CMPSCI 370: Introduction to Computer Vision

University of Massachusetts, Amherst
January 19, 2016

Instructor: Subhransu Maji

1

Administrivia

- **Lectures**
  - Tuesday/Thursday, 11:30 - 12:45, Hasbrouck 113
  - Honor's section: Tuesday 4:00 - 5:00, CS 142
- **Instructor:** Subhransu Maji
- **Office hours:** Monday, 3:00 - 5:00, CS 274
- **Website:** http://www-edlab.cs.umass.edu/~smaji/cmpsci370/
  - News, lecture slides, etc (check regularly)
  - Homework submission via Moodle

2

2

Administrivia

- **Grading policy**:
  - 370: homework (60%), mid-term (15%), final (25%)
  - 370HH: homework (45%), mid-term (10%), final (15%), project (20%)
- **Homework**
  - 5 in total (expect one every two weeks)
  - First one will be posted this Thursday
- **Course textbooks (recommended)**
  - Richard Szeliski, Computer Vision: Algorithms and Applications (available **online** as pdf) - readings will be from this
- **Necessary background**: Linear algebra, calculus, probability, programming in Matlab (image toolbox needed)

3

3

Homework #0

1.Figure out a way to run Matlab

- Obtain a student copy (Matlab suite 99$)
- Your lab machines might have it

2.Learn how to program in Matlab

- Plenty of online resources (the course website lists some)

- Alternatives: Python, Octave, JAVA, C++, ….

- Question: How many of you are familiar with Matlab?

4

4

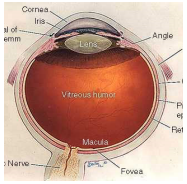
Before we start …

- Are there any questions?

Why Vision?

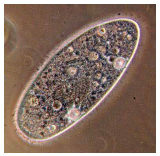
Why Vision? Light!

It is how we see other people, navigate our environment, communicate ideas, entertain, and **measure** the world around us.

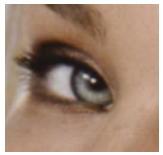


Source: A. Berg

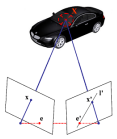
Why is light good for measurement?



Microscopy



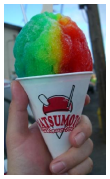
Surveillance



3D Analysis / Navigation

Remote Sensing

- Plentiful, sometimes free
- Interacts with many things, but not too many
- Goes generally straight over distance
- Very small → high spatial resolution
- Fast, but not too fast → time of flight sensors
- Easy to detect → cameras work, are cheap
- Comes in many flavors (wavelengths)



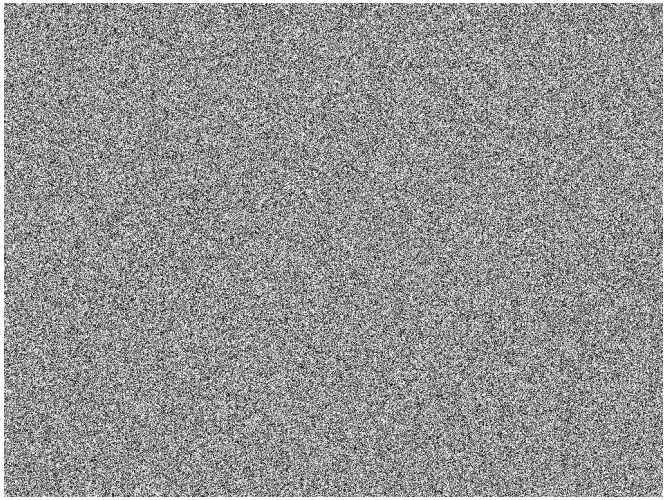
Source: A. Berg

The goal of computer vision

Extract properties of the world from visual data
(i.e., measurements of light)

