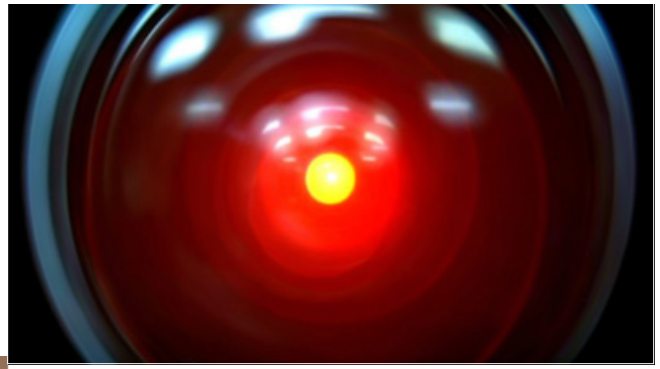
Lecture 1:
Introduction to “Computer Vision”

Professor Fei-Fei Li

Stanford Vision Lab

Welcome to CS231a: Computer Vision



Slide adapted from Svetlana Lazebnik

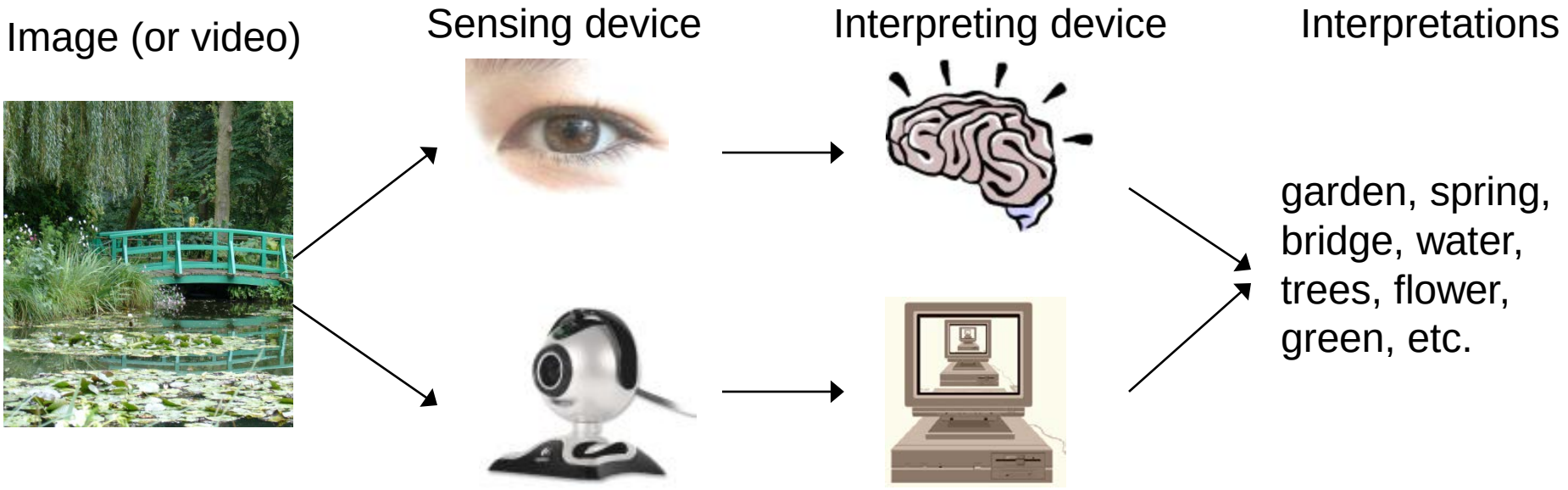
Today's agenda

- Introduction to computer vision
- Course overview

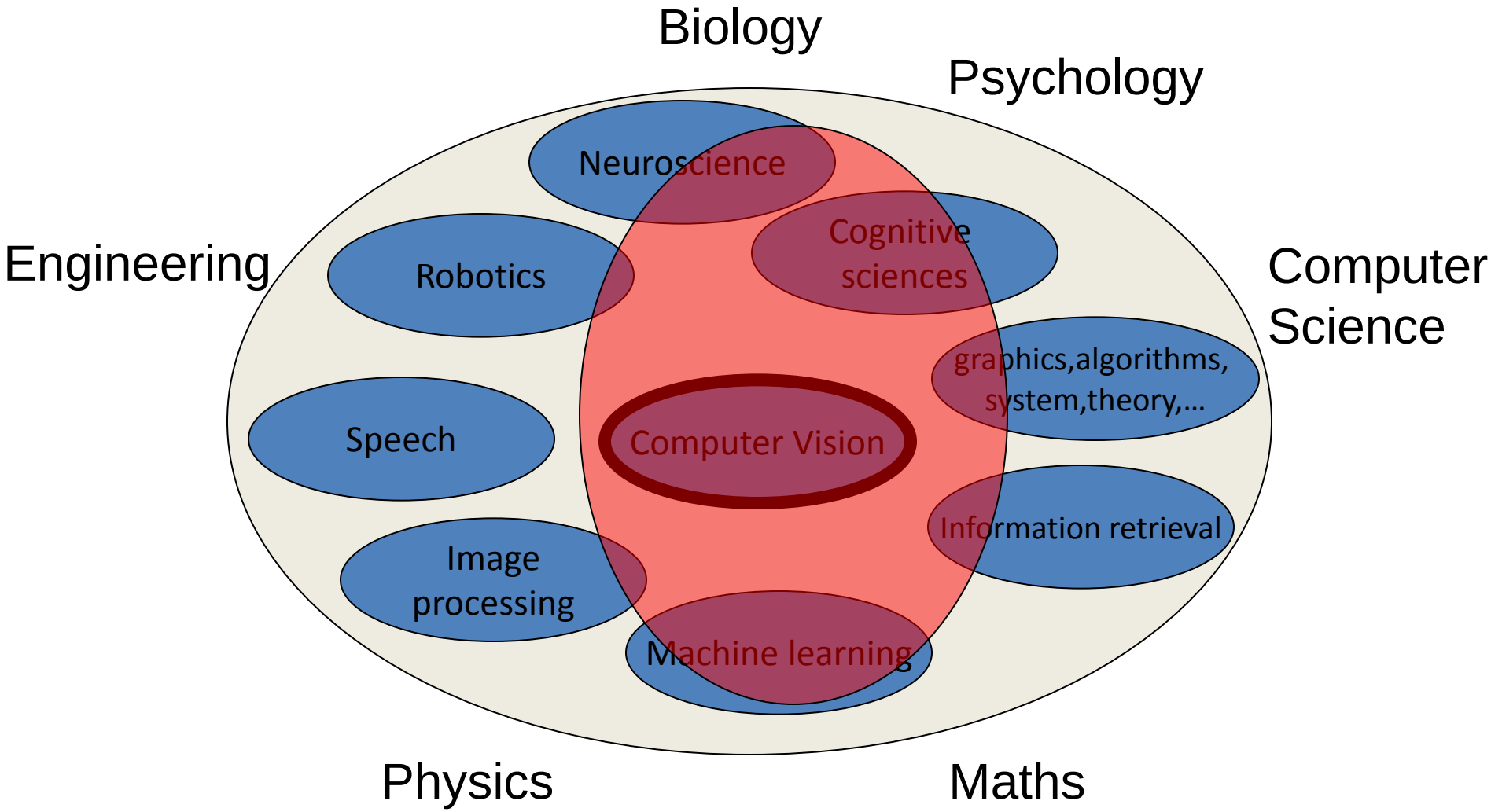
Quiz?

What about this?

What is (computer) vision?



What is it related to?



The goal of computer vision

- To bridge the gap between pixels and “meaning”

What we see

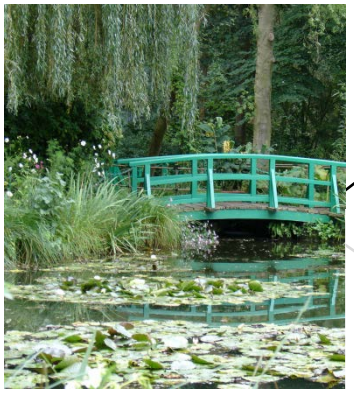
0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

What is (computer) vision?

