


Image from Microsoft's Virtual Earth

(see also: Google Earth)

Source: Steve Seitz and Neel Joshi

Vision for robotics, space exploration

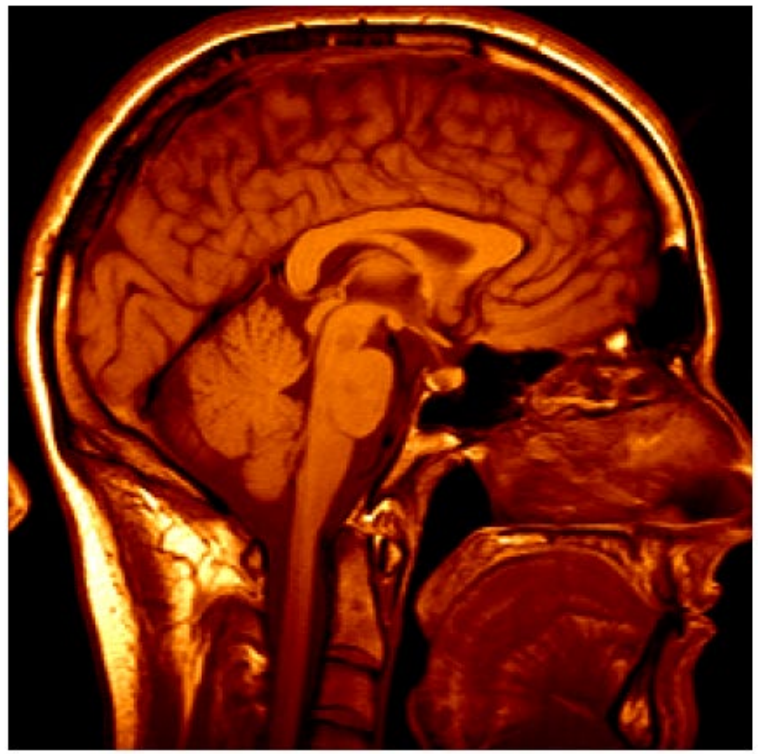
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Source: S. Seitz