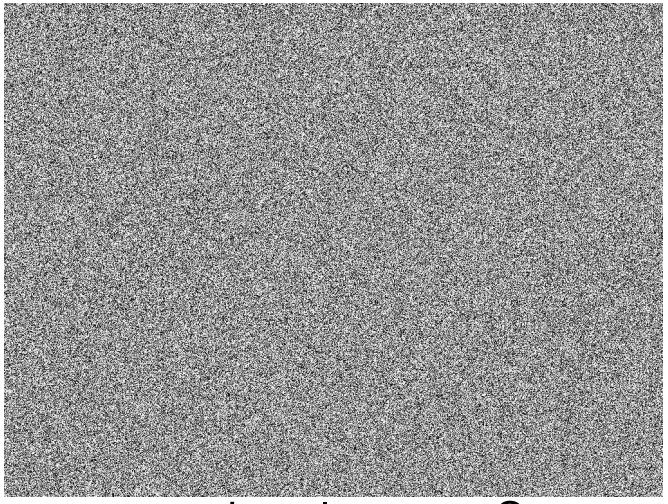
We are remarkably good at this!

An experiment …#1



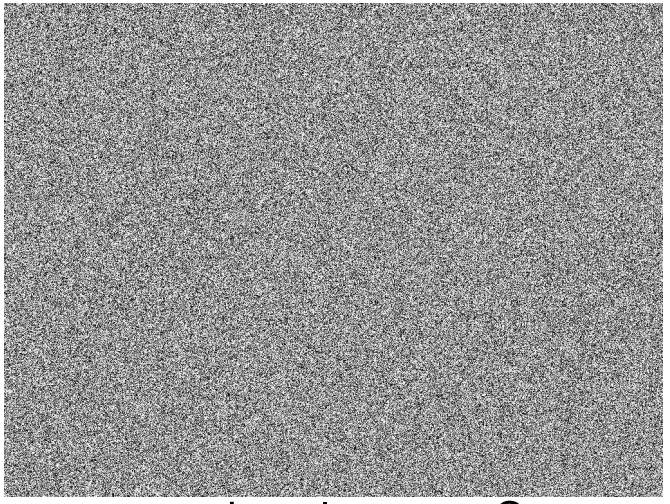
animal or not?

An experiment …#2



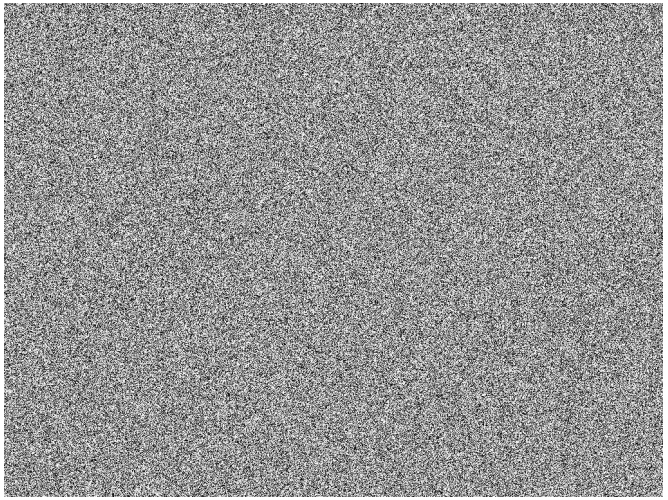
animal or not?

An experiment …#3



animal or not?

An experiment …#4

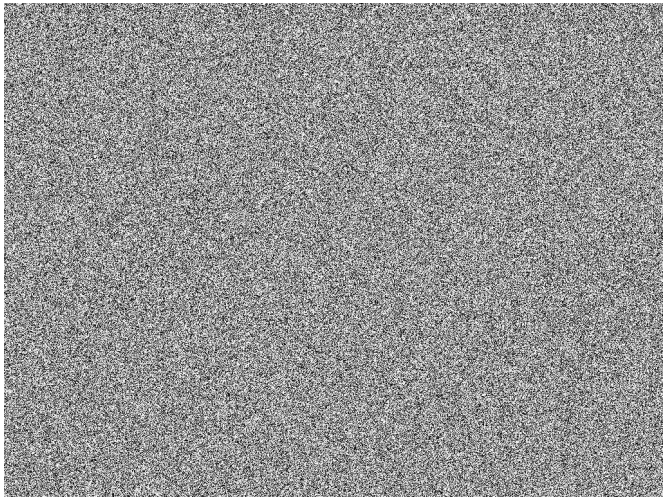


animal or not?