Image (or video)



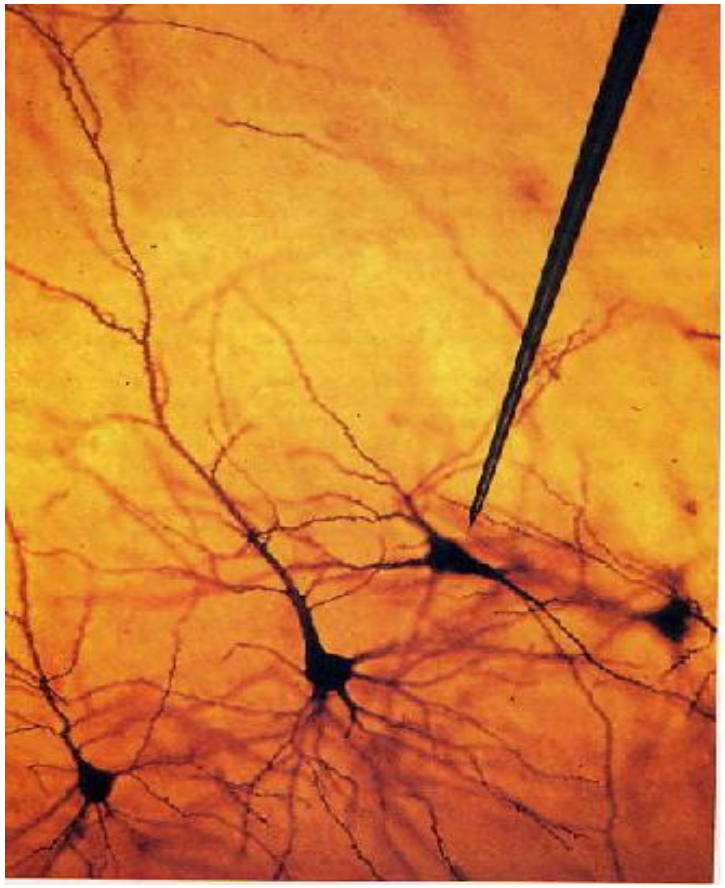
Sensing device

Interpreting device

Interpretations

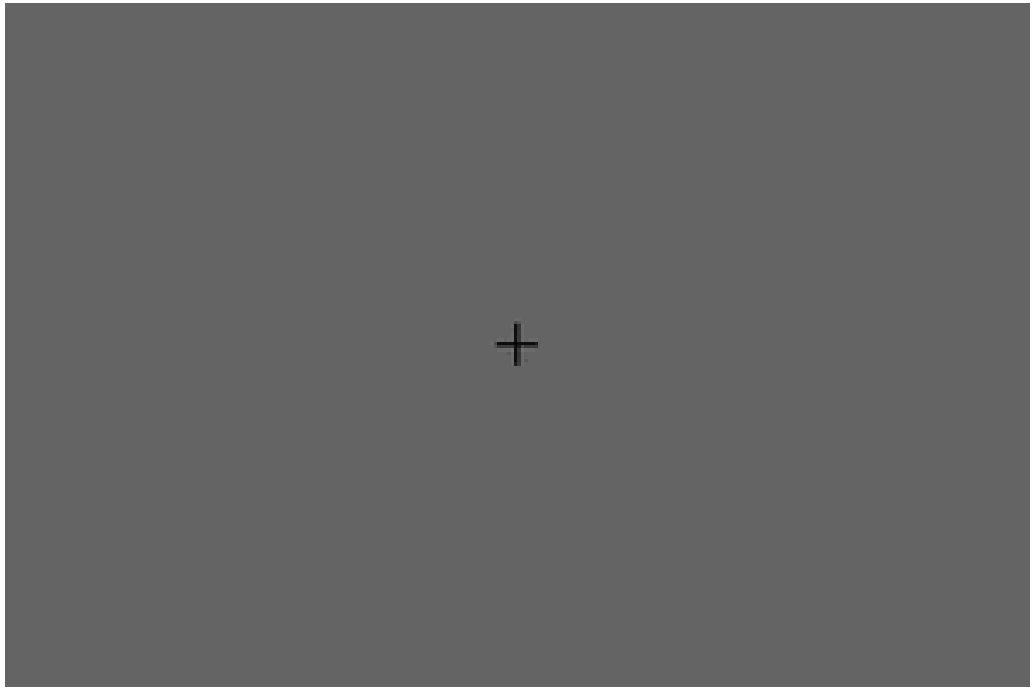
garden, spring, bridge, water, trees, flower, green, etc.

1981: Nobel Prize in medicine



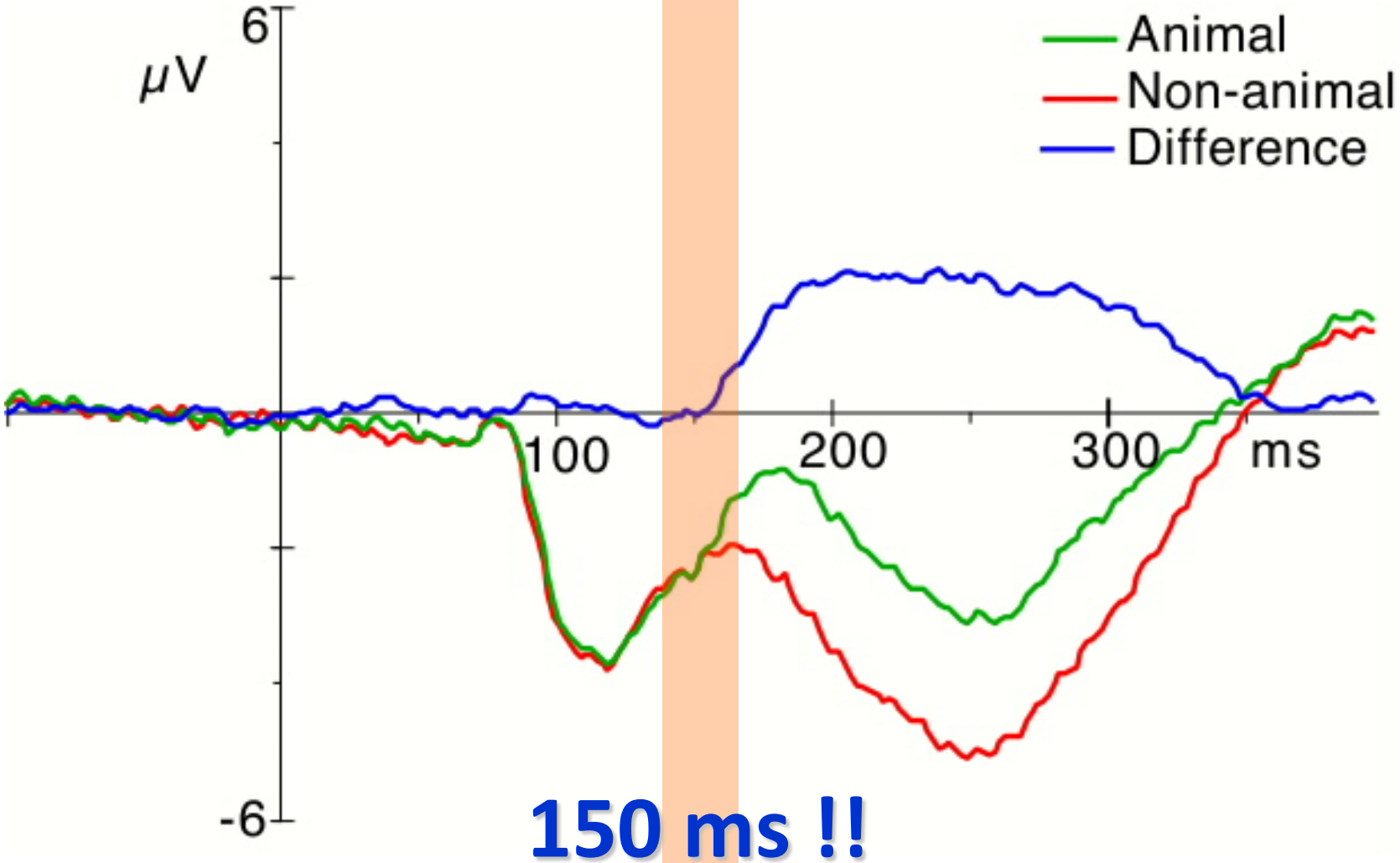
Hubel & Wiesel

Human vision is superbly efficient



Potter, Biederman, etc. 1970s

Thorpe, et al. *Nature,* 1996



150 ms !!