Medical imaging



3D imaging

MRI, CT

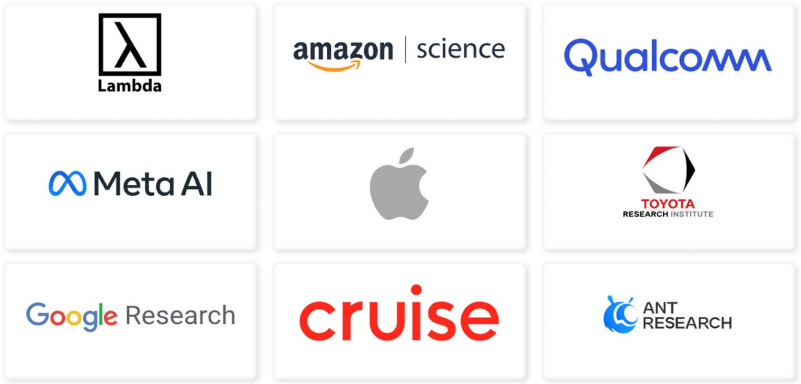
Image guided surgery

Grimson et al., MIT

Source: Steve Seitz and Neel Joshi

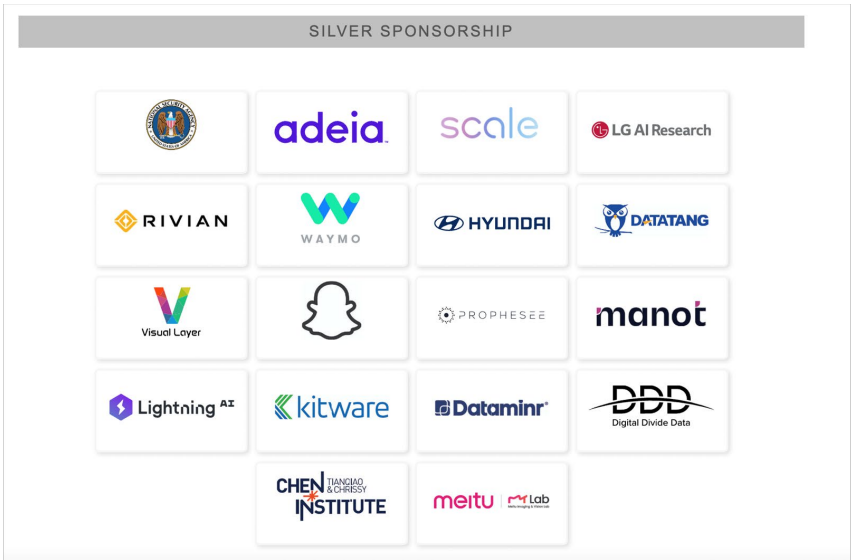
LOTS (!!!) of Industry

PLATINUM SPONSORSHIP



GOLD SPONSORSHIP

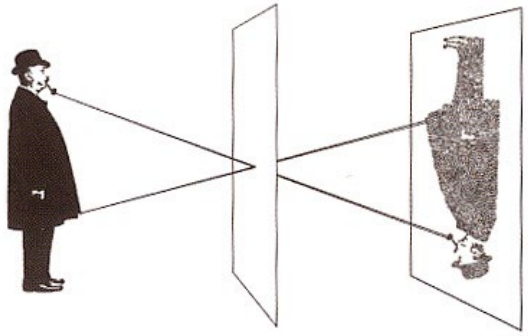
SILVER SPONSORSHIP



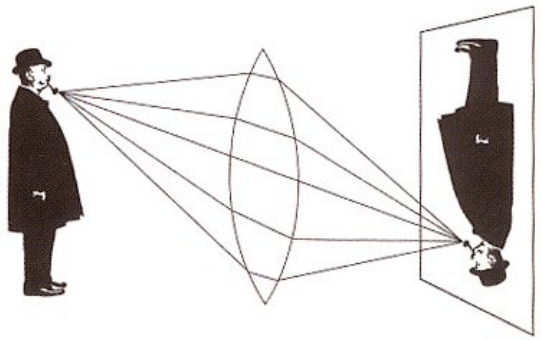
Growth in industry has been tremendous

Class Topics
(1) Image Acquisition

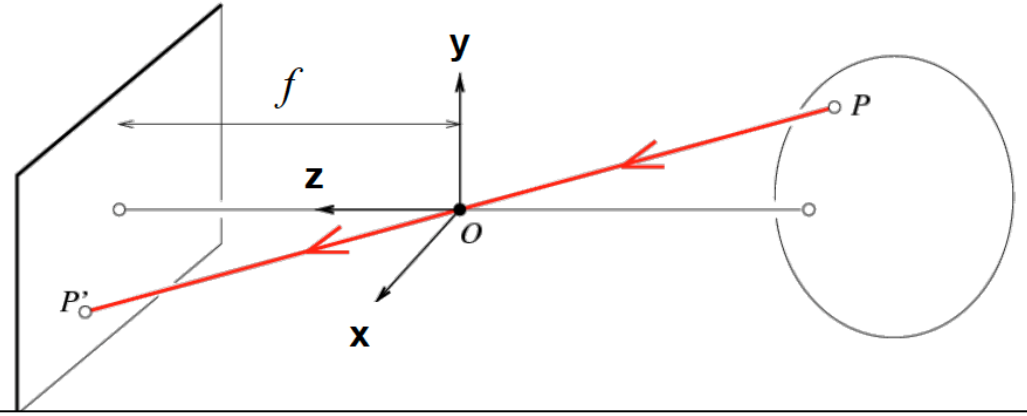
Photograph made with small pinhole



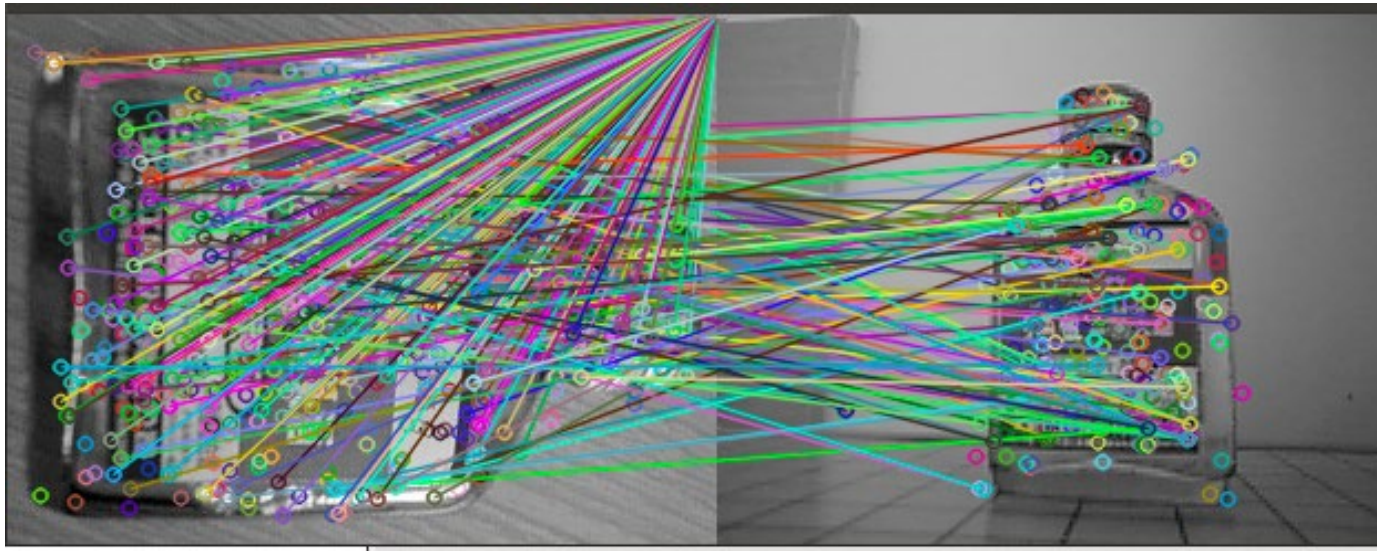
Photograph made with lens

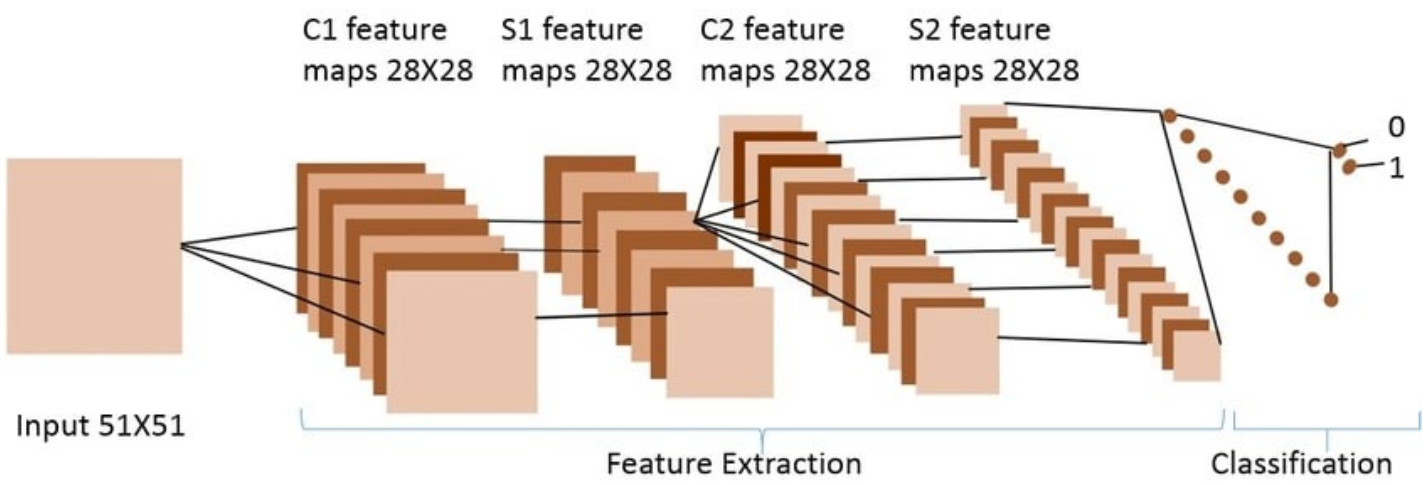


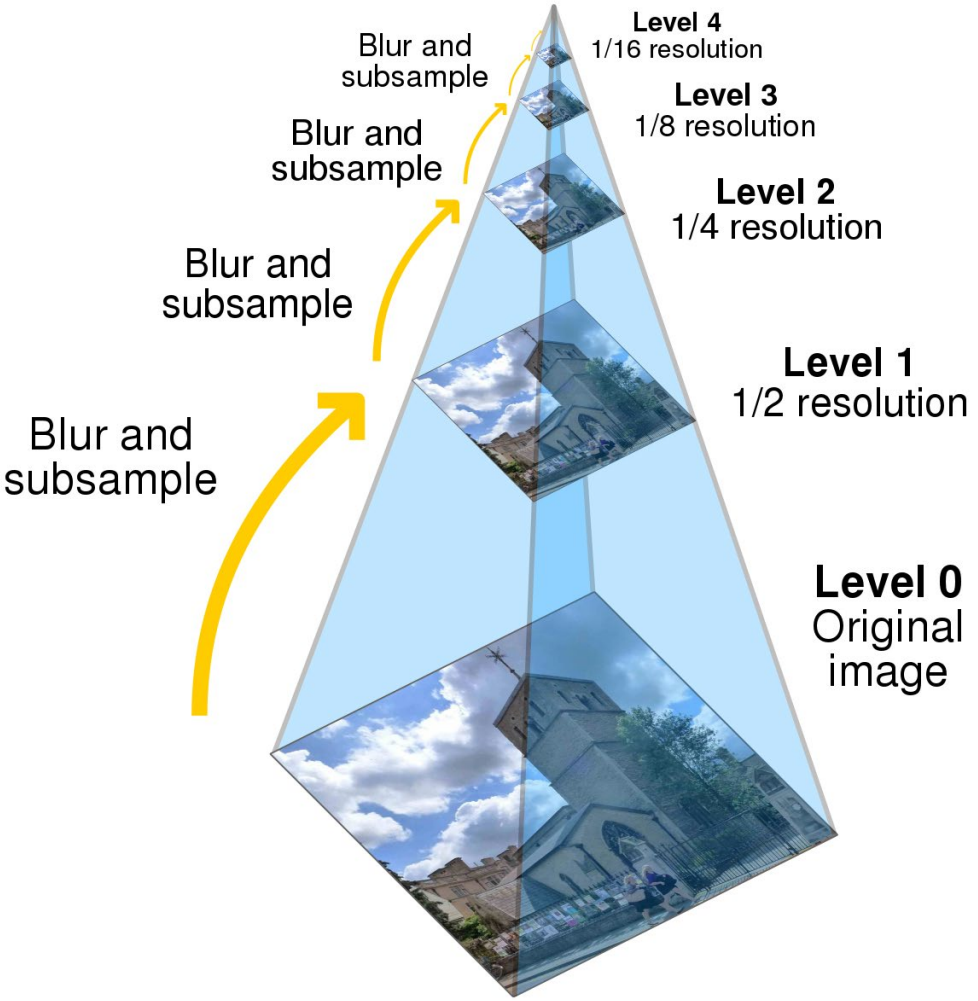
Modeling projection



Class Topics
(2) Image Processing & Features



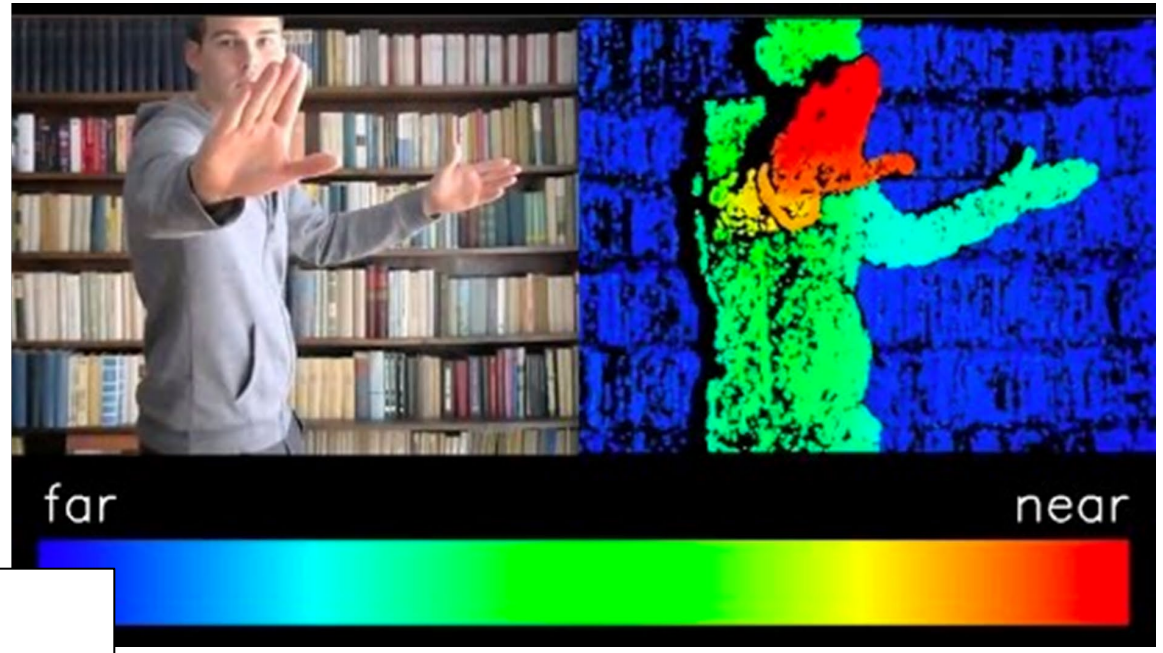


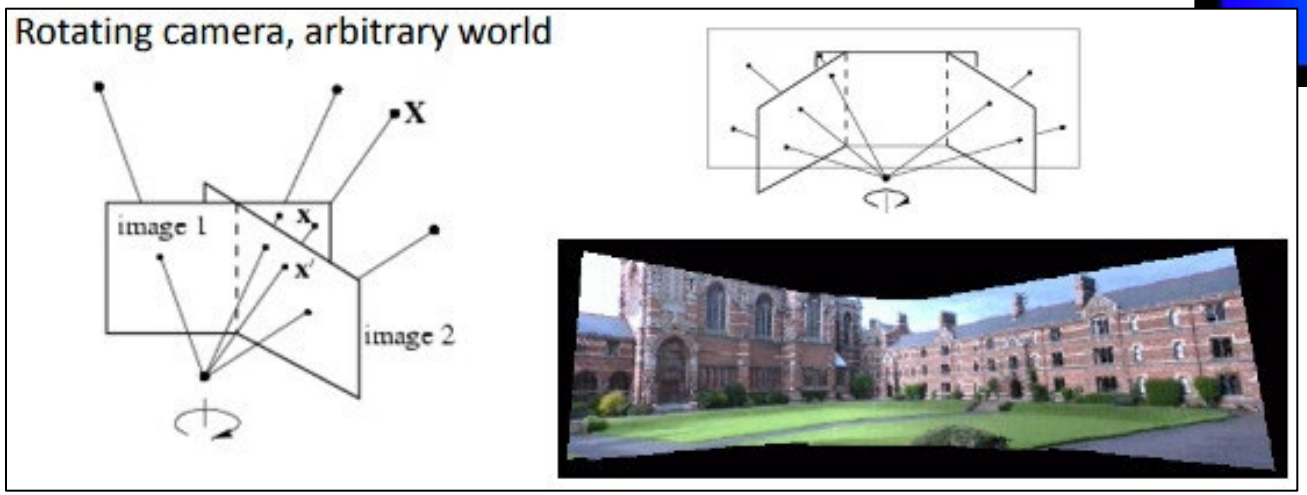


Class Topics
(3) Geometry (3D)

Topics:

- Epipolar Geometry and Stereo
- 3D geometry estimation
- Structure-from-motion

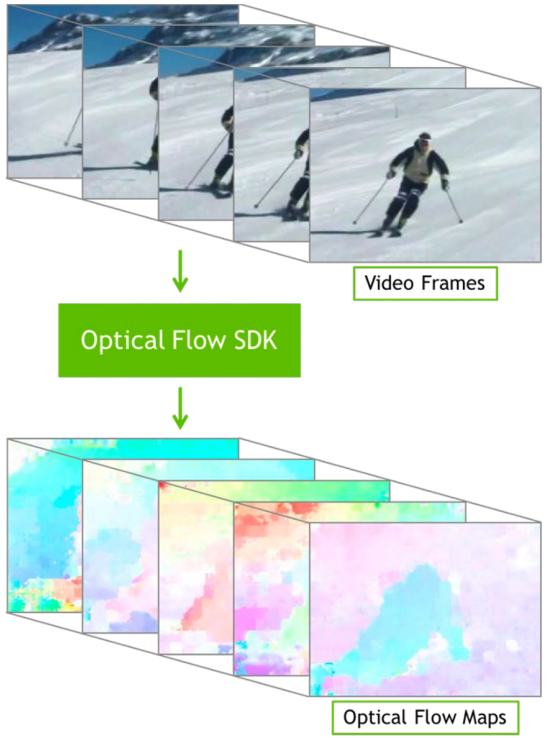


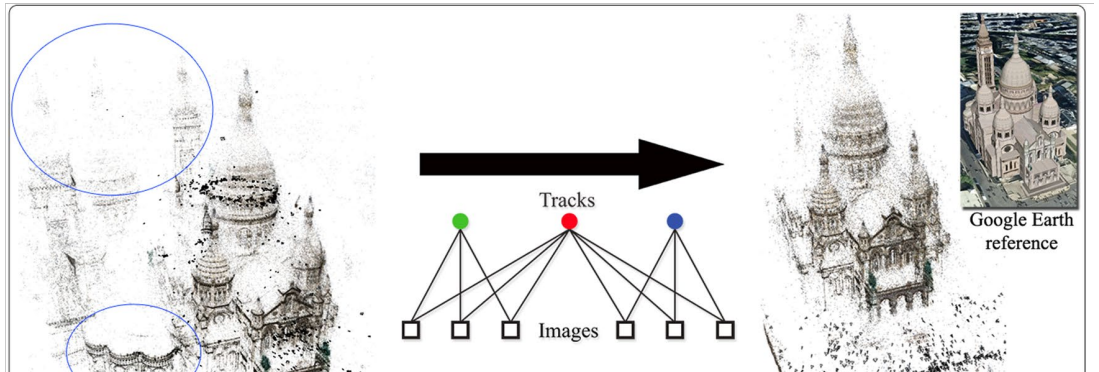


Class Topics
(4) Video (Temporal)

Topics:

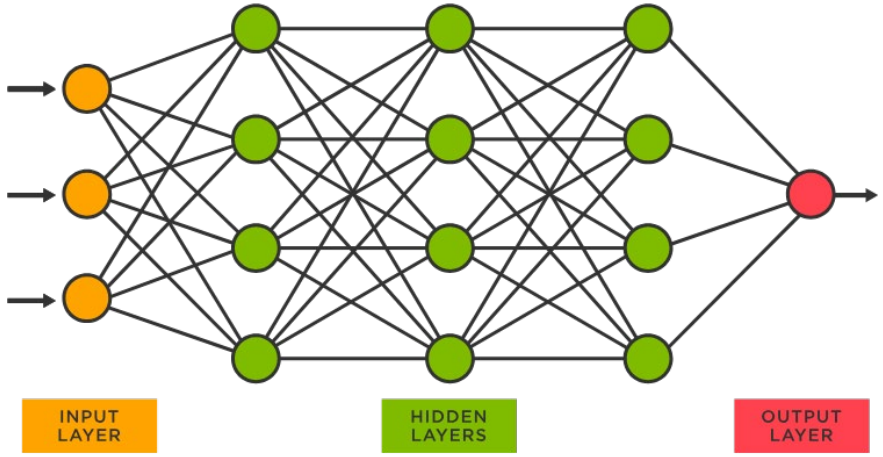
- Motion/optical flow
- Video segmentation, tracking
- Simultaneous localization and mapping (SLAM)





Class Topics
(5) Machine Learning for Vision

"Beach"



Advanced Topics: Vision + Language

Generating Text from Images

Caption Generator

A bird flying over water

Answering Questions about Images

What is the mustache made of?

AI System

bananas

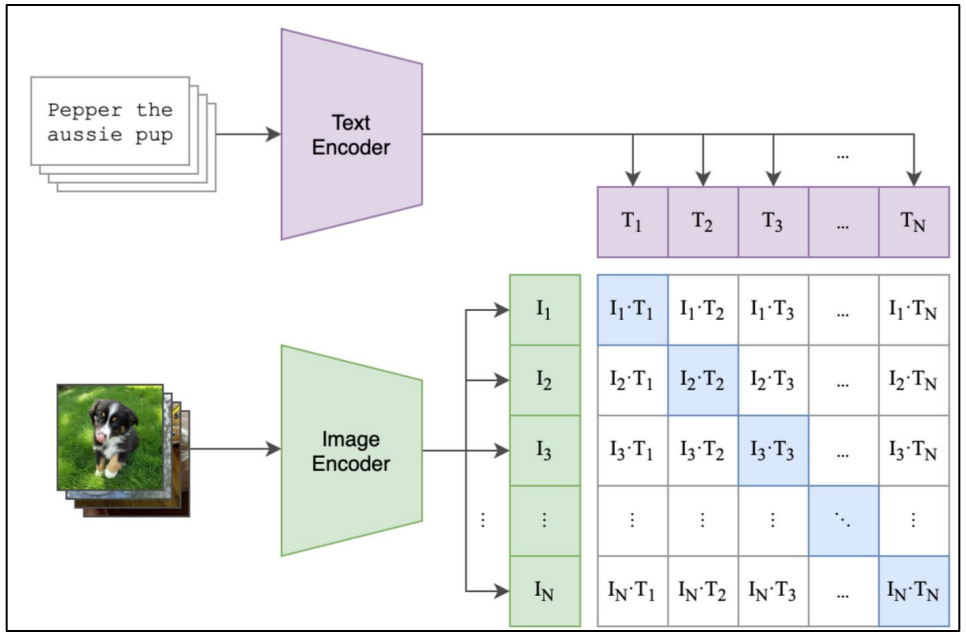
Generating Images from Text

"a corgi playing a flame throwing trumpet"