13

13

An experiment …#5

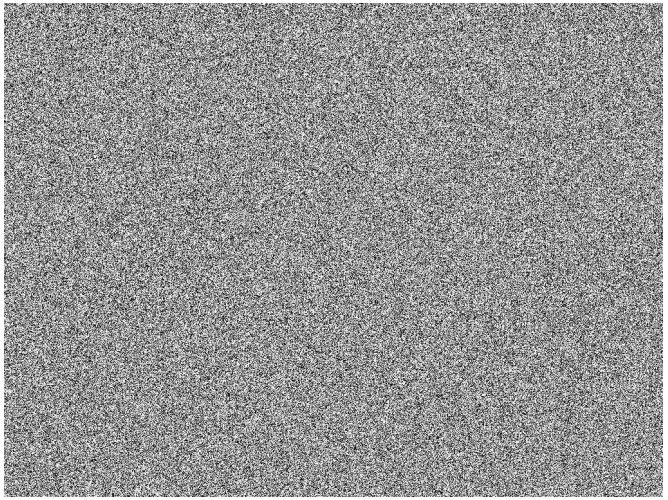


animal or not?

14

14

An experiment …#6



animal or not?

15

15

The images …

#1

#2

#3

#4

#5

#6

16

16

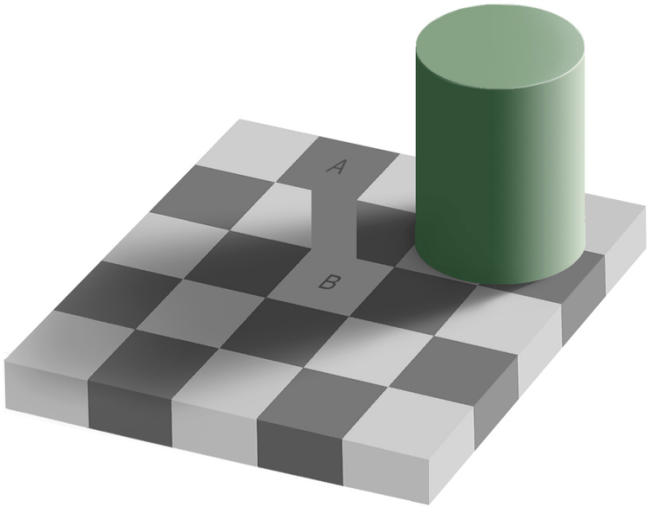
Human vision

- Amazingly good, fast and accurate
- Sometimes wrong, but often not in doubt
- Huge amount of bandwidth to the brain is visual data
- Large amount of the brain seems to be for processing visual data
- Vision is difficult!

Source: A. Berg 17

17

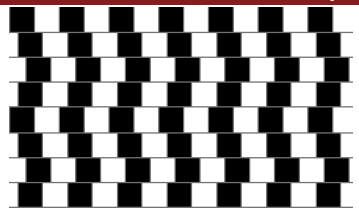
But we make mistakes …



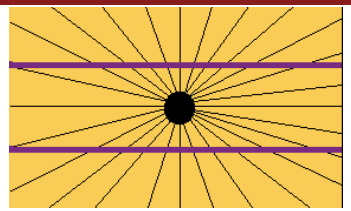
Checker shadow illusion - Edward H. Adelson 18

18

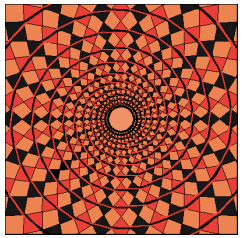
Other optical illusions



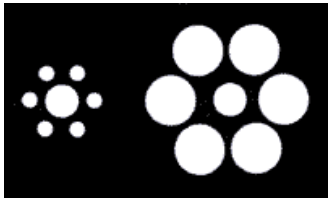
Are the horizontal lines parallel?



Are the purple lines straight?



Is this a spiral?



is the left circle (in the center) bigger?