Thorpe, et al. *Nature,* 1996

Change blindess

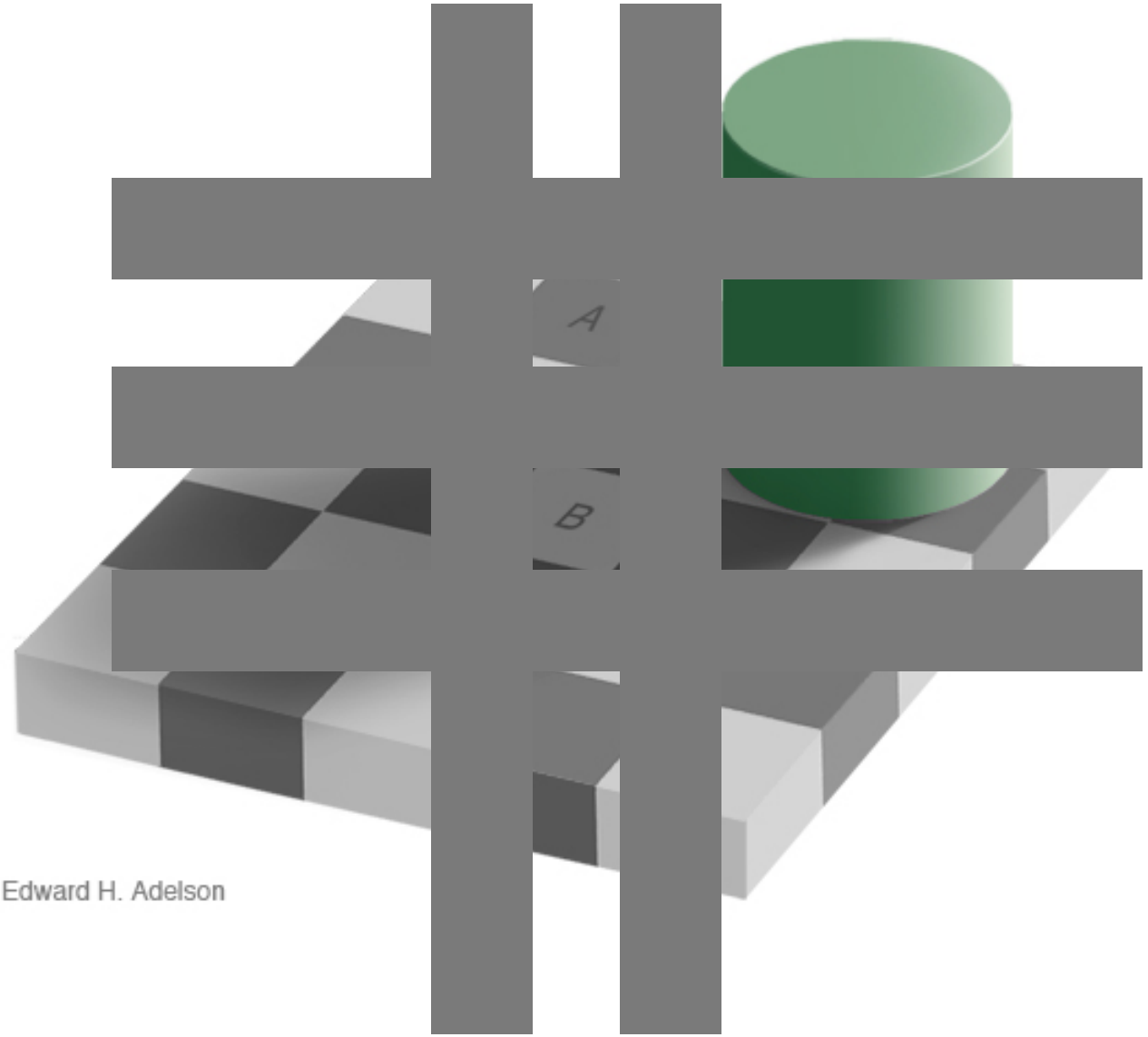
Rensink, O'regan, Simon, etc.

Change blindess

Rensink, O'regan, Simon, etc.

segmentation

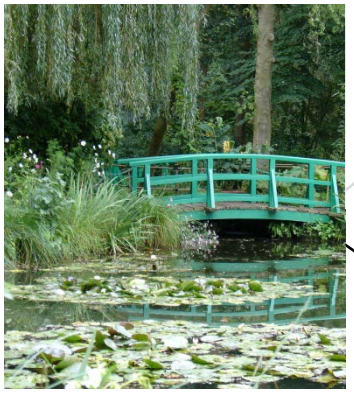
Perception



Edward H. Adelson

What is (computer) vision?

Image (or video)



Sensing device

Interpreting device

Interpretations

garden, spring, bridge, water, trees, flower, green, etc.

The goal of computer vision

- To bridge the gap between pixels and “meaning”

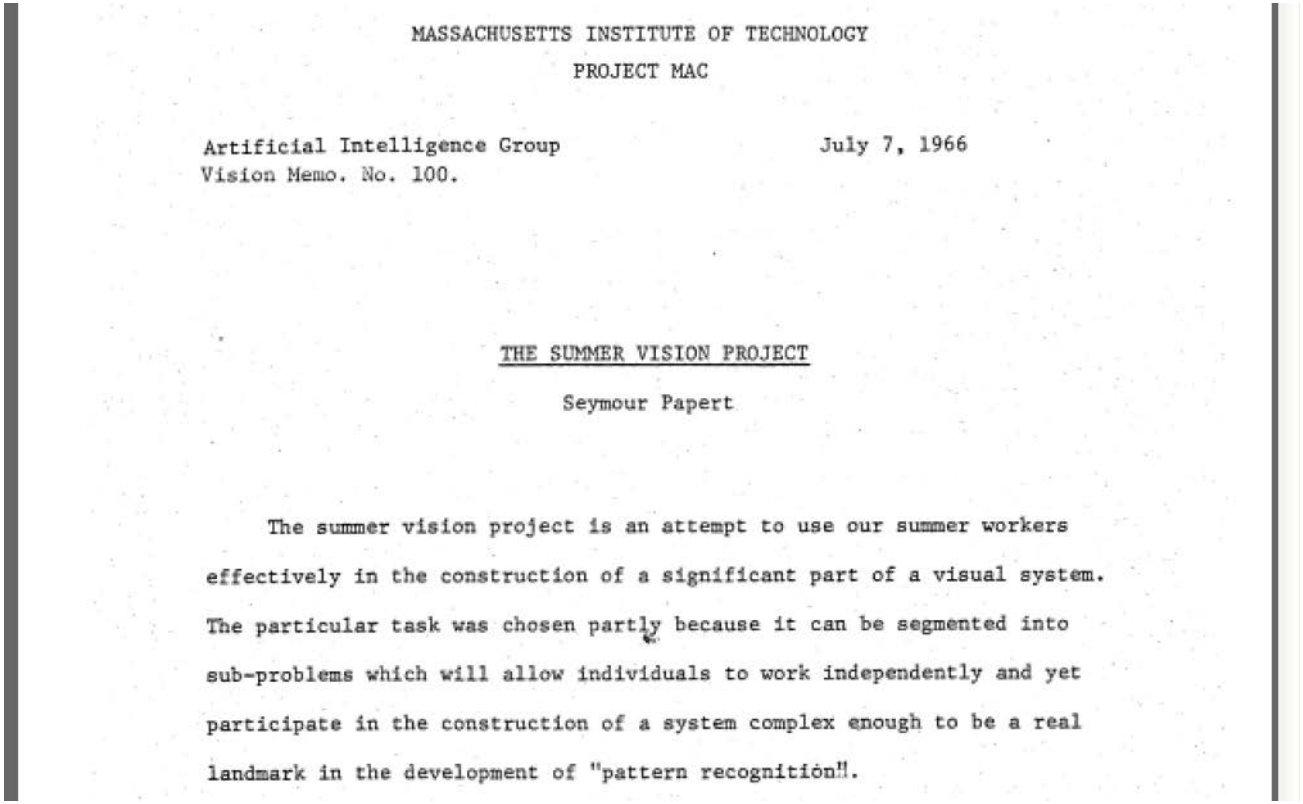
What we see

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

Source: S. Narasimhan

Origins of computer vision: an MIT undergraduate summer project



MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

THE SUMMER VISION PROJECT

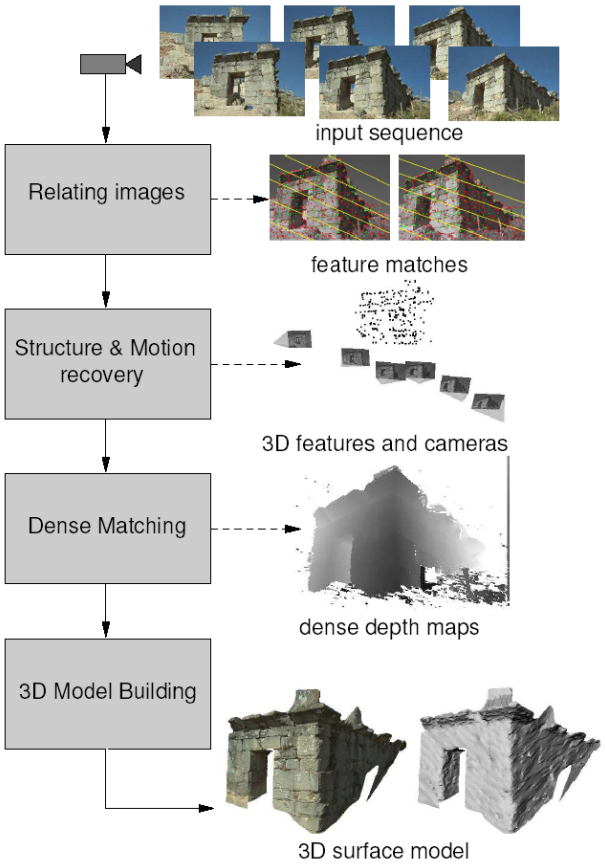
Seymour Papert

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".

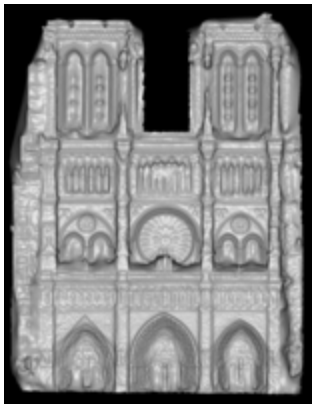
What kind of information can we extract from an image?