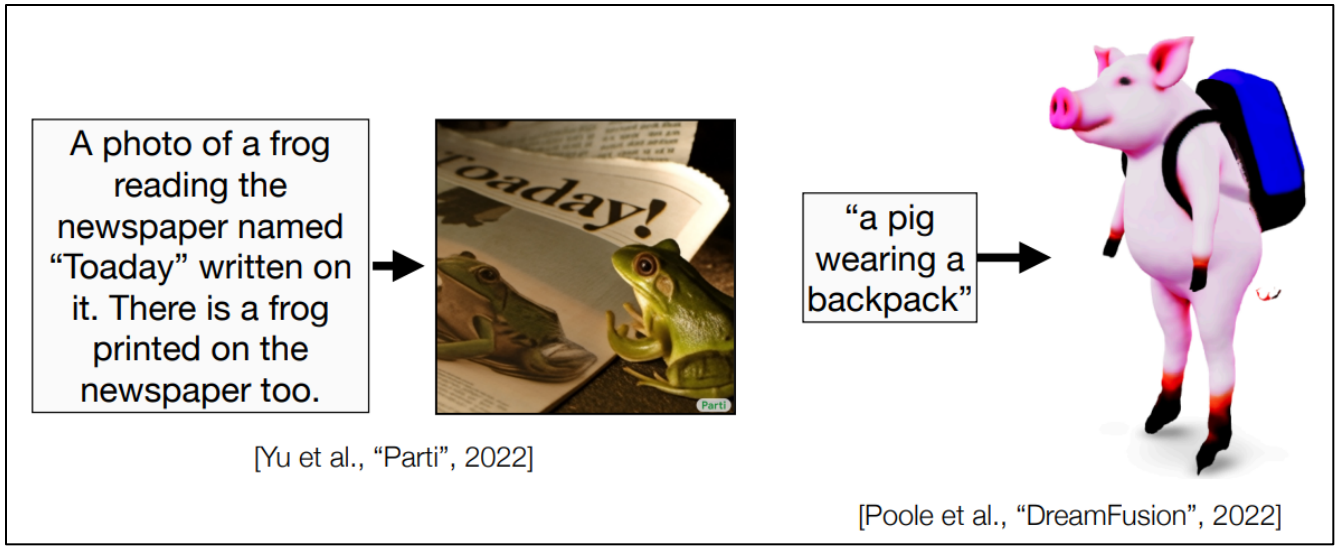
text encoder

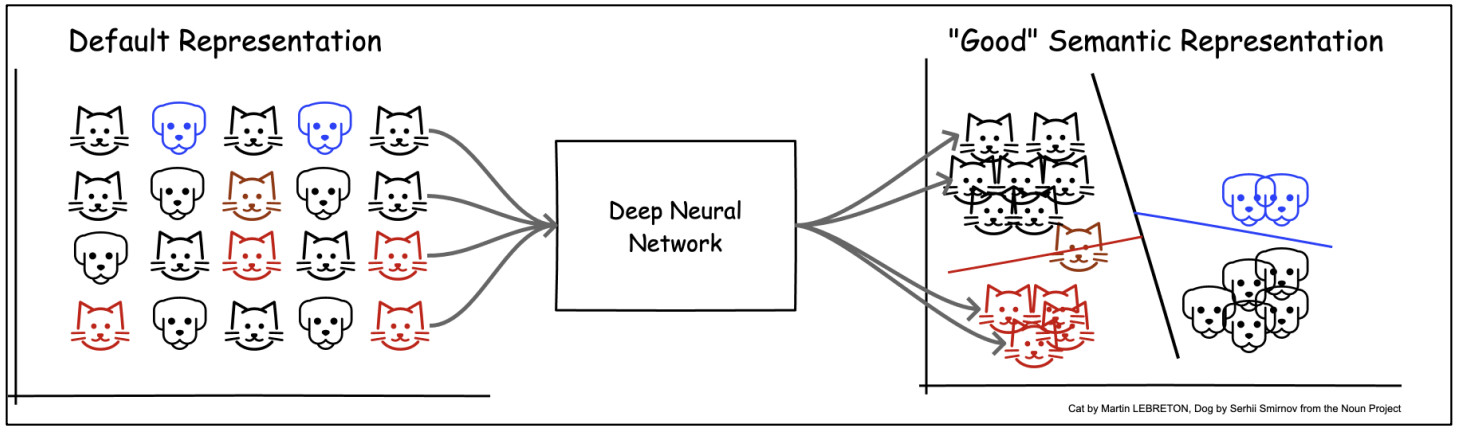
Advanced Topics

Representation Learning

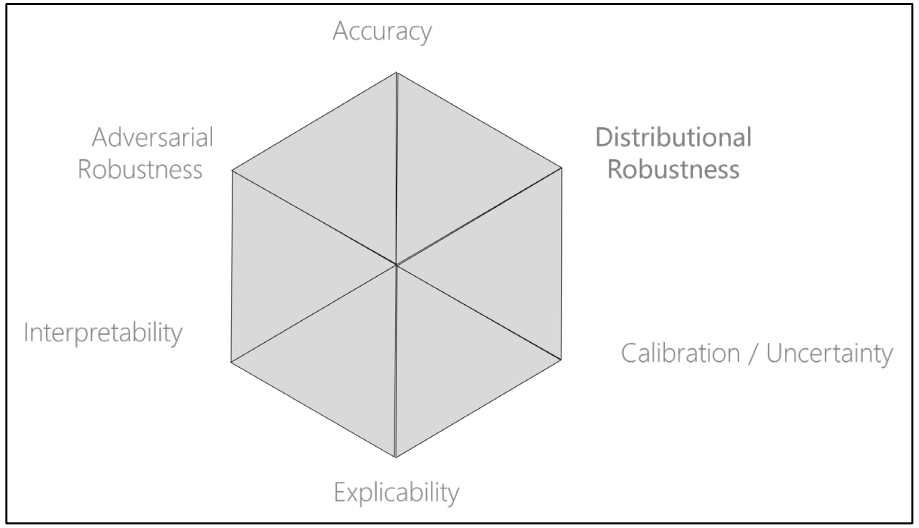


Generative Models





Robustness & Reliability



What we will learn to do

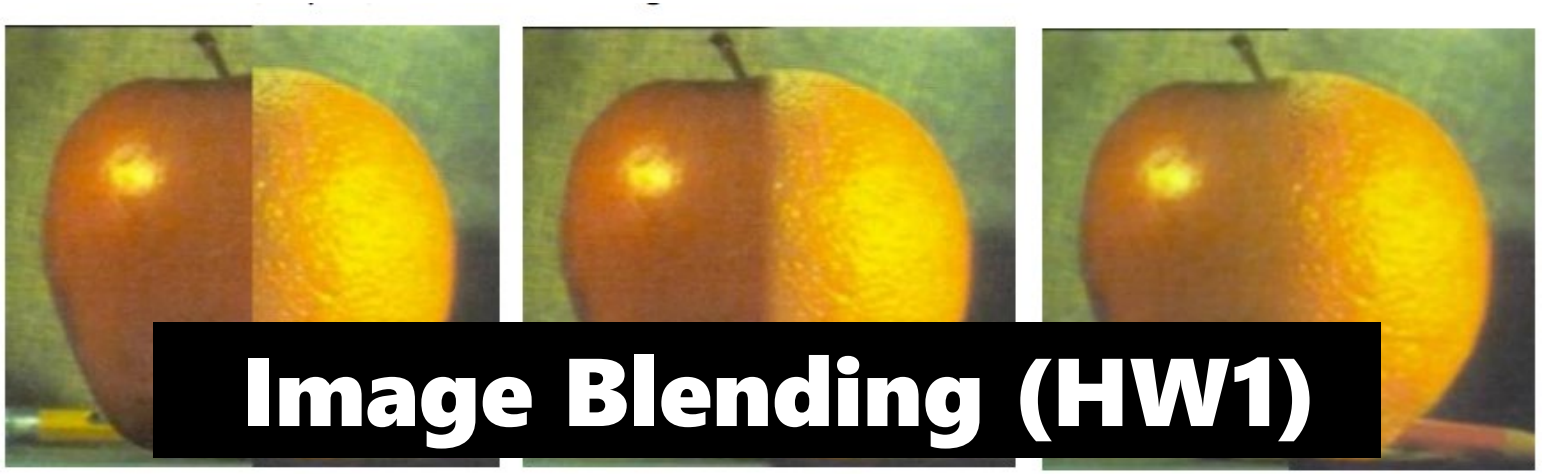


Image Blending (HW1)

Image Stitching / Panorama (HW2)

What we will learn to do

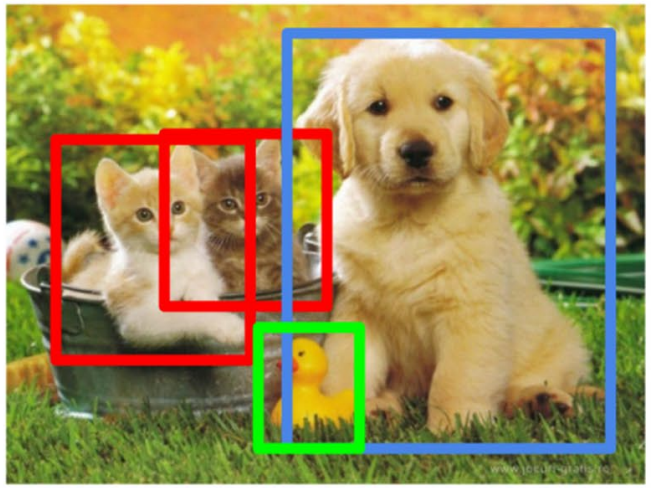
Classification

CAT

Classification + Localization

CAT

Object Detection



CAT, DOG, DUCK

Instance Segmentation

CAT, DOG, DUCK

Single object

Multiple objects

Course Logistics

Class Website

https://redirect.cs.umbc.edu/courses/graduate/691cv/

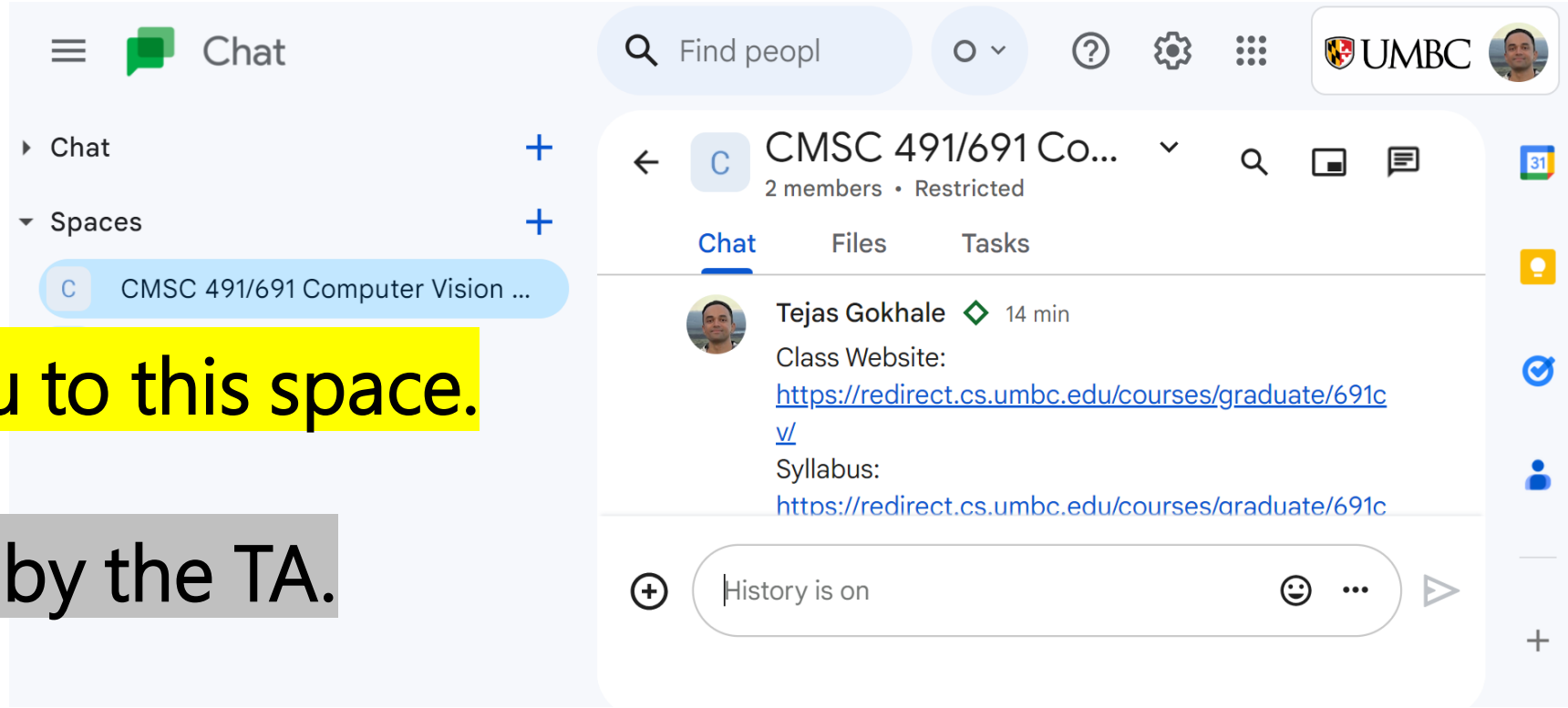
CMSC 491/691: Computer Vision

UMBC. Spring 2024

Instructor: Tejas Gokhale (OH: Wednesday 2 PM - 3:30 PM or by appointment);
Teaching Assistant: Sourajit Saha (OH: Monday 1:30 -- 3:30 PM & Tuesday 2:30 -- 4:30 PM);
Time: Monday and Wednesday 4:00pm - 5:15pm
Location: ENGR 231

Course Description | Schedule | Grading | Syllabus

Google Chat Space



We will add you to this space.

Monitored by the TA.

Use this for class-related discussions with other students and TA.