Are these failures of our vision system?

http://www.illusions.org 19

19

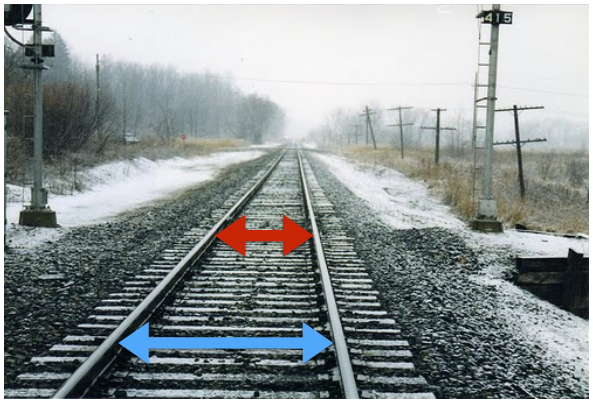
Vision as inverse of graphics

- Many possibilities — how do we solve this ambiguity?
  - Images are confusing, but they also reveal the structure of the world through numerous cues
  - Our job is to interpret the cues!

(following slides from J. Koenderink) 20

20

Cues: Linear perspective



http://kalisdigitalphotos.blogspot.com

Parallel lines merge at the horizon

Analyzing parallel lines to estimate space

21

21

Cues: Aerial (Atmospheric) perspective

Photo by **Éole Wind**

As the distance of the object from the viewer *increases*, the contrast between the object and its background *decreases*.

22

22

Cues: Occlusion ordering

Chicago loop, image source: **wikipedia**

23

23

Cues: Texture gradient

Gustave Caillebotte. Paris Street, Rainy Day, 1877, Art Institute of Chicago

24

24

Cues: Shading and Lighting

"The four seasons" sculpture set

25

25

Many other cues …

- **Motion parallax:** how things move relative to each other as we move. Objects near us move more than objects far away. Also provides *grouping* cues.
- **Familiar size:** Size of known things, e.g. faces gives us an estimate of the depth.
- **Defocus blur**: Far away objects are blurrier than nearer. Commonly used in photographs to create a perception of depth.
- **Elevation**: Distance from the horizon. Objects closer to the horizon are perceived to be farther.

26

26

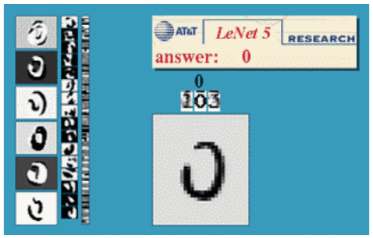
The study of computer vision

- Lots of tasks: detection, classification, segmentation, pose estimation, depth estimation, etc.
- Problems are often ill-posed. Most of the hard work is in crispy defining the problem you wish to solve.
- It is hard, ad-hoc. There are few theorems, but we rely on those from many other areas: optics, geometry, physics, etc.
- You are in good company:
  - Euclid, Alhazen, da Vinci, Kepler, Galileo, Descartes, Newton, Huygens, Maxwell, Helmholtz, Mach, Herring, Cajal, Minkowski, Hubel & Wiesel, Wald
- If that is not enough, there are many applications

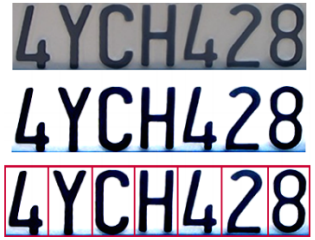
(following slides from Charless Flowkes) 27

27

Optical character recognition (OCR)



Digit recognition
yann.lecun.com



License plate readers
(google street view)

Automatic cheque readers
(Most bank ATMs)