- Metric 3D information
- Semantic information

Vision as measurement device



Pollefeys et al.



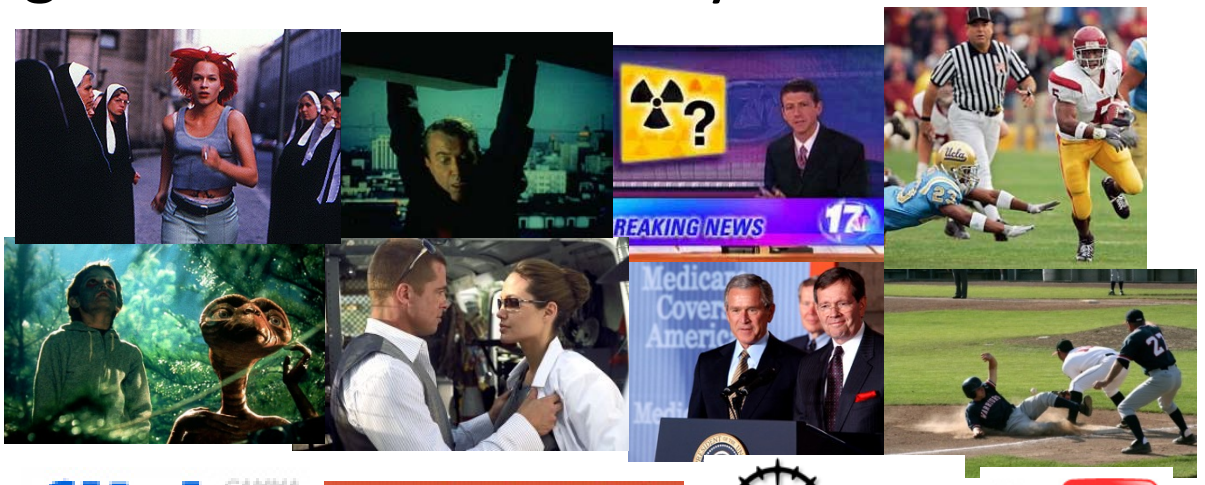
Goesele et al.

Vision as a source of semantic information

Slide credit: Kristen Grauman

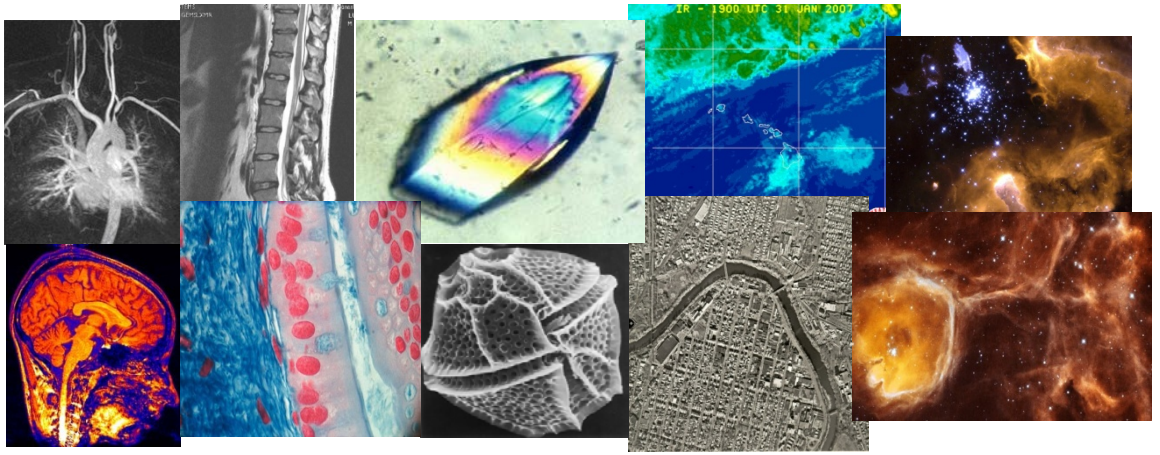
Why study computer vision?

- Vision is useful: Images and video are everywhere!



Google Image Search Picasa™ flickr™ webshots picsearch™ YouTube Broadcast Yourself™

Surveillance and security



Medical and scientific images

Why study computer vision?

- Vision is useful
- Vision is interesting
- Vision is difficult
  - Half of primate cerebral cortex is devoted to visual processing
  - Achieving human-level visual perception is probably "AI-complete"

Why is computer vision difficult?

Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

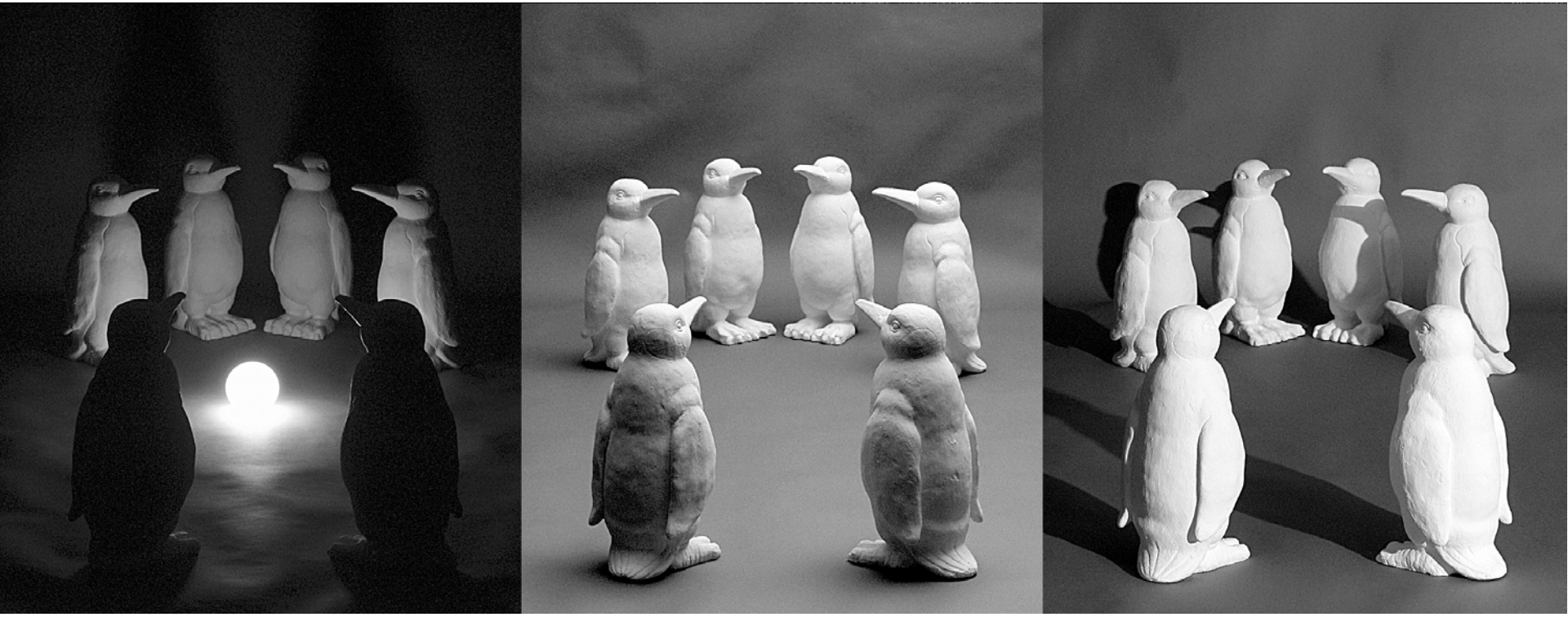
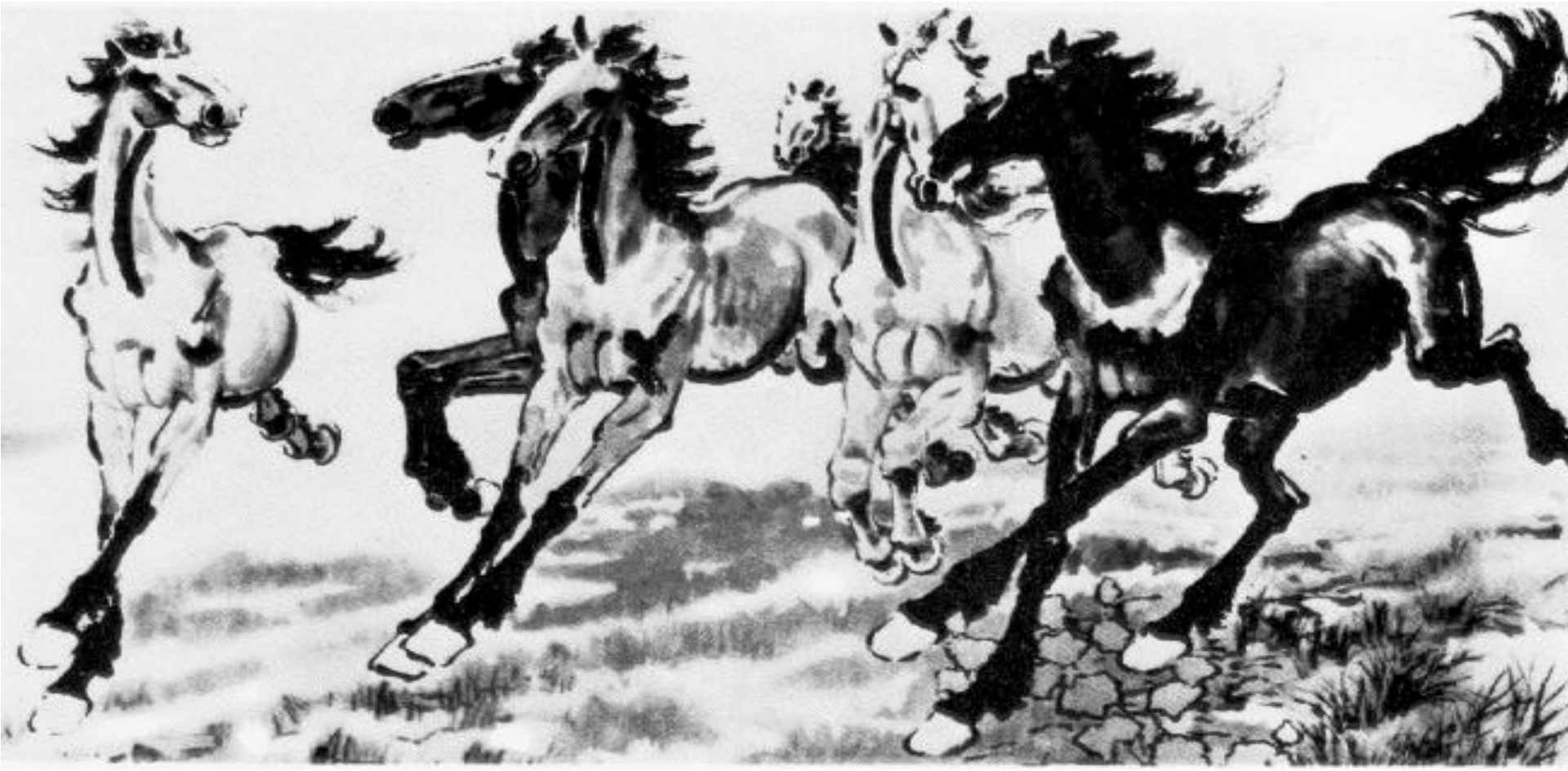