Follow academic integrity guidelines. Don't post actual code, answers, grades

Grading

Grading Components

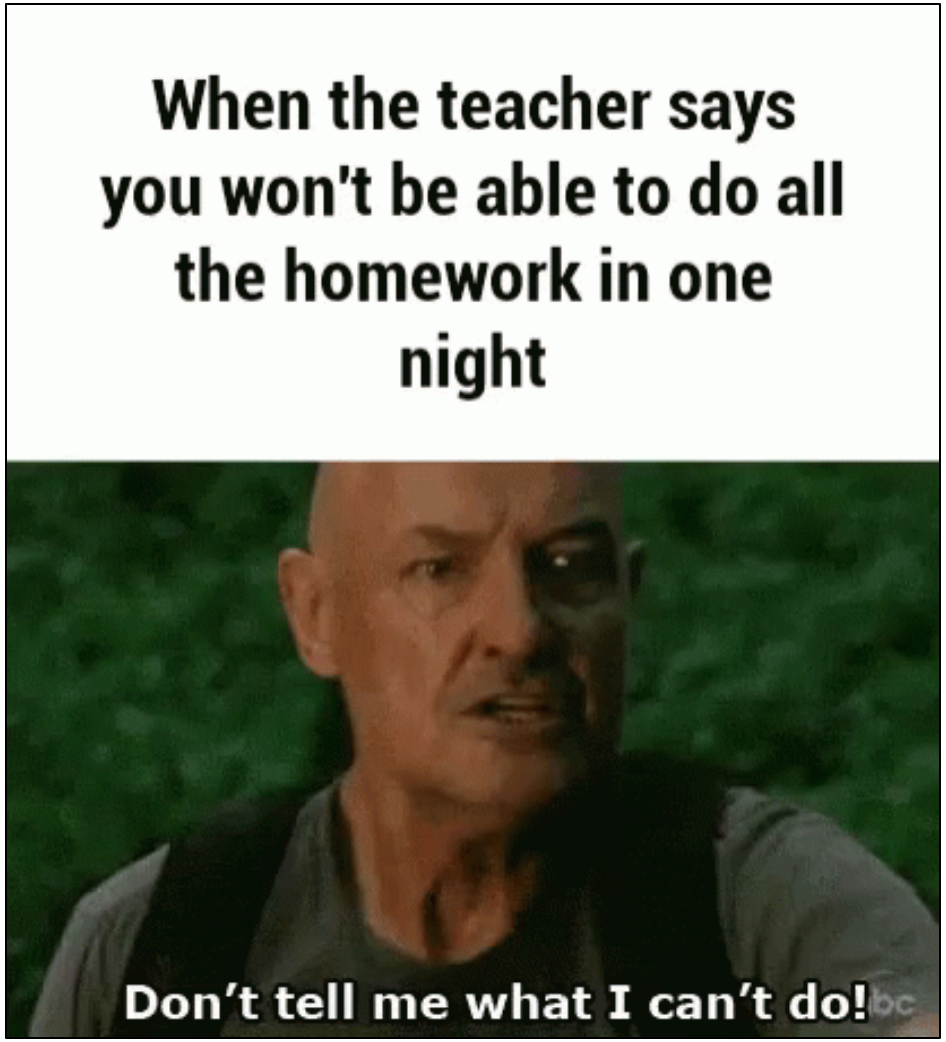
Homework	3-5 assignments including conceptual questions, exercises, and Python implementation and system design. Additionally, 691 students will be required to submit a literature review on a given topic.	40%
Project	Course project in groups of 3-4 (smaller or individual groups only for PhD students with the professor's consent).	40%
	• 10% for project proposal • 5% for midterm presentation • 10% for final presentation • 15% for final report	
Scribing	Each student will typeset lecture notes for one lecture in the semester	5%
Midterm Exam	Date will be announced in class **Tentative Plan: Apr 1**	15%

There is no Final Exam

If you get at least	your minimum grade will be
95	A
80	B
70	C
65	D

Homework

- Conceptual Questions (pen & paper)
- Implementation in Python
  - building cool systems for image blending, panorama, image recognition, generation, etc.
- Literature Review:
  - 4-page Literature Review about the HW Topic
  - CMSC 691: REQUIRED
  - CMSC 491: OPTIONAL (extra credit)



Project


EVERY GROUP PROJECT
DOES 99%
OF THE WORK
HAS NO
IDEA WHAT'S
GOING ON THE
WHOLE TIME
DISAPPEAR
AT THE VERY
UP AGAIN TIL
THE VERY END

Project

Each student will be graded separately

Project Topic

- Pick from our list
- Choose your own *

* *needs approval from Tejas*

Group Size: 3 to 4 students

- *Declare group by March 1*
- *PhD / MS thesis students: can work alone**

* *needs approval from Tejas*

Deliverables

(1) Group Formation	03/01	(2) Project Proposal	03/08
(3) Midterm Update (Video)	04/17	(4) Final Presentation (in class)	05/08
(5) Submit Slides (PDF)	05/17	(6) Submit Report (8 pg. CVPR format)	05/22

Scribing

- All students are required to scribe at least **twice** during the semester.
- You can sign-up for two preferred lectures *(signup sheet: QR code)*
- Scribing = high-quality detailed notes during the lectures in that week, typeset using Overleaf/LaTeX
  (template will be shared, hand-drawn figures are allowed).
- ***Due Dates:***
  - *Notes for Monday lectures are due before class next Monday*
  - *Notes for Wednesday lectures are due before class next Wednesday*

Extra Credit

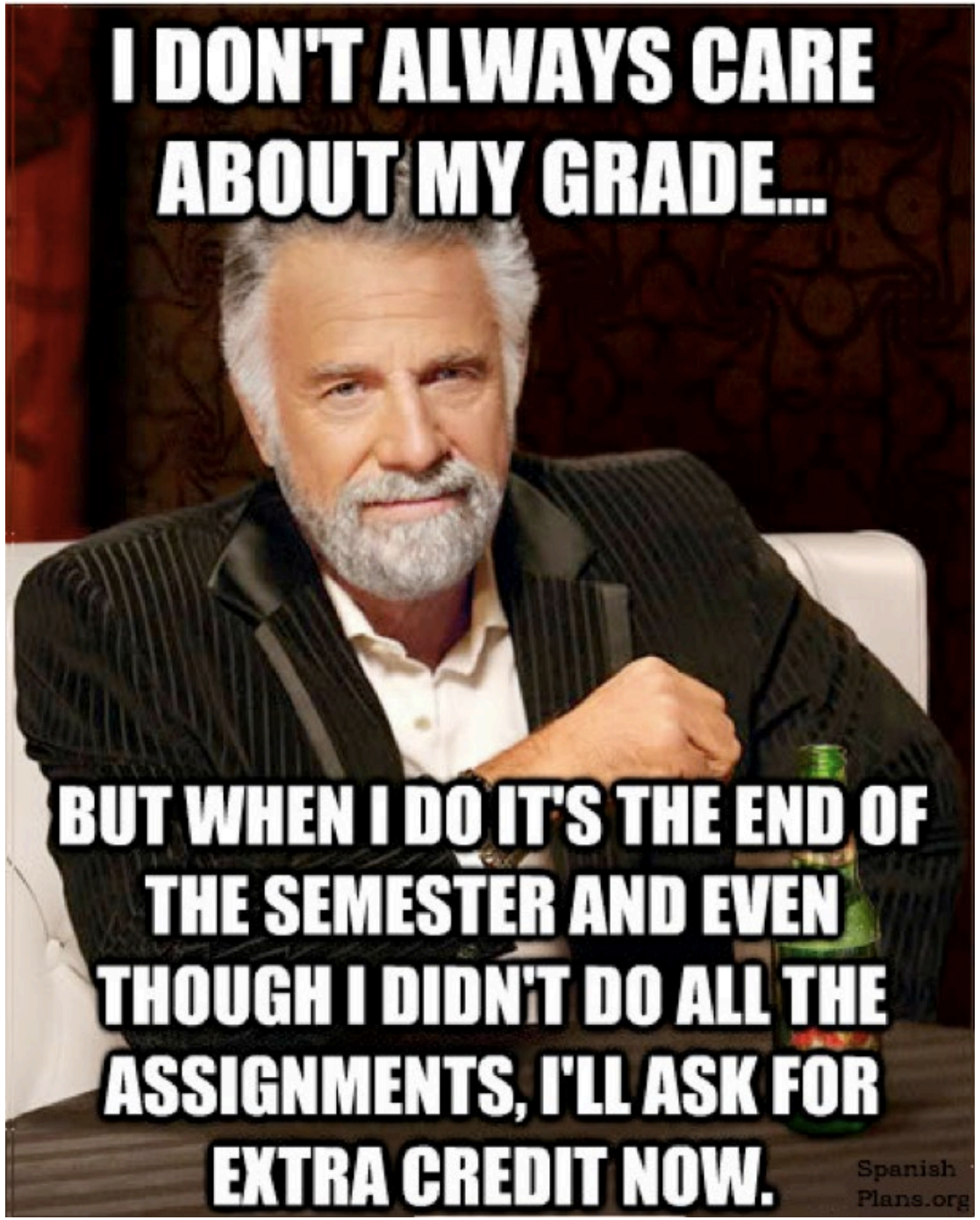
- HW will have some optional problems (but for extra credit!)
  - Open ended questions or tasks where your creativity is required
  - We may showcase best outcomes of these in class
- Classes will have pop quizzes about fundamentals
- Extra credit is capped at 10%
  - 5 extra points and 91 in the rest of the class ➔ your final grade is 96
  - 18 extra points and 91 in the rest of the class ➔ your final grade is 101

Deadlines & Late Days

- Each homework and project component will have a deadline.
- Late Days: each student will get 7 late days
  - Each late day extends the deadline by 24 hours
  - Using a late day does not influence the grade.
  - Can be used for homeworks and scribing only
  - Late submissions turned in *after all 7 late days have been exhausted* **will not be evaluated and will receive 0 points.**
- Late days are provided to help you deal with illness or injury, personal emergencies, paper deadlines, interviews, and computer problems.
- Do not use the late days as an excuse for procrastination ☺

Academic Integrity

- Homework must be done independently. Sharing your work constitutes cheating.
- Use of "AI" writing assistants is prohibited.
- Consult UMBC's policy on plagiarism and other forms of cheating: https://academicconduct.umbc.edu/resources-for-students/
- See the syllabus and course website for consequences



Seek Help Early !

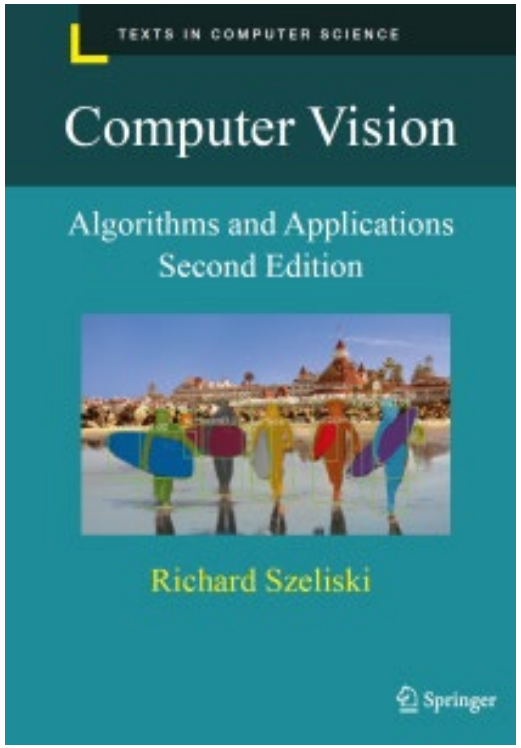
Attending Classes

- Attendance is mandatory
  - Exceptions: health reasons and personal emergencies.
  - Impossible to do a good job at scribing, homework, and midterm (60% of your grade) without your attendance and attention in class.
- Perks:
  - Classes will have pop quizzes for extra credit -- I wouldn't miss that opportunity!
  - Meet your future project team – new friends – I'm still in touch with my CV project teammates from 2016!
  - *Do you really want to miss these memes?*

Recommended Background

- Linear algebra + calculus + geometry + prob/stats (required)
  - 491/691 should not be your first introduction to these topics
  - Without these tools, you are likely to struggle with the course.
- Python programming with numerical libraries like `numpy`
  - Homework 1 will allow you to quickly learn OpenCV basics
  - TA will give a tutorial on computer vision with PyTorch
- Useful resources to brush up on these topics
  - deeplearningbook.org/contents/linear_algebra.html
  - https://www.deeplearningbook.org/slides/02_linear_algebra.pdf
  - deeplearningbook.org/contents/prob.html
  - https://www.deeplearningbook.org/slides/03_prob.pdf

Readings

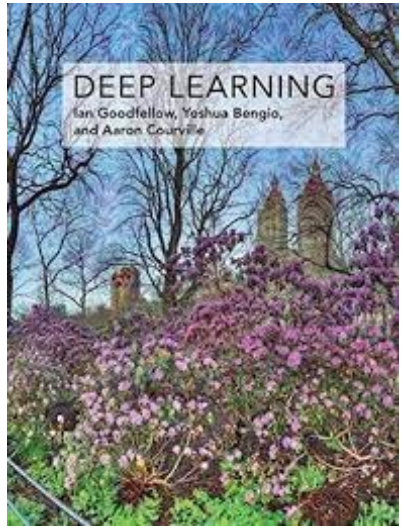


https://szeliski.org/Book/

Free download

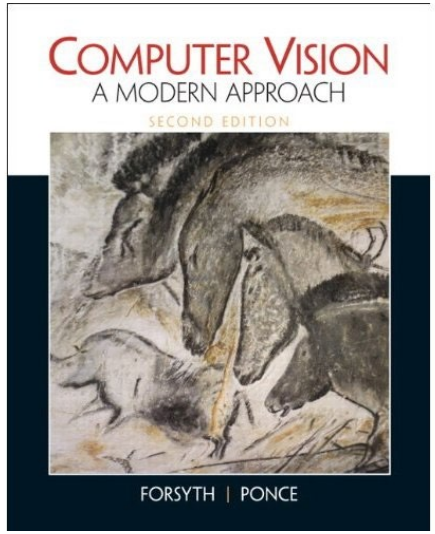
Topic-specific reading will be provided (pdf)

Other useful resources:



www.deeplearningbook.org/

Free download



GPU Computing

- Homeworks do not require GPUs
- Projects (depending on the topic you choose) might need GPUs
- GPUs are very expensive. I DO NOT EXPECT YOU TO UPGRADE YOUR COMPUTER JUST FOR THIS CLASS !!!
- Google Colab is your friend
  - Free, but has usage limits (*per email address*)
  - You can consider *purchasing* Colab pro (but it is not a requirement for the class).
- More updates later on during the semester

CMSC 491 vs 691

- 491 is the undergrad version
  691 is the grad version
- We are confident that 491 students are as capable (if not more so) than their 691 classmates
- *No difference in class materials, exams, quizzes, and scribing expectations*
- Homework: additional parts for 691
- Grad projects will be evaluated at a higher standard

Any Questions so Far?