Sudoku grabber
http://sudokugrab.blogspot.com/

Source: S. Seitz, N. Snavely 28

28

Biometrics

Fingerprint scanners are now on many new laptops and other devices

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

Source: S. Seitz 29

29

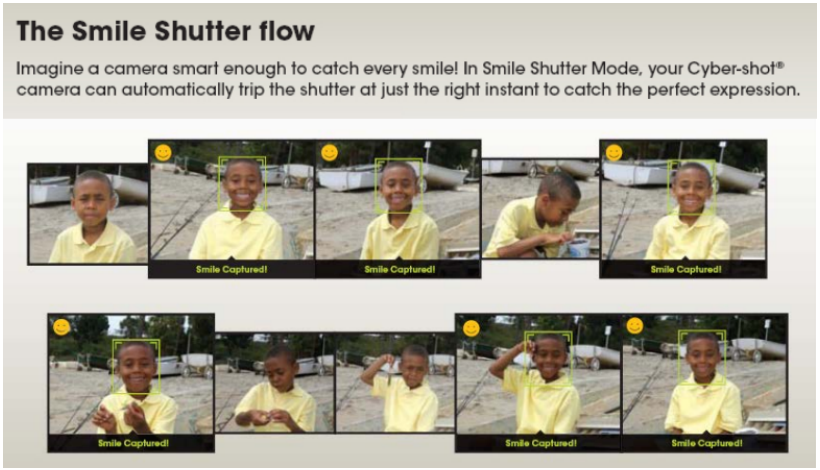
Face detection

Face detection is on many cameras these days

Source: S. Seitz 30

30

Smile detection



Source: S. Seitz 31

31

Face recognition

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

Source: S. Seitz 32

32

Instance recognition

Source: S. Seitz 33

33

Automotive safety

- **Mobileye** : Vision systems on high end BMW, GM, Volvo models
  - Pedestrian collision warning
  - Forward collision warning
  - Lane departure warning
  - Headway monitoring and warning

Source: A. Shashua, S. Seitz 34

34

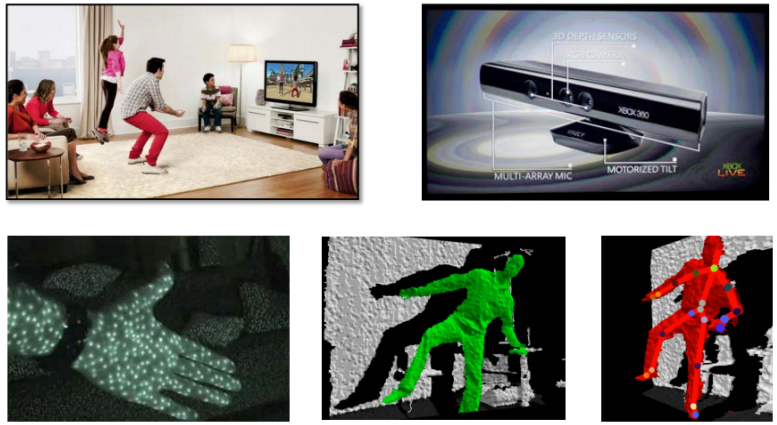
Self-driving cars

Source: L. Lazebnik 35

35

Interactive interfaces

Microsoft Kinect depth sensors



Source: L. Lazebnik 36

36

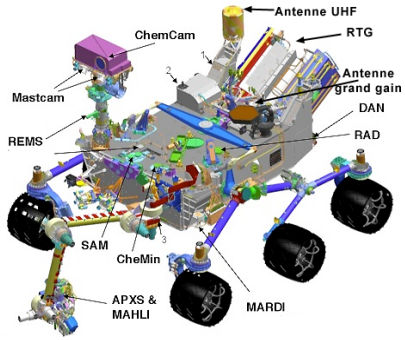
Large-scale 3D reconstruction

Photo Tourism: Exploring Photo Collections in 3D

YouTube link

Source: S. Seitz, N. Snavely

Vision for robotics, space exploration



NASA's Curiosity Rover has 17 cameras as a part of its sensing system

http://en.wikipedia.org/wiki/Curiosity_(rover)

What this course is about?