image credit: J. Koenderink

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

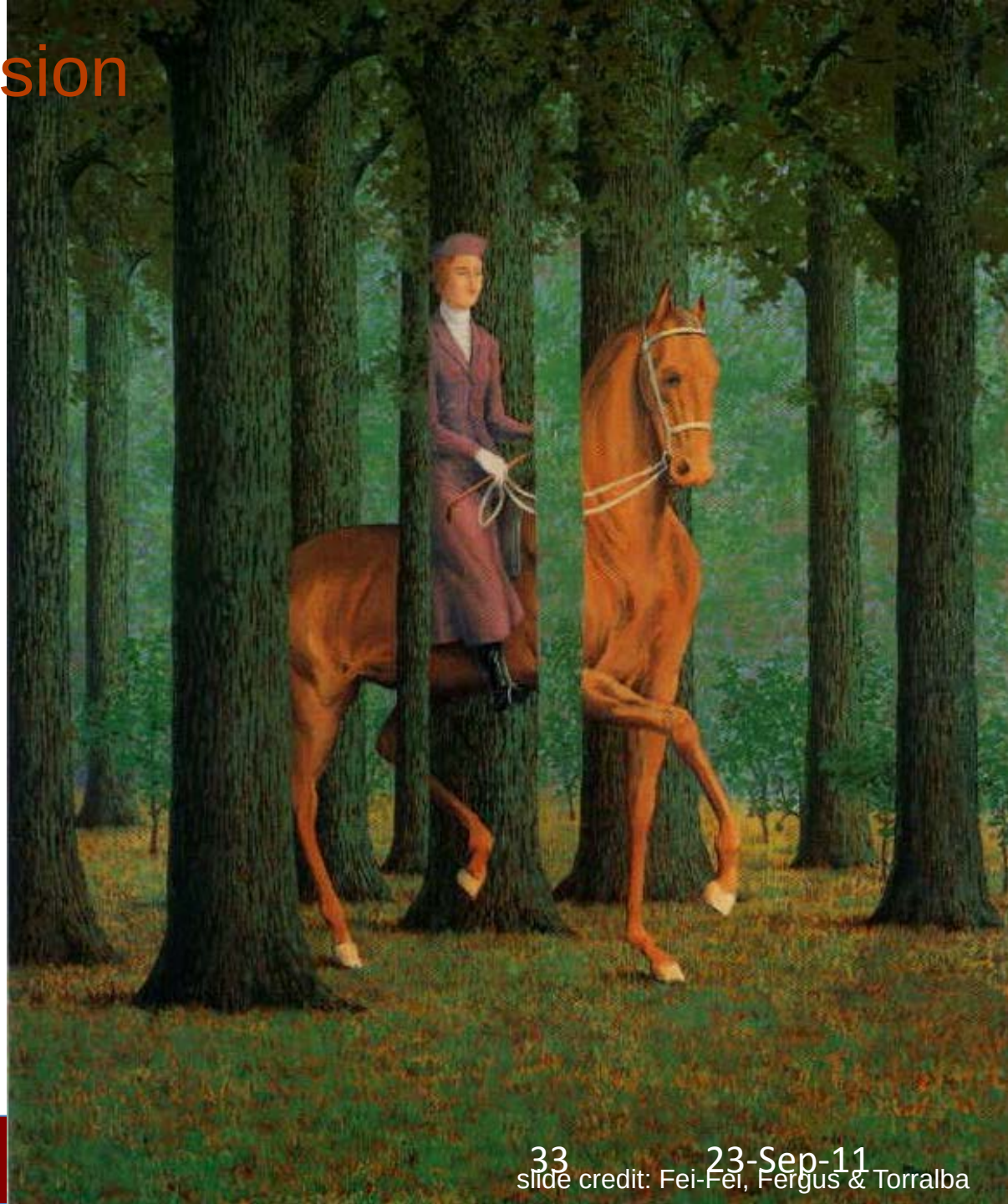
Challenges: deformation



Xu, Beihong 1943

slide credit: Fei-Fei, Fergus & Torralba

Challenges: occlusion



Magritte, 1957

slide credit: Fei-Fei, Fergus & Torralba

Challenges: background clutter

slide credit: Svetlana Lazebnik

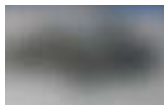
Challenges: Motion

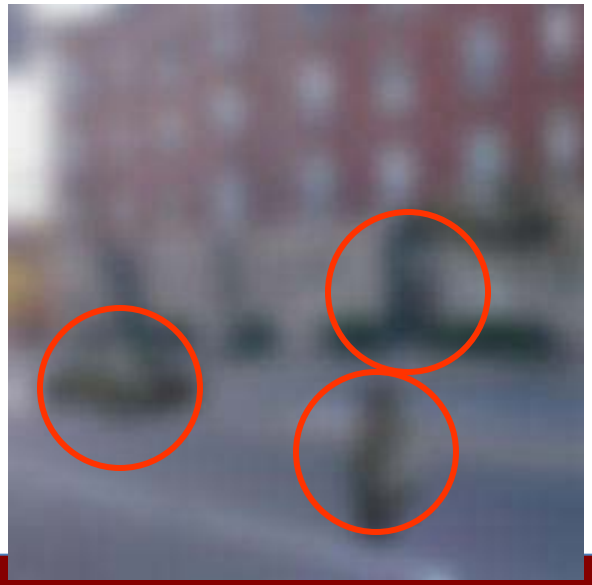
slide credit: Svetlana Lazebnik

Challenges: object intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity





slide credit: Fei-Fei, Fergus & Torralba

Challenges or opportunities?

- Images are confusing, but they also reveal the structure of the world through numerous cues
- Our job is to interpret the cues!