Timeline of Buzzwords in Vision

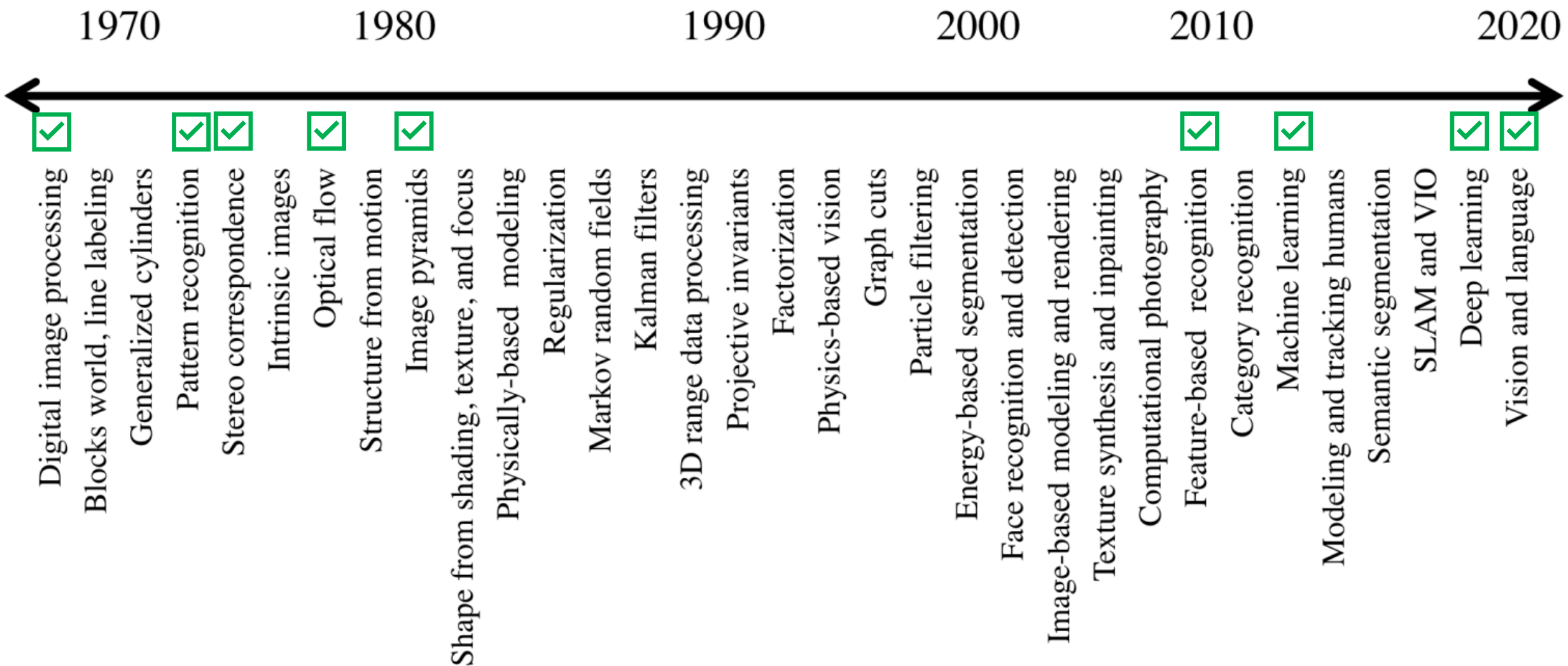
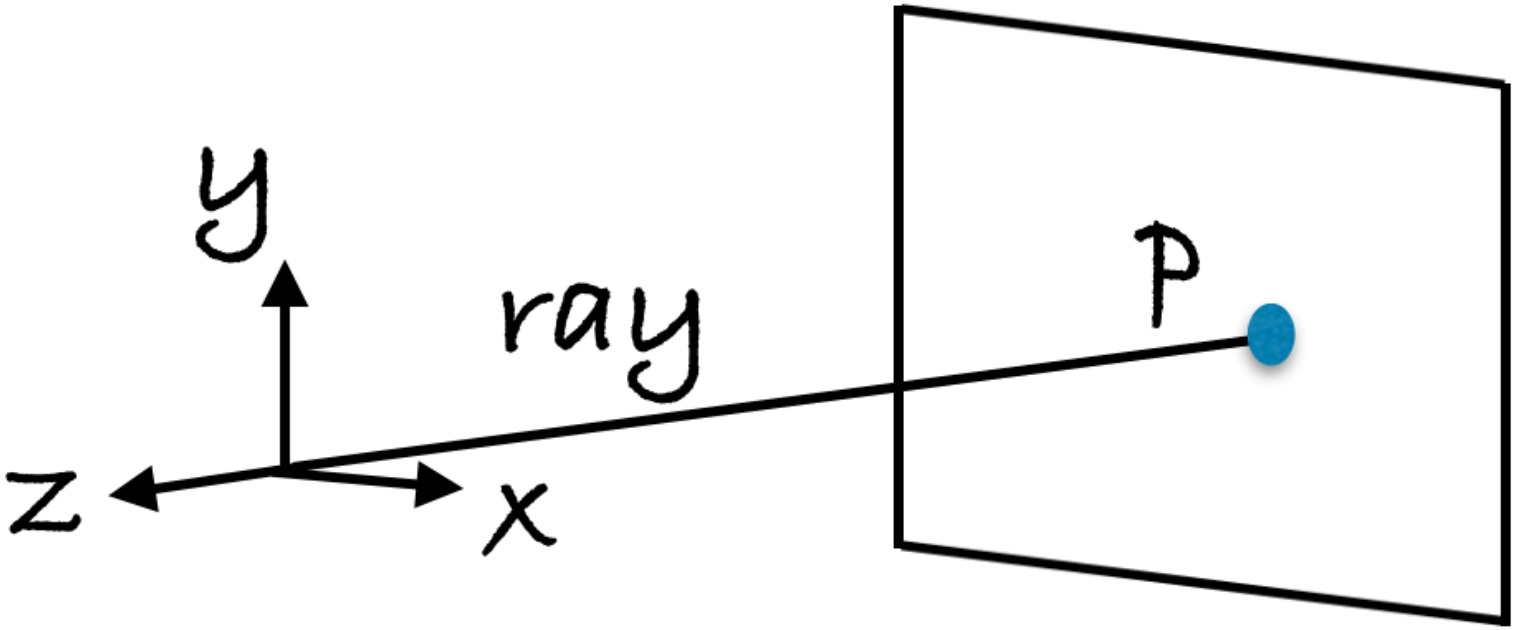


Image Formation



Let’s say we have a sensor…

digital sensor
(CCD or CMOS)

… and an object we like to photograph

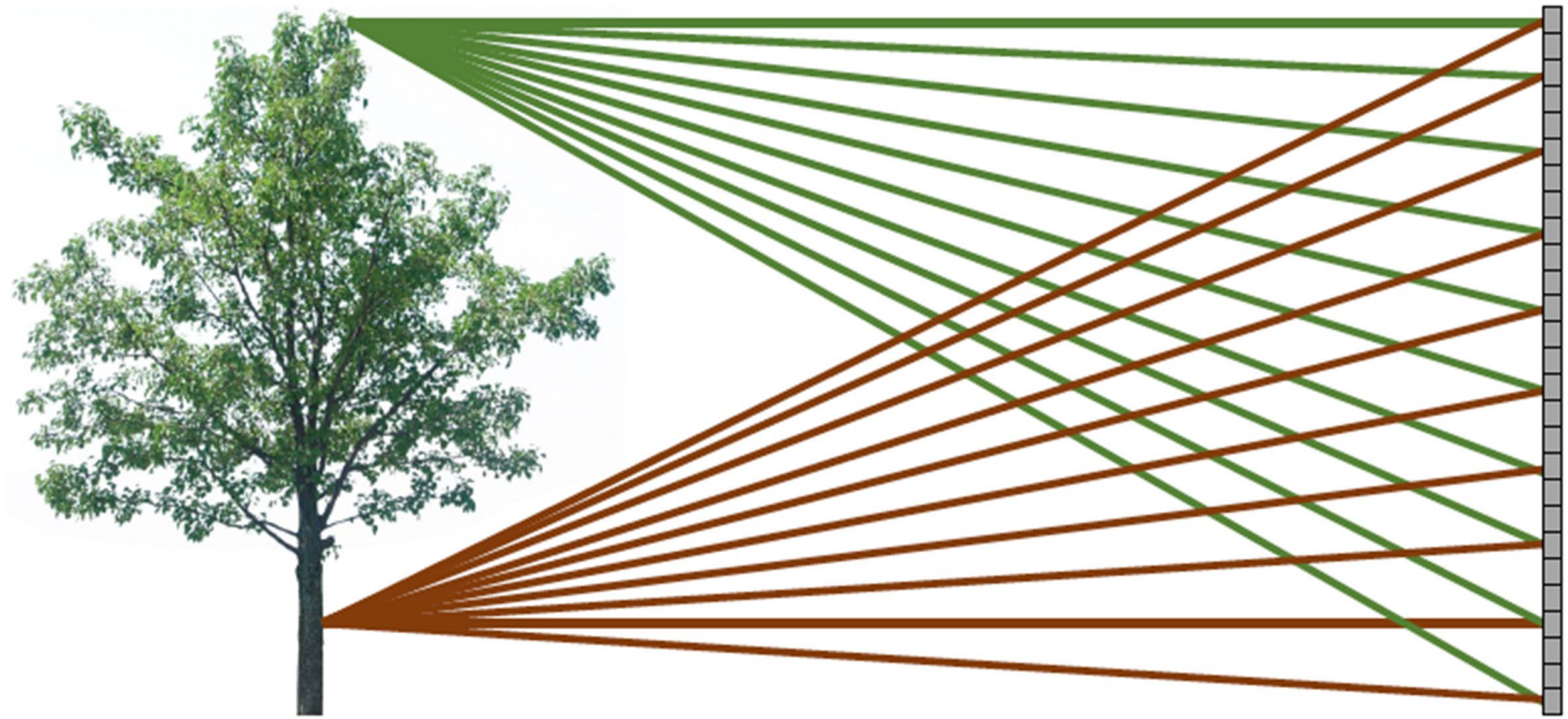
real-world object

digital sensor (CCD or CMOS)

What would an image taken like this look like?

Bare-sensor imaging

real-world object



digital sensor (CCD or CMOS)

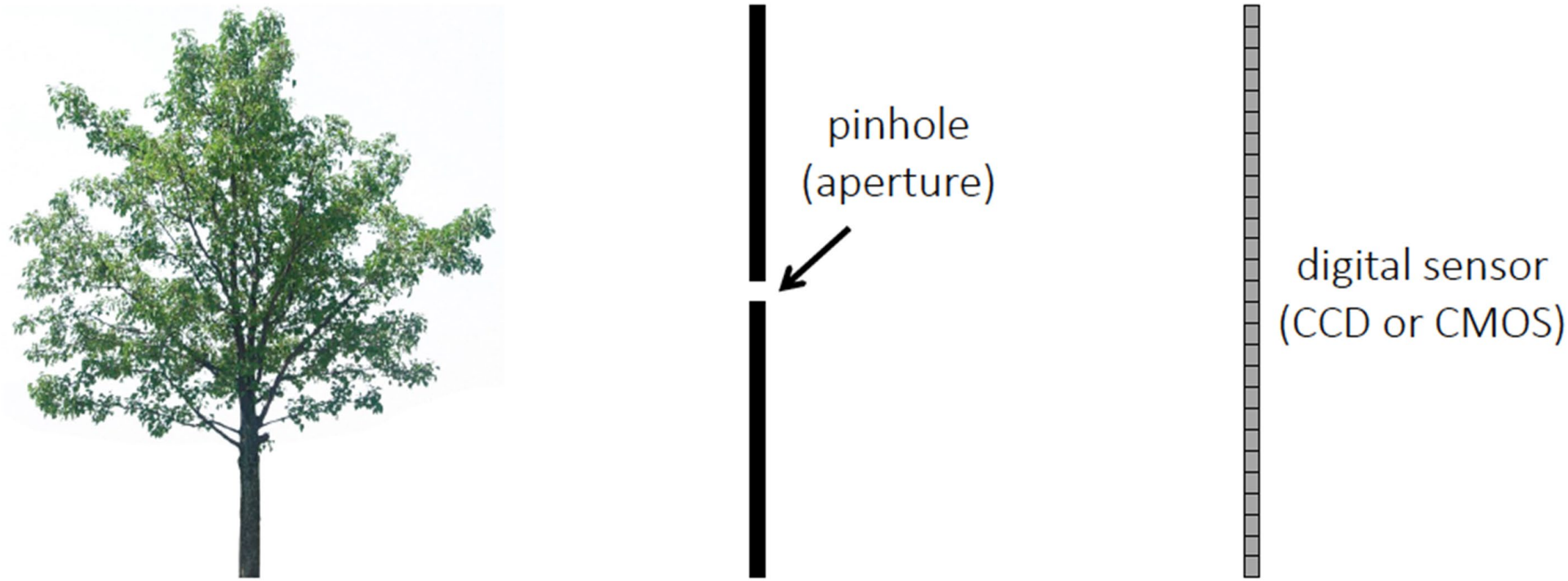
All scene points contribute to all sensor pixels

What does the image on the sensor look like?