- Course overview
  - I. **Early vision:** image formation, sensing, light and shading, filtering
  - II. **Mid-level vision :** grouping, perceptual organization
  - III. **Multi-view geometry**
  - IV. **Recognition**
  - V. **Additional topics** (time permitting)
- **Goal:** To develop vision researchers. You can come up with a reasonable solution to various vision problems (and implement it yourself).
- We are not going to cover:
  - Graphics: Physics of light transport, material properties, rendering
  - Computational photography: design of sensing devices, etc
  - How the human vision system works

I. Early vision

- Basic image formation and processing

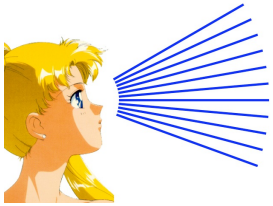


image formation

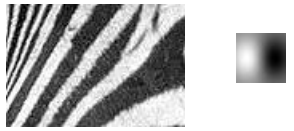
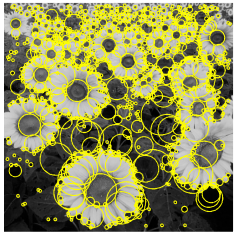


image filtering

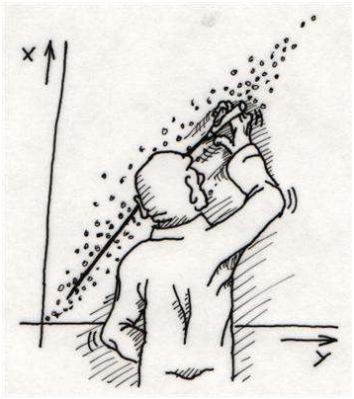


feature extraction, key-point detection

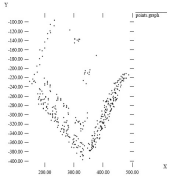
Source: L. Lazebnik

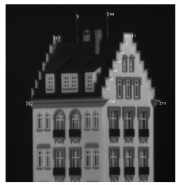
II. Mid-level vision

- Model fitting and grouping

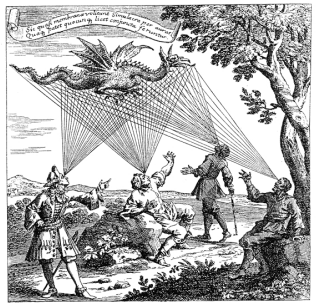


III. Multi-view geometry

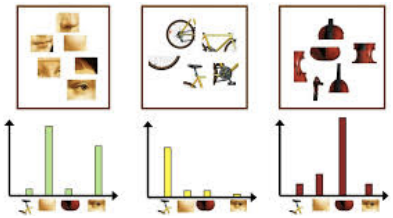




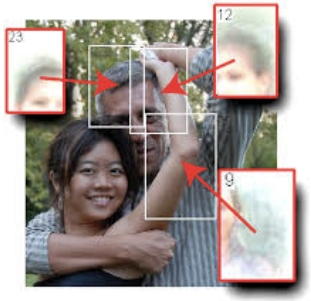
structure from motion



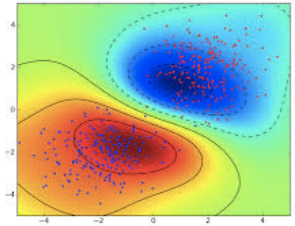
IV. Recognition



bag-of-word models

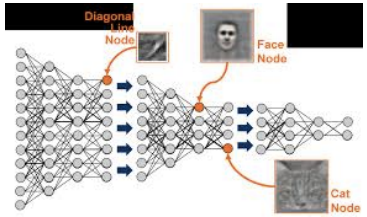


part-based models



learning

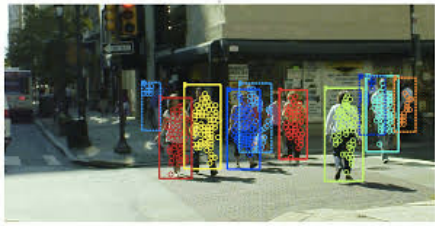
V. Additional topics



Deep learning

Human-centric vision

Optical flow



Tracking

For next class …

- Familiarize yourself with MATLAB (more information is on the course page)
  - **Student copy** is 99$ from Matlab's page
  - **UMASS IT** (100% free): https://www.it.umass.edu/support/software
- Readings:
  - **The speed of processing in the human visual system**, Thorpe et al., Letters to Nature, 1996
  - Chapter 1 in RS textbook

45