Image source: J. Koenderink

Depth cues: Linear perspective

slide credit: Svetlana Lazebnik

Depth cues: Aerial perspective

slide credit: Svetlana Lazebnik

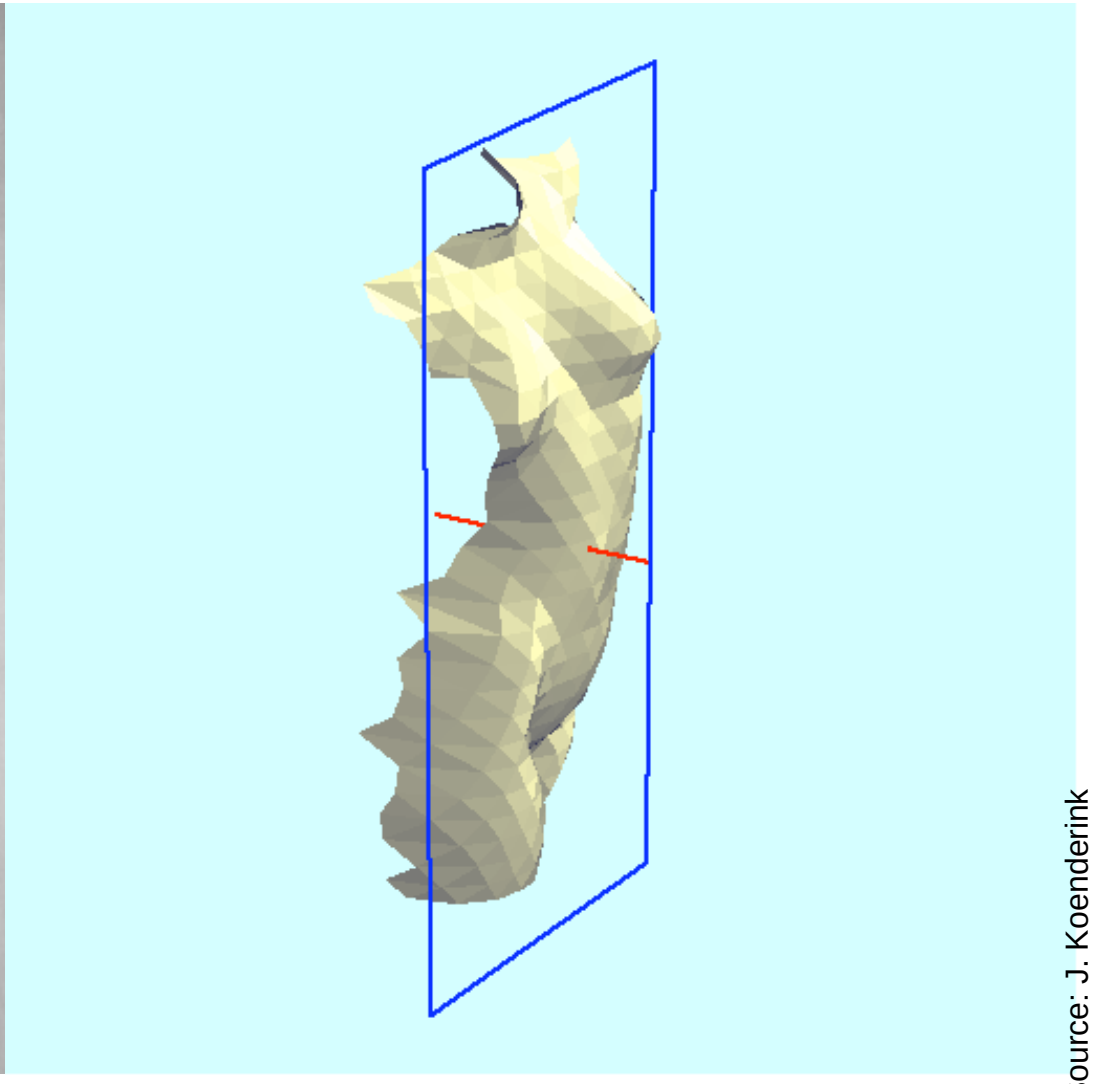
Depth ordering cues: Occlusion

Source: J. Koenderink

Shape cues: Texture gradient

slide credit: Svetlana Lazebnik

Shape and lighting cues: Shading



Source: J. Koenderink

Position and lighting cues: Cast shadows

Source: J. Koenderink

Grouping cues: Similarity (color, texture, proximity)

slide credit: Svetlana Lazebnik

Grouping cues: “Common fate”

Image credit: Arthus-Bertrand (via F. Durand)

Bottom line

- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a particular 2D picture

Bottom line

- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a particular 2D picture

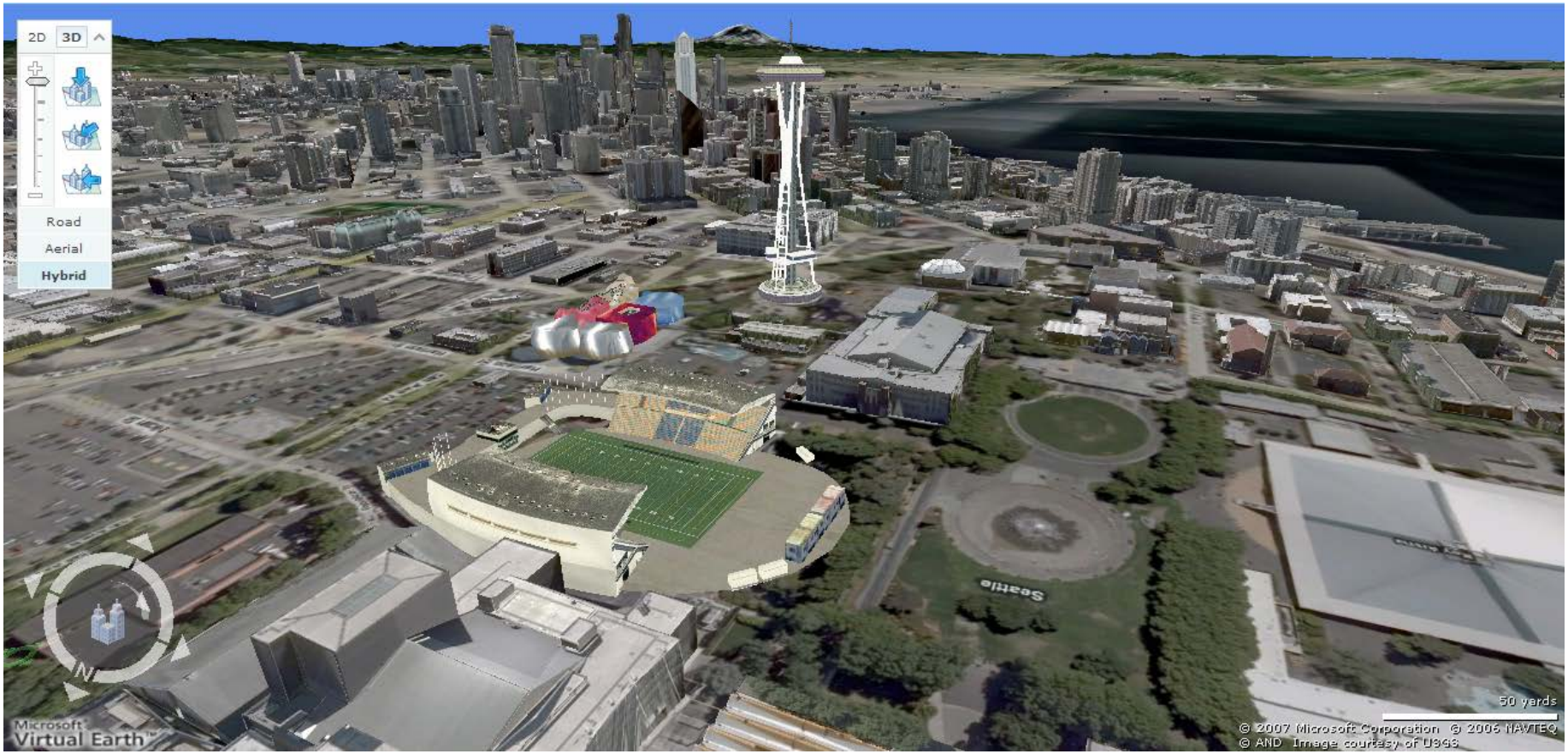
- Possible solutions
  - Bring in more constraints (more images)
  - Use prior knowledge about the structure of the world
- Need a combination of different methods