Bare-sensor imaging

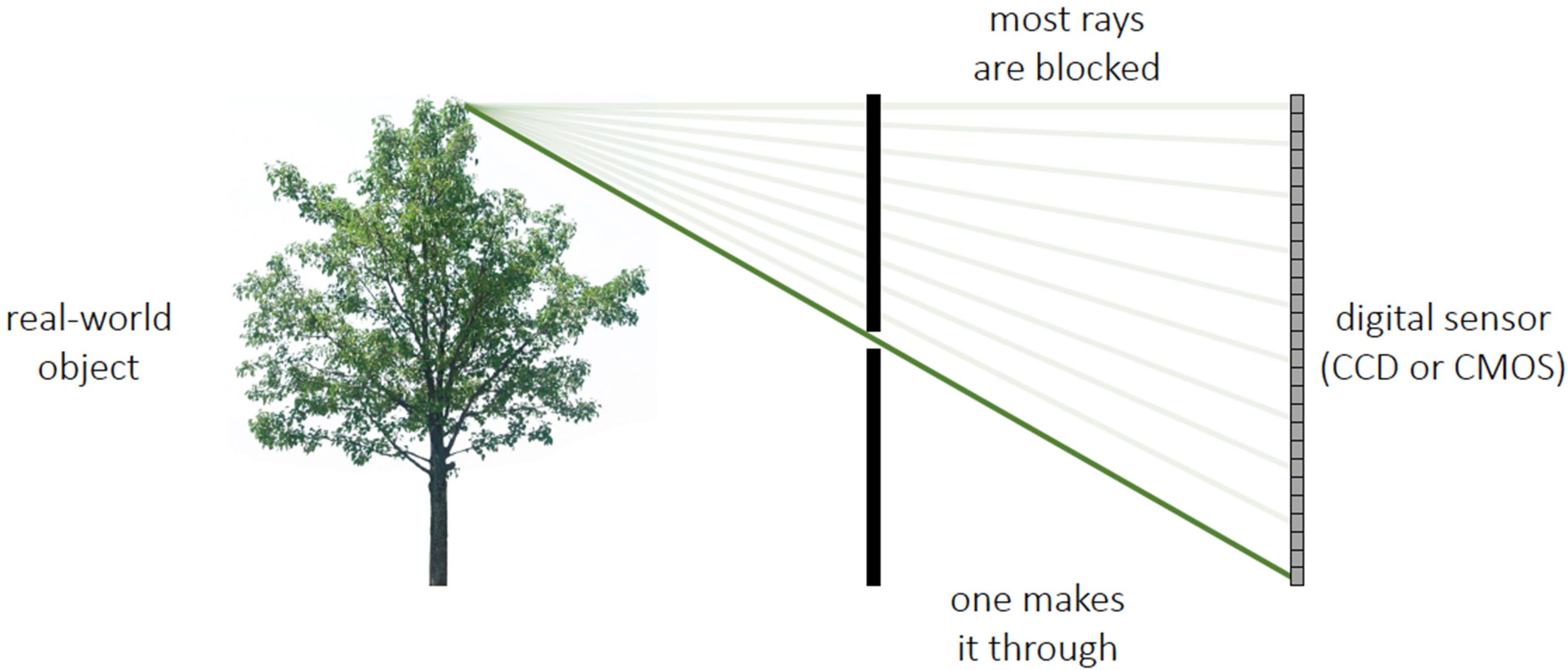
All scene points contribute to all sensor pixels

Let’s add something to this scene

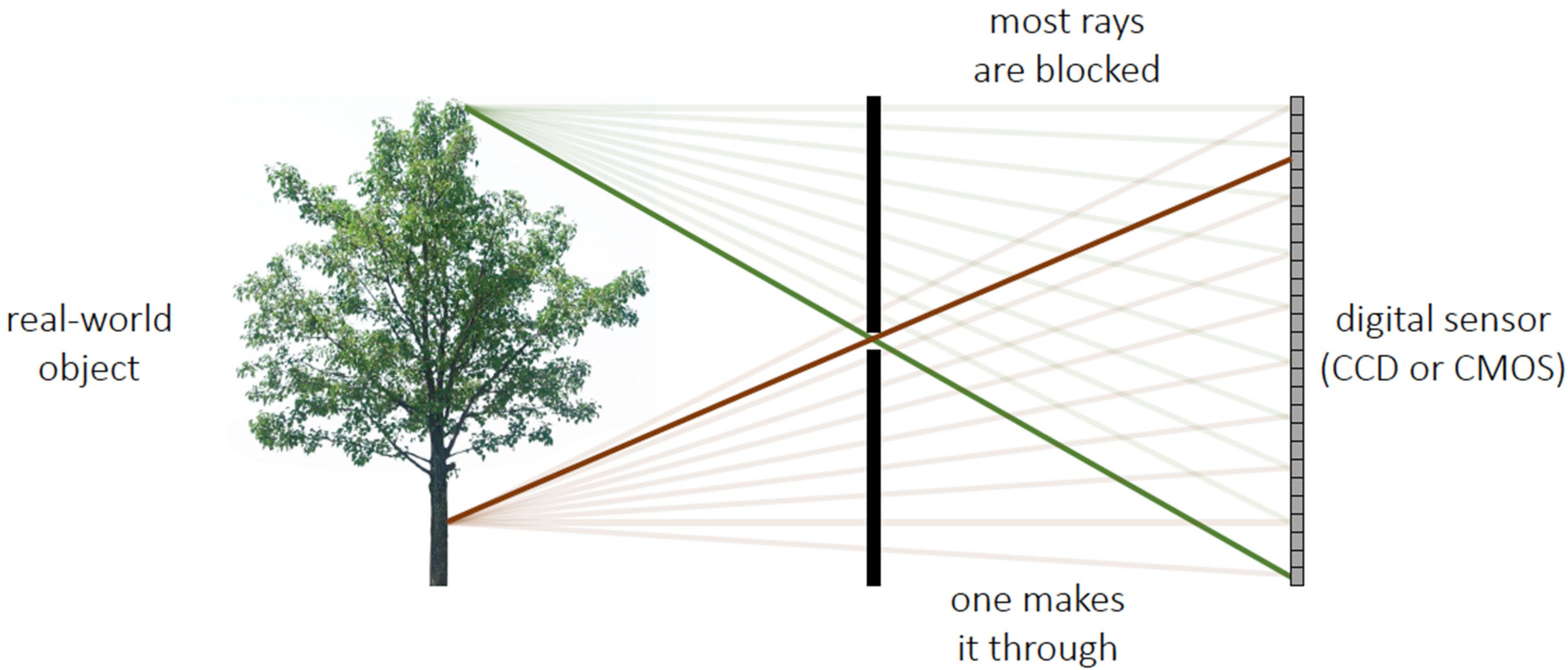


What would an image taken like this look like?

Pinhole imaging

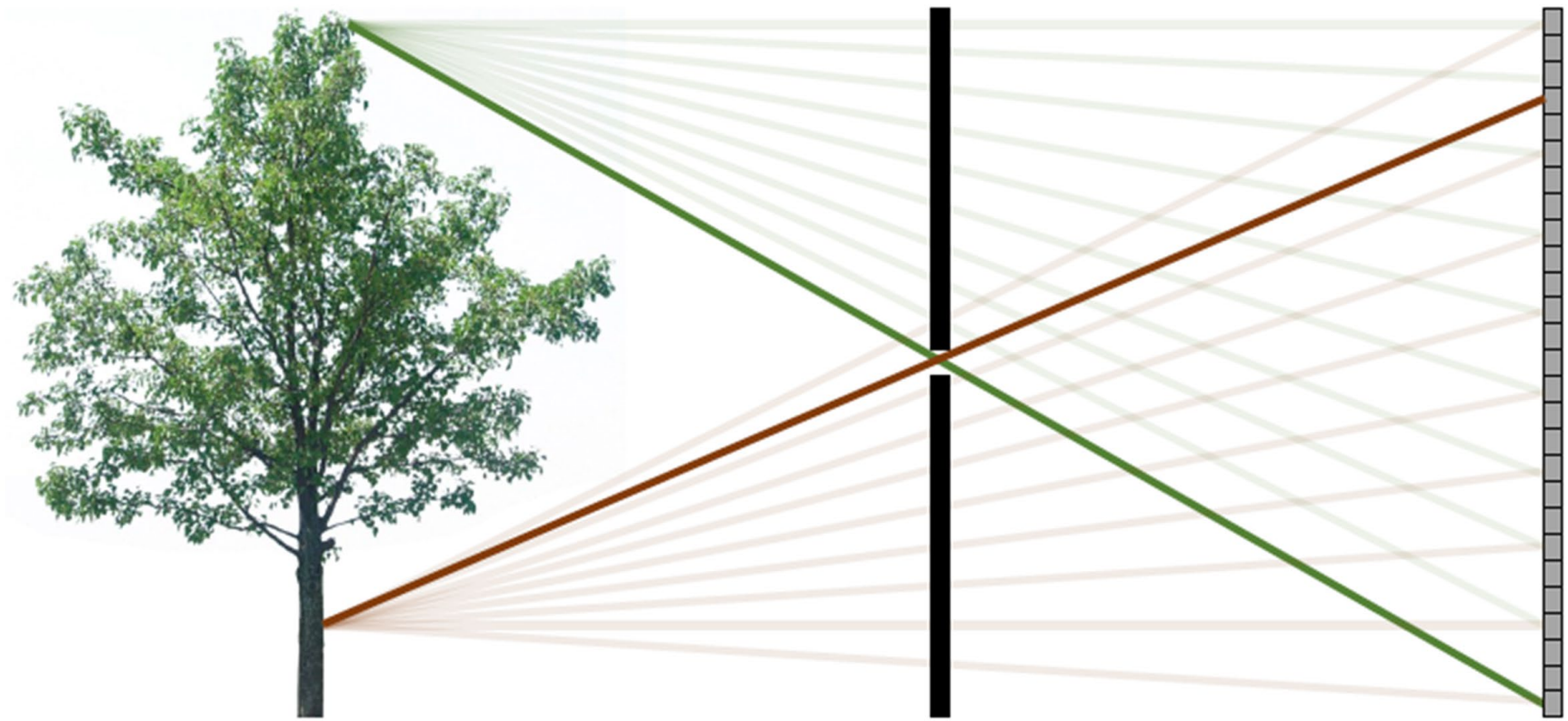


Pinhole imaging



Pinhole imaging

real-world object



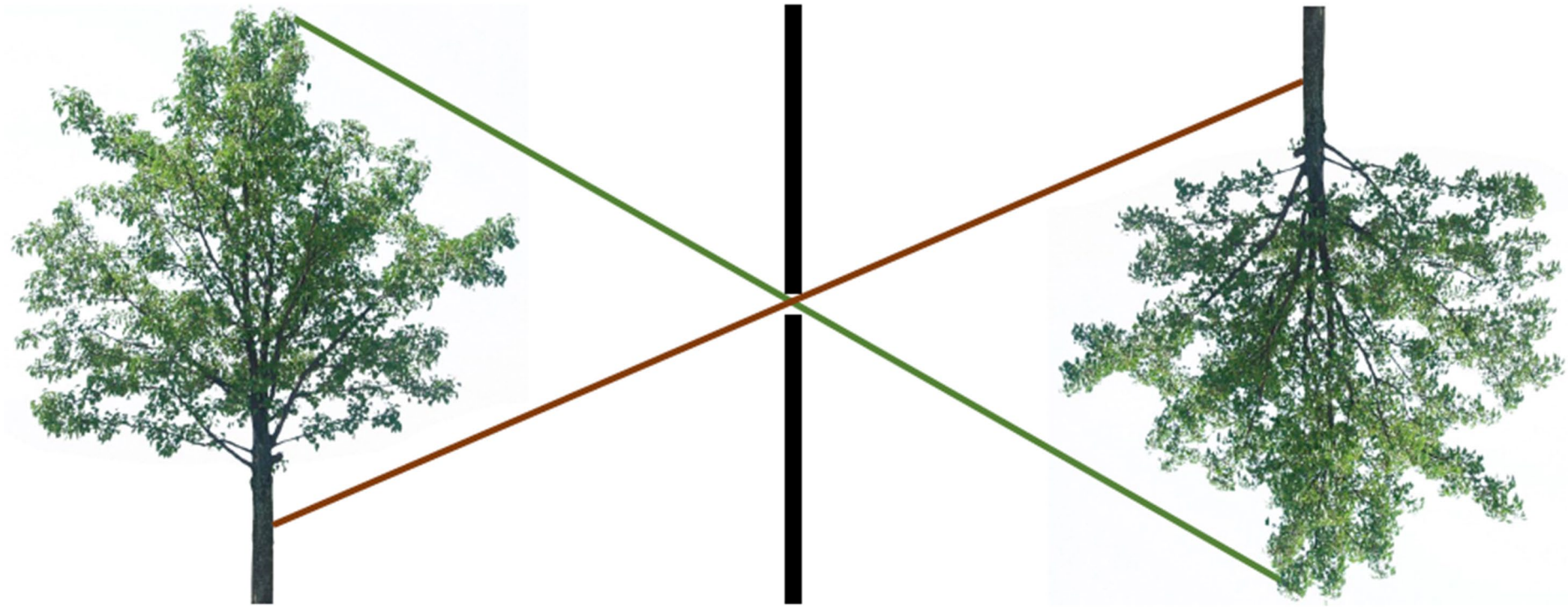
digital sensor (CCD or CMOS)

Each scene point contributes to only one sensor pixel

What does the image on the sensor look like?

Pinhole imaging

real-world object



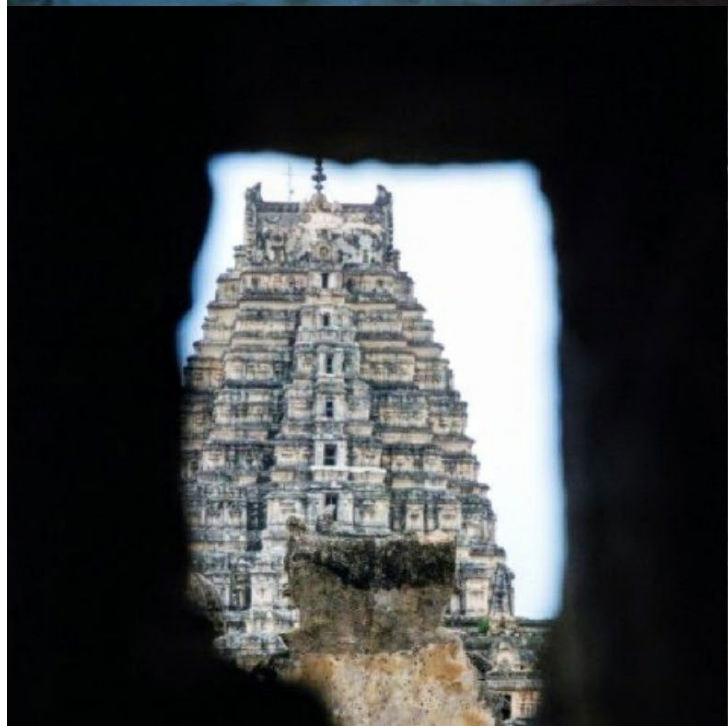
copy of real-world object (inverted and scaled)

“Camera Obscura”

- Camera obscura = “dark chamber” in Latin
- Theorized to be the reason for distortions in prehistorical neolithic paintings …
- Mozi (China, ~ 400 BC)
- Aristotle (Greece, ~ 350 BC)

Ancient Architecture or Accidental Pinhole Camera?

- Shiva Temple in Hampi (India)
- inverted image can be seen
- Built ~ 600 AD



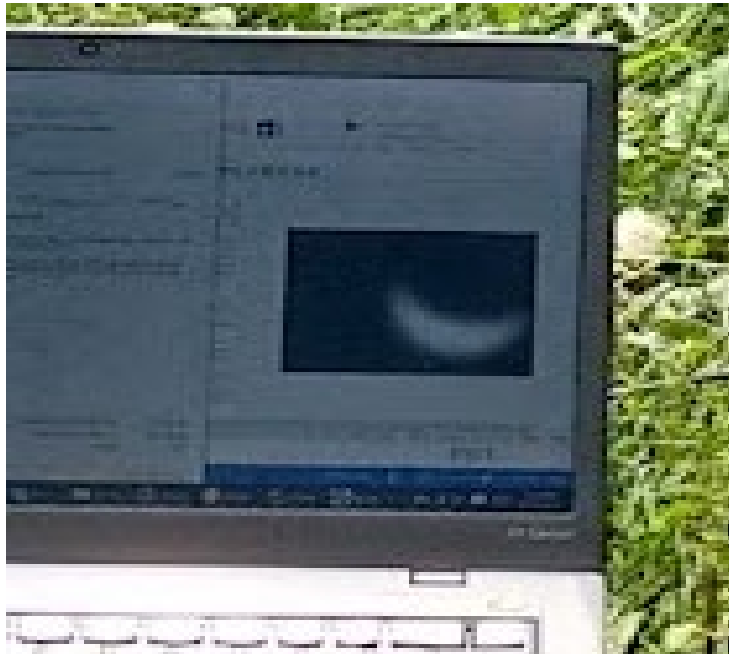
Make your own pinhole camera while the sun shines (Ancient Computer Vision Proverb)

Paper with a hole in it … some creative friends … an auspicious day …

Where are they now? Faculty at UMBC, Faculty at UC Riverside; Fellow at Allen Institute; Research Scientist at Adobe …

“Make your own pinhole camera while the sun shines” (Ancient Computer Vision Proverb)

“Great American Eclipse” of August 21, 2017



“Make your own pinhole camera while the sun shines” (Ancient Computer Vision Proverb)

Images produced by natural pinholes [edit]

(Images of the eclipse created by natural pinholes formed by tree leaves)

North Cascade mountains (British Columbia and Washington).

East Wenatchee, Washington

Moon, Pennsylvania

Pinhole camera terms

real-world object

barrier (diaphragm)

pinhole (aperture)

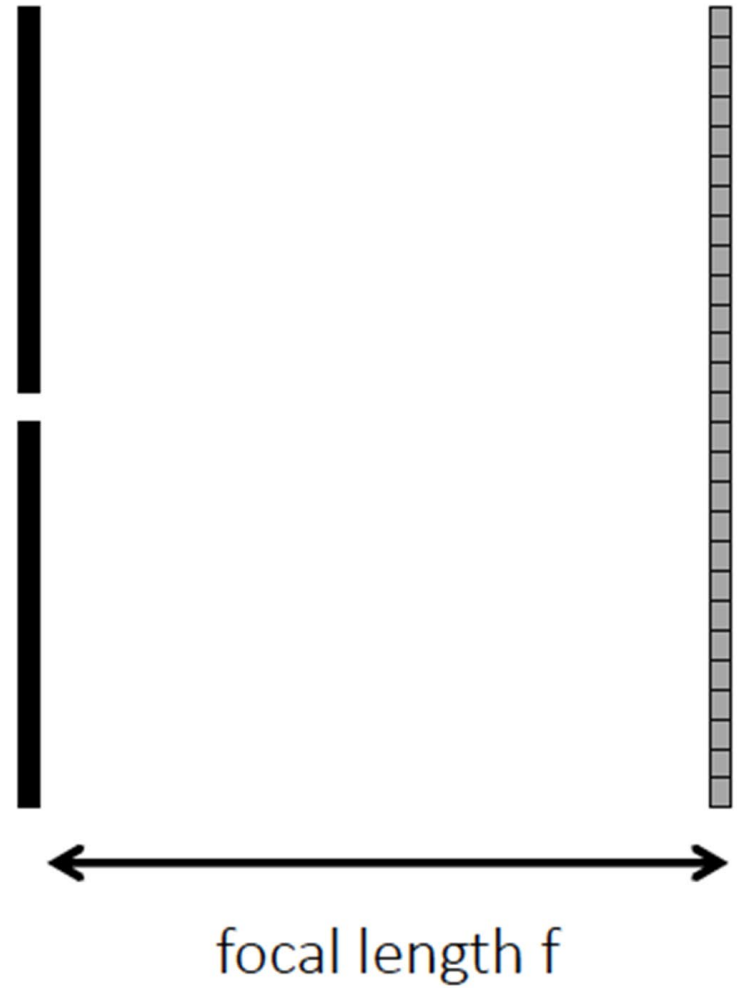
camera center (center of projection)

image plane

digital sensor (CCD or CMOS)

Focal length

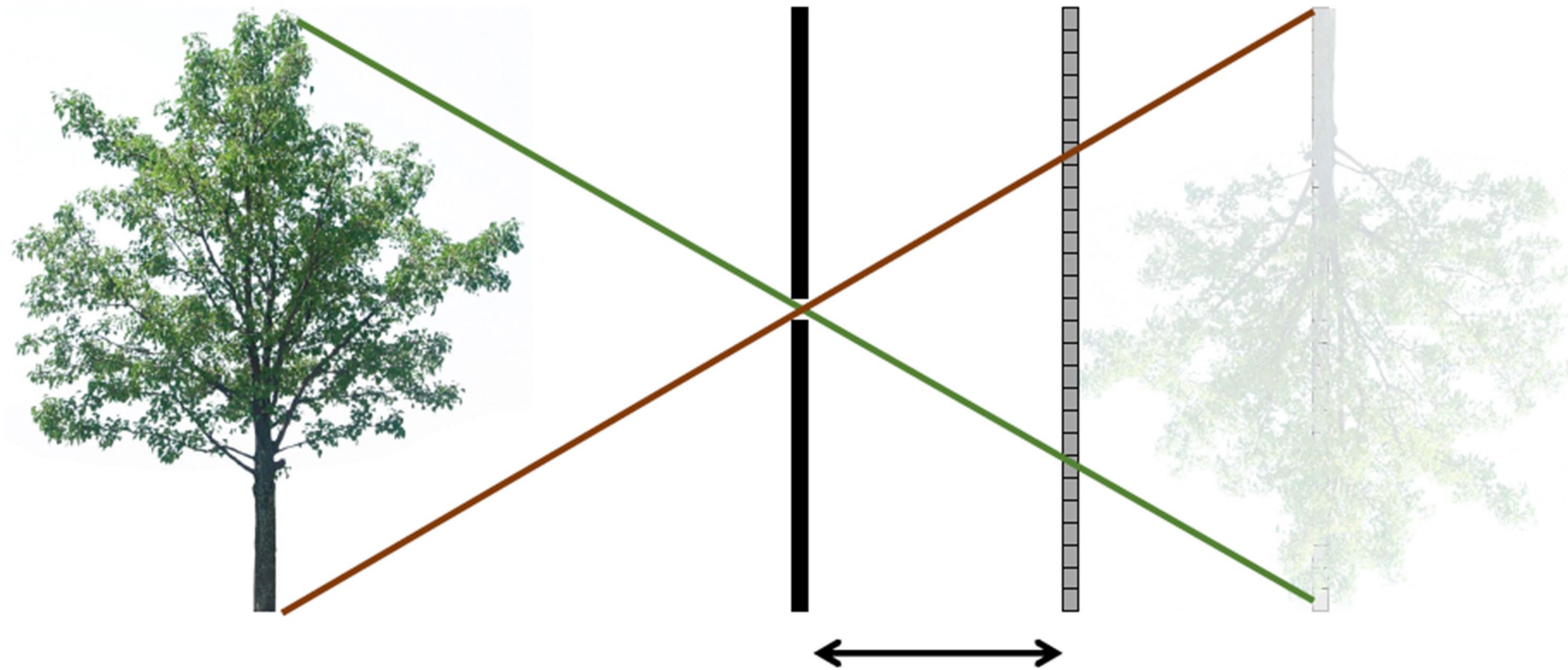
real-world object



Focal length

What happens as we change the focal length?

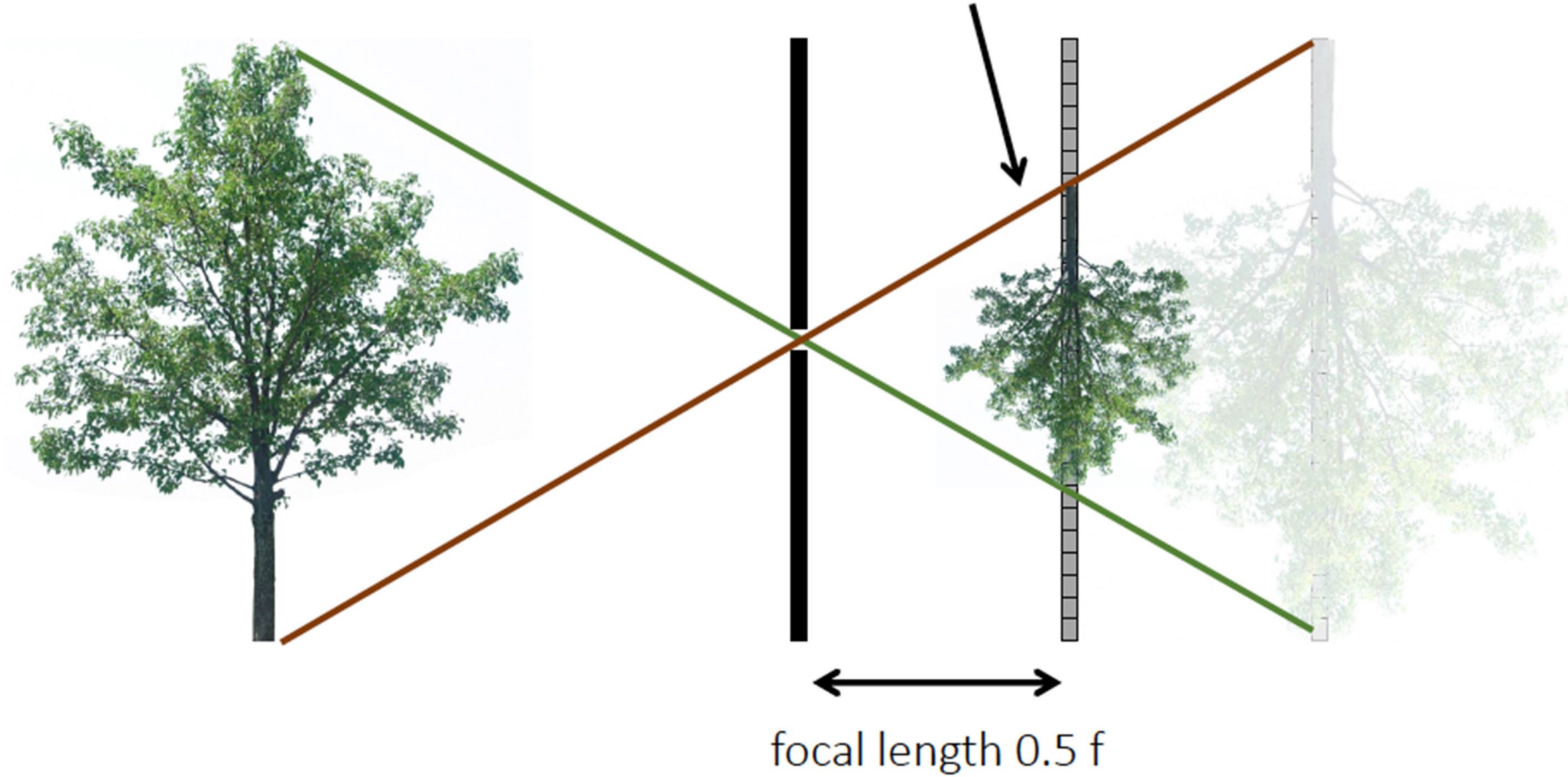
real-world object



focal length 0.5 f

Focal length

What happens as we change the focal length?