Computer Vision in the Real World

Special effects: shape and motion capture

Source: S. Seitz

3D urban modeling



Bing maps, Google Streetview

Source: S. Seitz

3D urban modeling: Microsoft Photosynth

http://labs.live.com/photosynth/

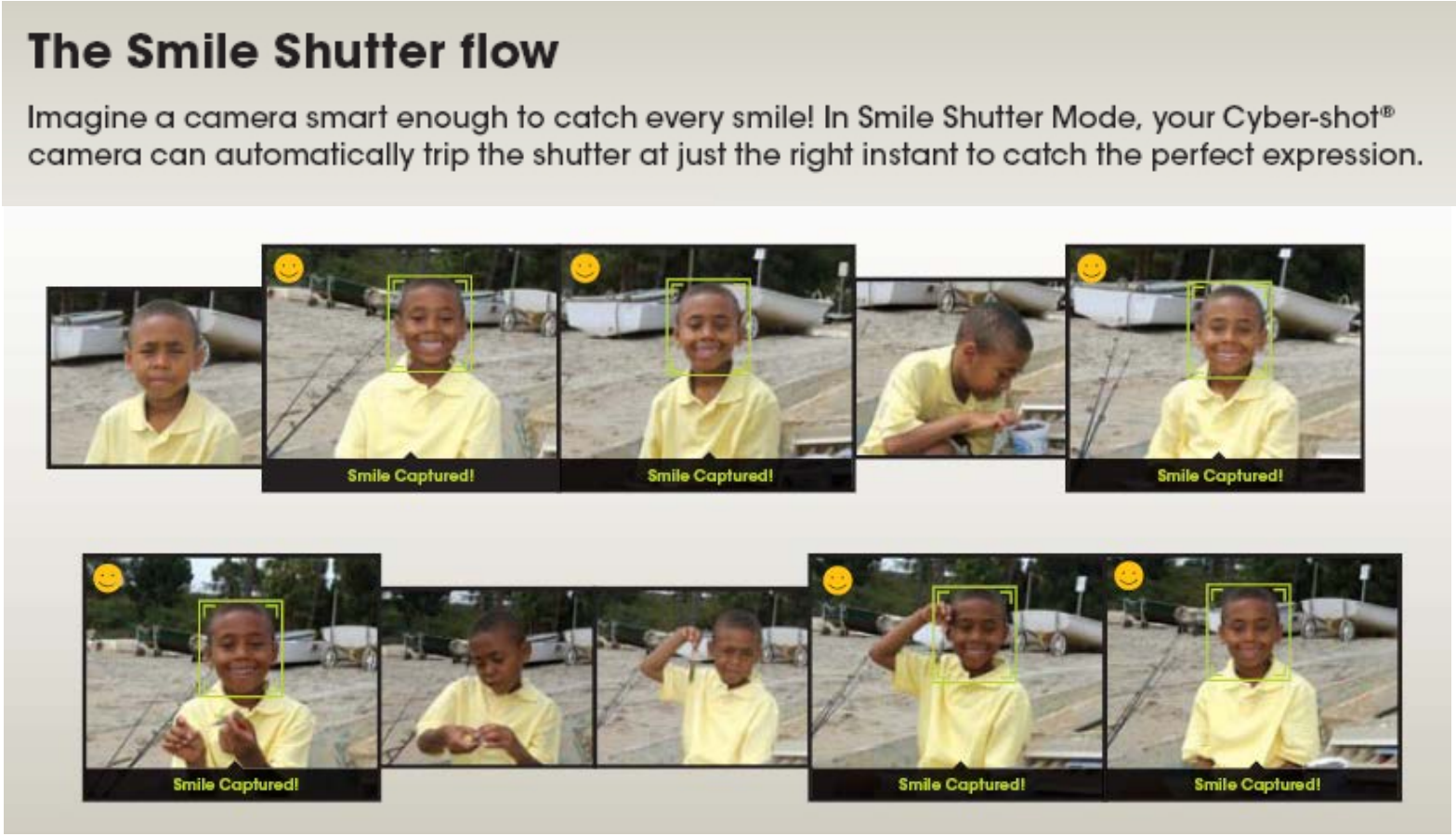
Source: S. Seitz

Face detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

Source: S. Seitz

Smile detection



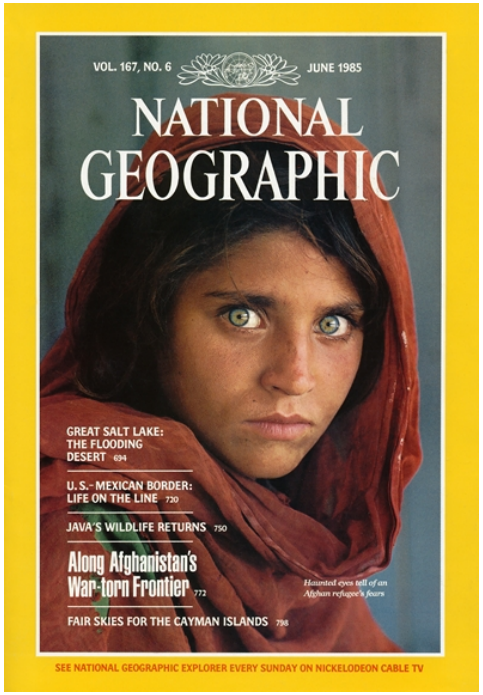
[Sony Cyber-shot® T70 Digital Still Camera]

Source: S. Seitz

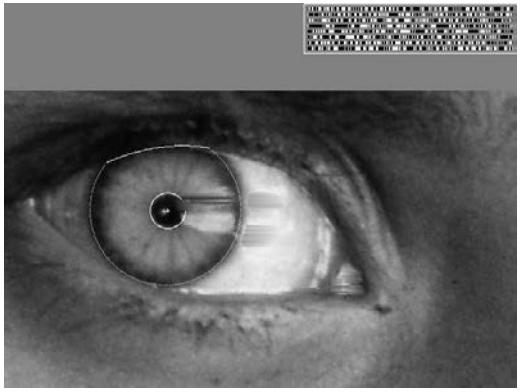
Face recognition: Apple iPhoto software

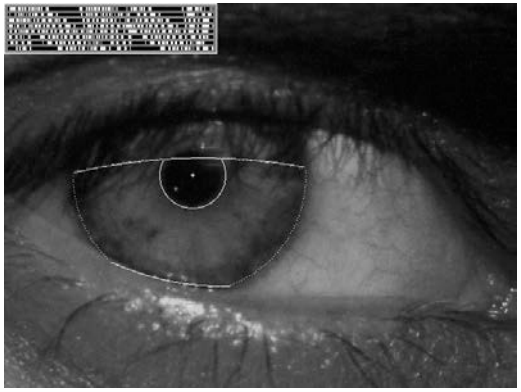
http://www.apple.com/ilife/iphoto/

Biometrics



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





Source: S. Seitz

Biometrics

Fingerprint scanners on many new laptops, other devices

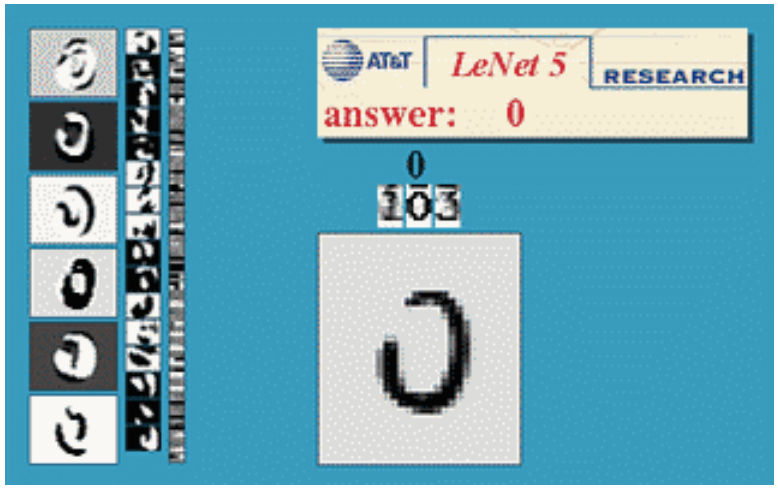
Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Source: S. Seitz