Pinhole size

What happens as we change the pinhole diameter?

real-world object

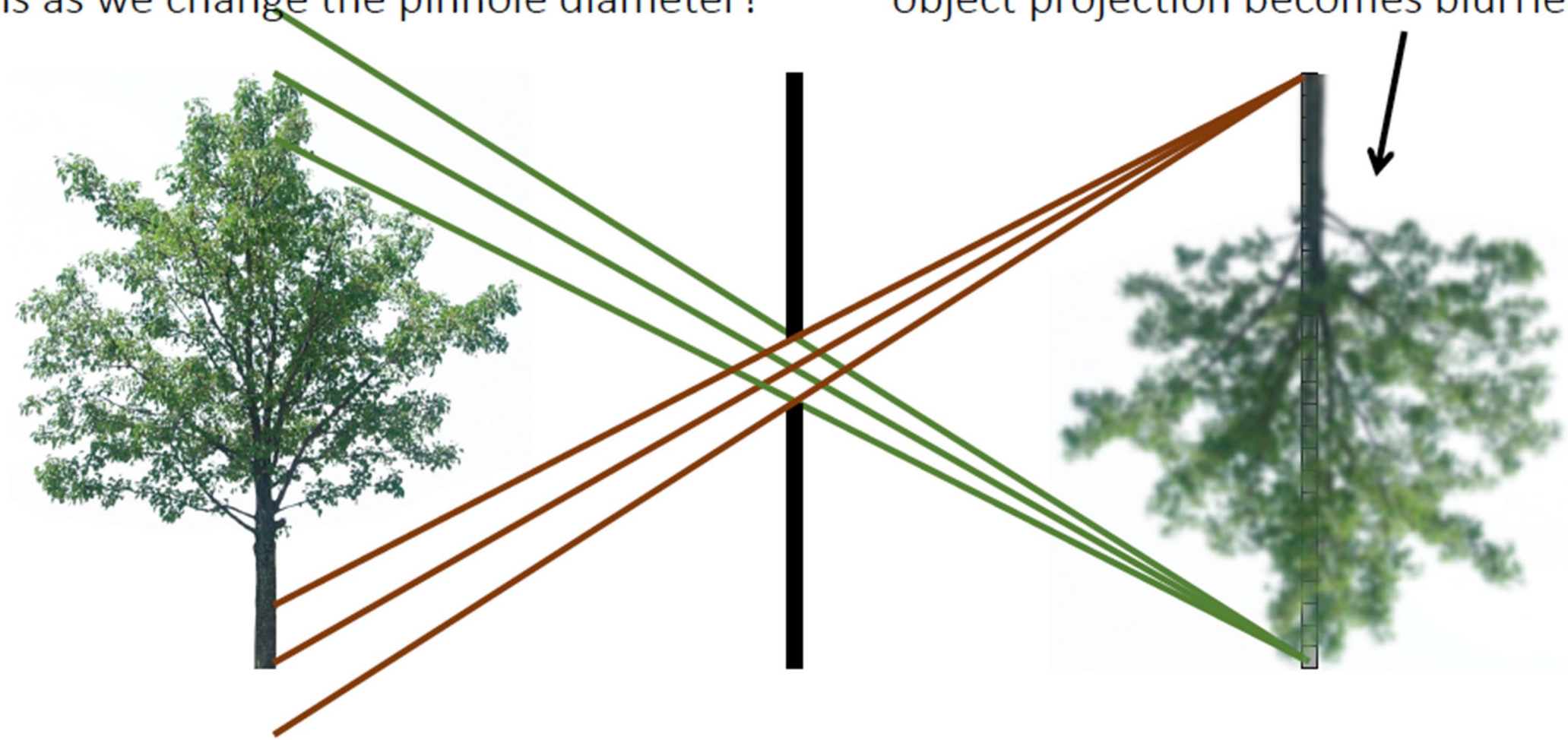
pinhole diameter

Pinhole size

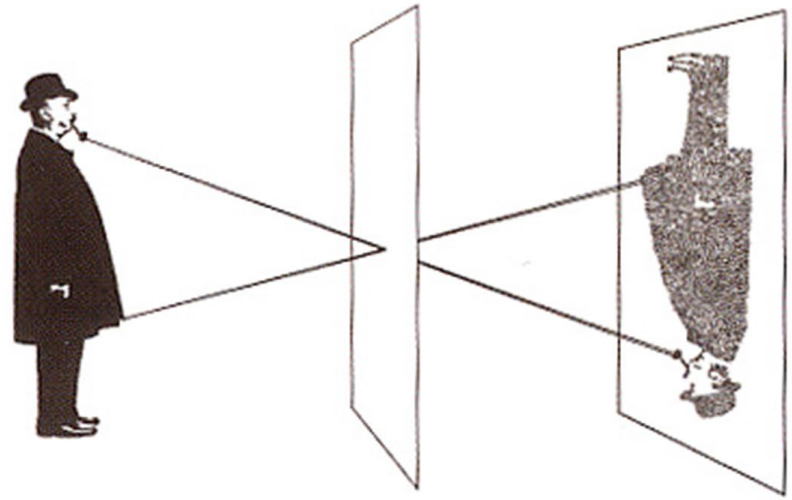
What happens as we change the pinhole diameter?

object projection becomes blurrier

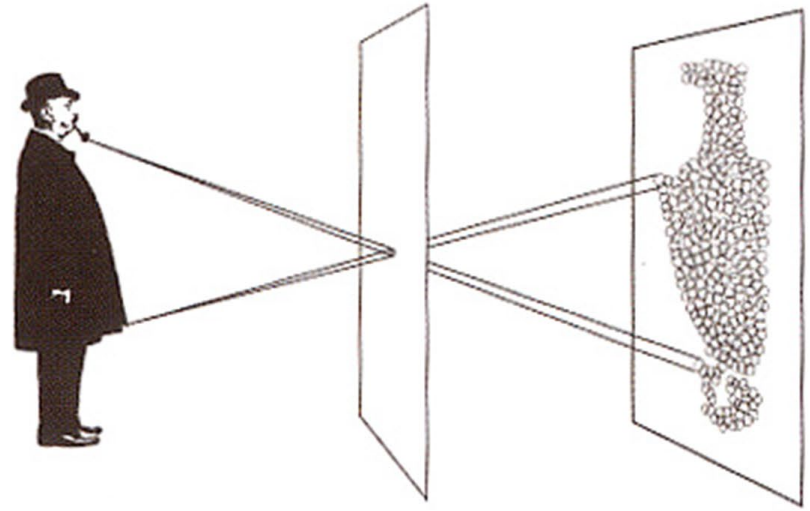
real-world object



Photograph made with small pinhole



Photograph made with larger pinhole



Problems with Pinholes

- Pinhole size (aperture) must be "very small" to obtain a clear image.
- However, as pinhole size is made smaller, less light is received by image plane.
- If pinhole is comparable to wavelength λ of incoming light, DIFFRACTION blurs the image!
- Sharpest image is obtained when:

pinhole diameter $d = 2\sqrt{f'\lambda}$

Example: If f' = *50mm,*

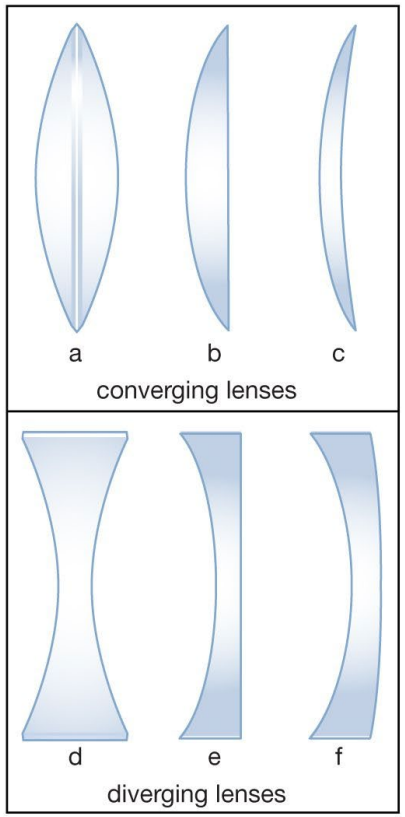
= *600nm* (red),

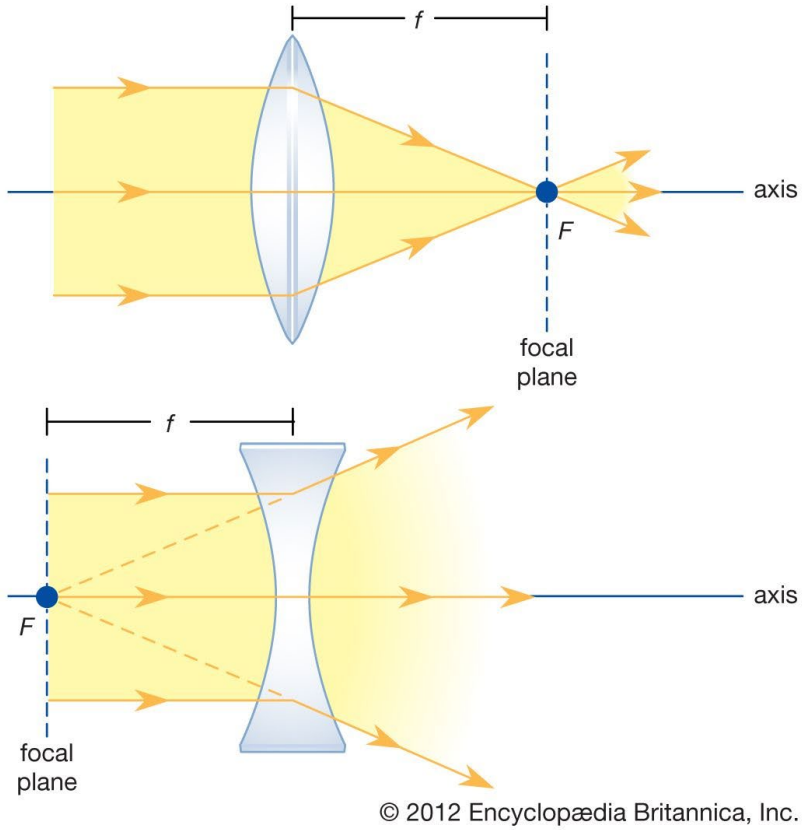
d = *0.36mm*

Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]

Pinholes have Problems … What's the Solution?

Lenses!

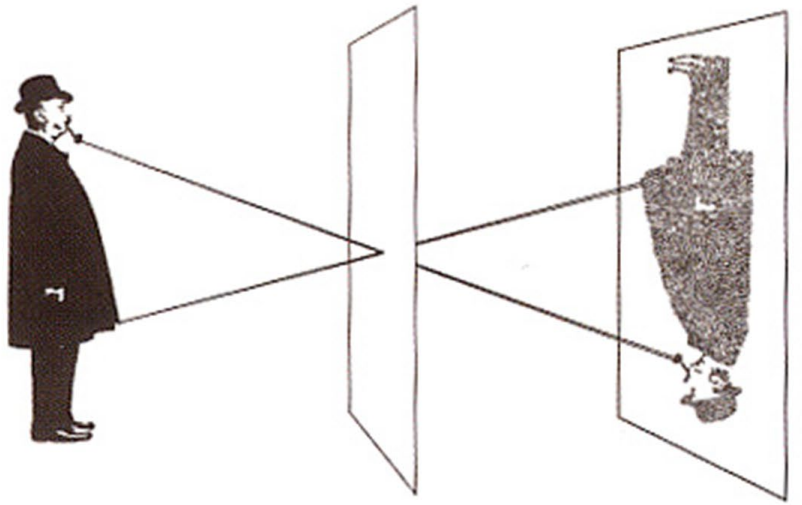




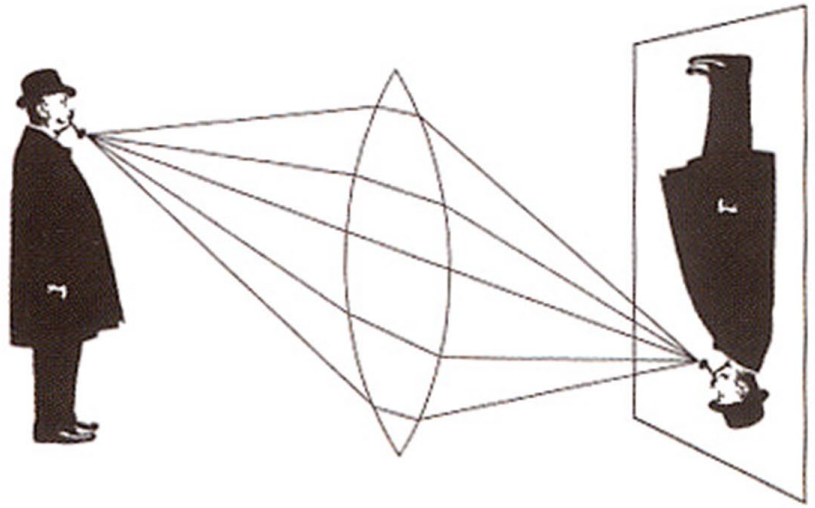
Lenses are Cool

We'll study them in next class …

Photograph made with small pinhole



Photograph made with lens



Frequently Asked Questions

How do I succeed in this class?

- Use pen-and-paper. Draw! Draw! Draw!
- Computer vision is a very popular *and open-source* topic
  - Feel free to read or watch lectures from other professors
- Focus on grasping fundamentals!
  - HW will become easy once you know the concepts behind the problems and projects.
- Ask for help
  - TA Office Hours
  - Tejas Office Hours
  - Google Chat Space

Frequently Asked Questions

Can I join your research lab?

- Joining (See FAQ on my website)
  - Take this class and talk to me during office hours about your interests
  - Computer Vision has a low barrier to entry – all motivated STEM majors have enough background to start learning.
    (CMPE, CMSC, ENEE, MATH, PHYS, STAT are probably closest aligned)
- Will I get paid for research ?
  - Depends (I currently only have funding for PhD students)
  - Undergraduate Research & Prestigious Scholarships - UMBC
  - CWIT Scholars – Center for Women in Technology – UMBC
  - You can also do research for credit (e.g. CMSC 299, 499, 698, 699)

More Questions?

- I’m here till 5:30 PM today.