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs

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

Source: S. Seitz

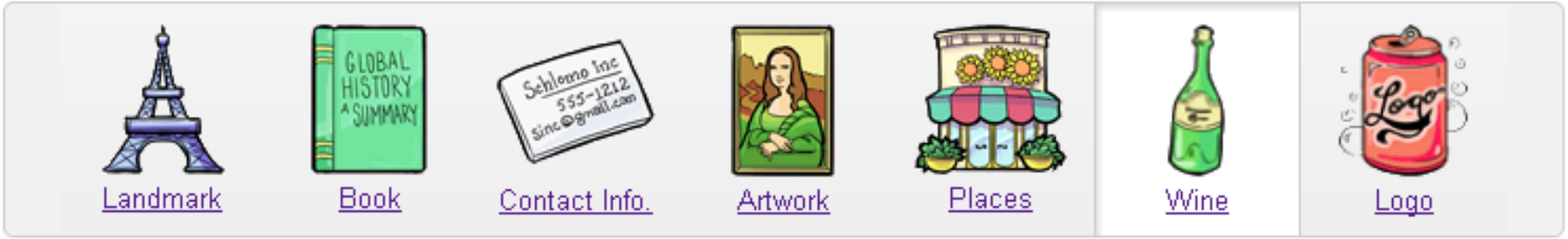
Toys and Robots

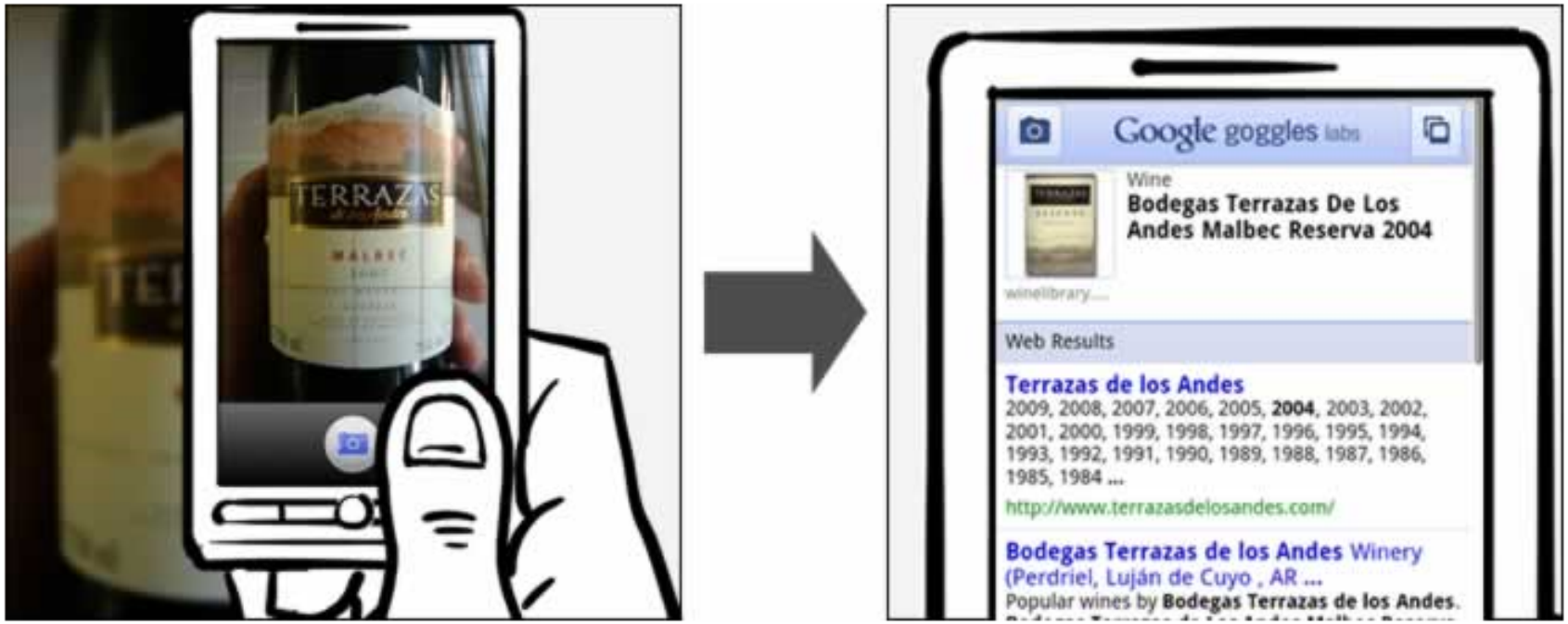
Mobile visual search: Google Goggles

Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.

Google goggles labs

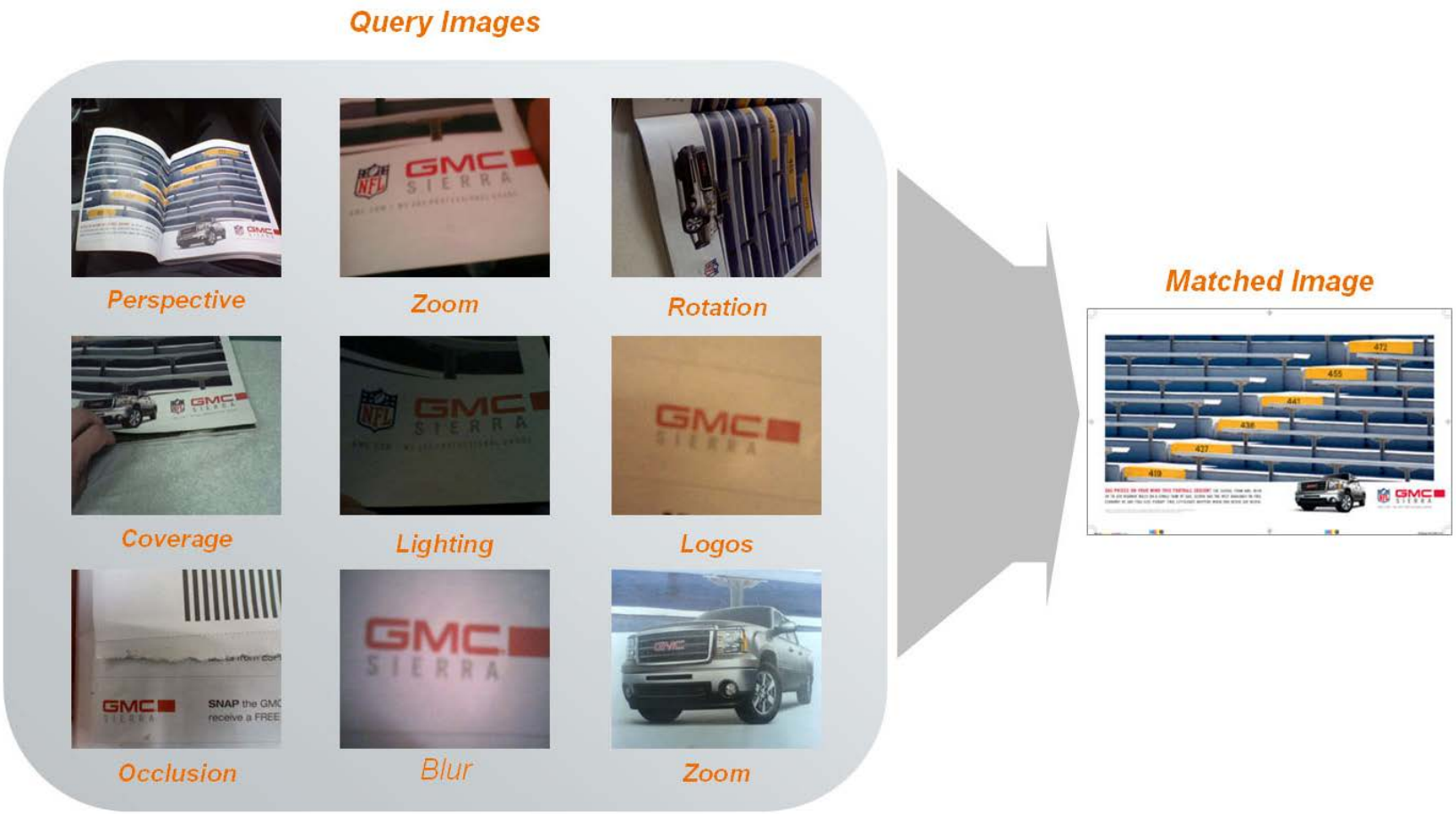




Mobile visual search: iPhone Apps

snaptell

kooaba



Automotive safety

- Mobileye: Vision systems in high-end BMW, GM, Volvo models
  - "In mid 2010 Mobileye will launch a world's first application of full emergency braking for collision mitigation for pedestrians where vision is the key technology for detecting pedestrians."

Source: A. Shashua, S. Seitz

Vision in supermarkets

[LaneHawk by EvolutionRobotics]

"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk, you are assured to get paid for it… "

Source: S. Seitz

Vision-based interaction (and games)

Microsoft's Kinect

Sony EyeToy

Assistive technologies

Source: S. Seitz

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

The computer vision industry

- A list of companies here:

  http://www.cs.ubc.ca/spider/lowe/vision.html

Today's agenda

- Introduction to computer vision
- Course overview

TA's of the class

- Kevin Tang
  - cs231a-aut1112-staff@lists.
  - Office hour: Thur 4-5pm
- Jiahui Shi
  - cs231a-aut1112-staff@lists.
  - Office hour: Mon 4-5pm
- Yongwhan Lim
  - cs231a-aut1112-staff@lists.
  - Office hour: Fri 3:30-4:30pm
- Hao Su
  - cs231a-aut1112-staff@lists.
  - Office hour: Tues 5-6pm

Syllabus

- Go to website…

Course Project: overview

- 40% of your grade
- Form your team:
  - either 2 people or 1 person
  - but the quality is judged regardless of the number of people on the team
  - be nice to your partner: do you plan to drop the course?

Course Project: overview (continued)

- Start immediately
- Some important dates:
  - Mon, Oct 17
    - Finalize team
    - Project proposal due
  - Mon, Nov 7
    - Milestone due (2-3 pages)
  - Tues, Dec 13
    - Final program code and writeup submission
  - Thurs, Dec 15
    - Presentation

Course Project: overview (continued)

- Original research ideas encouraged
- Useful datasets:
  - ImageNet (www.image-net.org)
  - PASCAL
- Need Fei-Fei's approval
  - Email is the best way
  - Do it BEFORE Jan 27

Grading policy

- Problem Sets: 40%
  - We have 5 problem sets
  - Homework 0: very important! (more details…)
  - Late policy
    - 5 free late days – use them in your ways
    - Afterwards, 25% off per day late
    - Not accepted after 3 late days per PS
  - Collaboration policy
    - Read the student code book, understand what is ‘collaboration’ and what is ‘academic infraction’
- Midterm Exam: 20%
  - In class: Mon, Oct 31

Grading policy

- Course project: 40%
  - presentation: 5%
  - write-up: 10%
    - clarity, structure, language, references: 3%
    - background literature survey, good understanding of the problem: 3%
    - good insights and discussions of methodology, analysis, results, etc.: 4%
  - technical: 15%
    - correctness: 5%
    - depth: 5%
    - innovation: 5%
  - evaluation and results: 10%
    - sound evaluation metric: 3%
    - thoroughness in analysis and experimentation: 3%

- A word about 'the curve'