Class overview and intro to CV

Subhransu Maji

CMPSCI 670: Computer Vision

September 6, 2016

Course background

- What is the course about?
  - Physics and geometry of image formation
    - Understand how cameras work (and design new sensors)
  - Finding (and exploiting) patterns in visual data
    - Examples: object detection, image classification
    - It is hard, ad-hoc. There are few theorems, but we rely on those from many other areas: optics, geometry, physics, machine learning, optimization, statistics, etc.

- Why is computer vision so cool?
  - You are in good company: Euclid, Alhazen, da Vinci, Kepler, Galileo, Descartes, Newton, Huygens, Maxwell, Helmholtz, Mach, Herring, Cajal, Minkowski, Hubel & Wiesel, Wald
  - Broad applicability: robotics, biometrics, search, etc.
  - Open area, lots of room for new work

Course goals

- By the end of the semester, you should be able to:
  - Look at a problem and identify if CV is an appropriate solution
  - If so, identify what types of algorithms might be applicable
  - Apply those algorithms
  - Conquer the world

- In order to get there, you will need to:
  - Do a lot of math (calculus, linear algebra, probability)
  - Do a fair amount of programming
  - Work hard (this is a 3-unit course)

Topics covered

- Sensing light and image representation
  - Image formation, cameras, color, light, shading
- Basic image processing
  - linear filtering; detecting lines, corners, and blobs
- Recognition + other topics
  - model fitting, designing image representations, machine learning
  - applications: detection, segmentation, tracking, etc.

- Not a zoo tour!
- Not an introduction to tools!
- You will learn how these techniques work and how to implement them

Requirements and grading

- Weekly homework assignments: **20%**
  - About 12 in total, graded at 0, 0.5 or 1
  - Completed individually
  - May not be late at all
- Mini-projects: **50%**
  - Four or five in total
  - Completed individually (but can be discussed with others)
  - Can be 24 hours late, with a 50% mark down
- Project: **25%**
  - Canned or your choice, teams of two or more
  - Proposal, presentation (or poster), written report
- Class and forum participation: **5%**

Who should take this course?

- Is this the right course for you?
  - Do you have all the pre-requisites?
    - good math and programming background
  - Balance of theory vs. practice. Other courses being offered:
    - 589/689 — Machine learning
    - 697L — Deep learning seminar (Focus is on CNNs)
    - 690IV — Intelligent visual computing (Focus is on computer graphics)
- Still not sure?
  - talk to me after class
- Wait listed?
  - Will decide on a case by case basis

Course logistics

- My office hours: Tuesday 2:30 - 3:30pm, CS 274
- TA: Tsung-Yu Lin (office hours: tbd)
- Course website: http://www-edlab.cs.umass.edu/~smaji/cmpsci670
  - Class slides, links to homework assignments will the posted here
  - Check regularly for announcements
- Moodle for homework submissions and grades
- Piazza for discussions
  - If you use it, I will

- Textbooks (recommended):
  - Forsyth and Ponce, Computer Vision: A Modern Approach, 2nd edition
  - Richard Szeliski, Computer Vision: Algorithms and Applications (available online as pdf). I'll post readings from this

Things you need to know now!

- Finish homework 00
  - Due 8 September (that's Thursday! before class)
  - Not graded but required
  - Submit in *.pdf* format *only* via **moodle**
    - ➡ Those who are not yet on moodle may email me

- Get started on MATLAB
  - Acquire Matlab (student license for 100$)
  - Intro to MATLAB programming

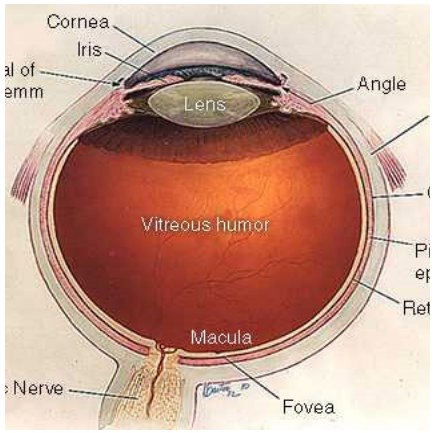
- Read the web page!

Now, on to some real content …

(but first, questions?)

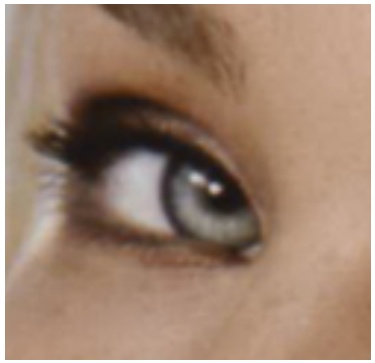
Why vision? Light!

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

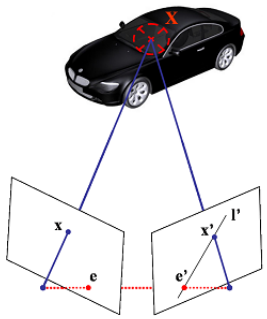


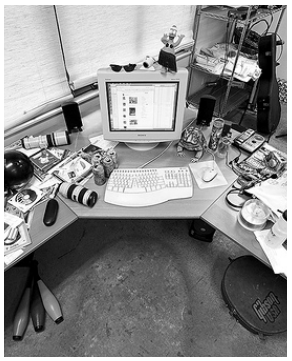
Why is light good for measurement?

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)

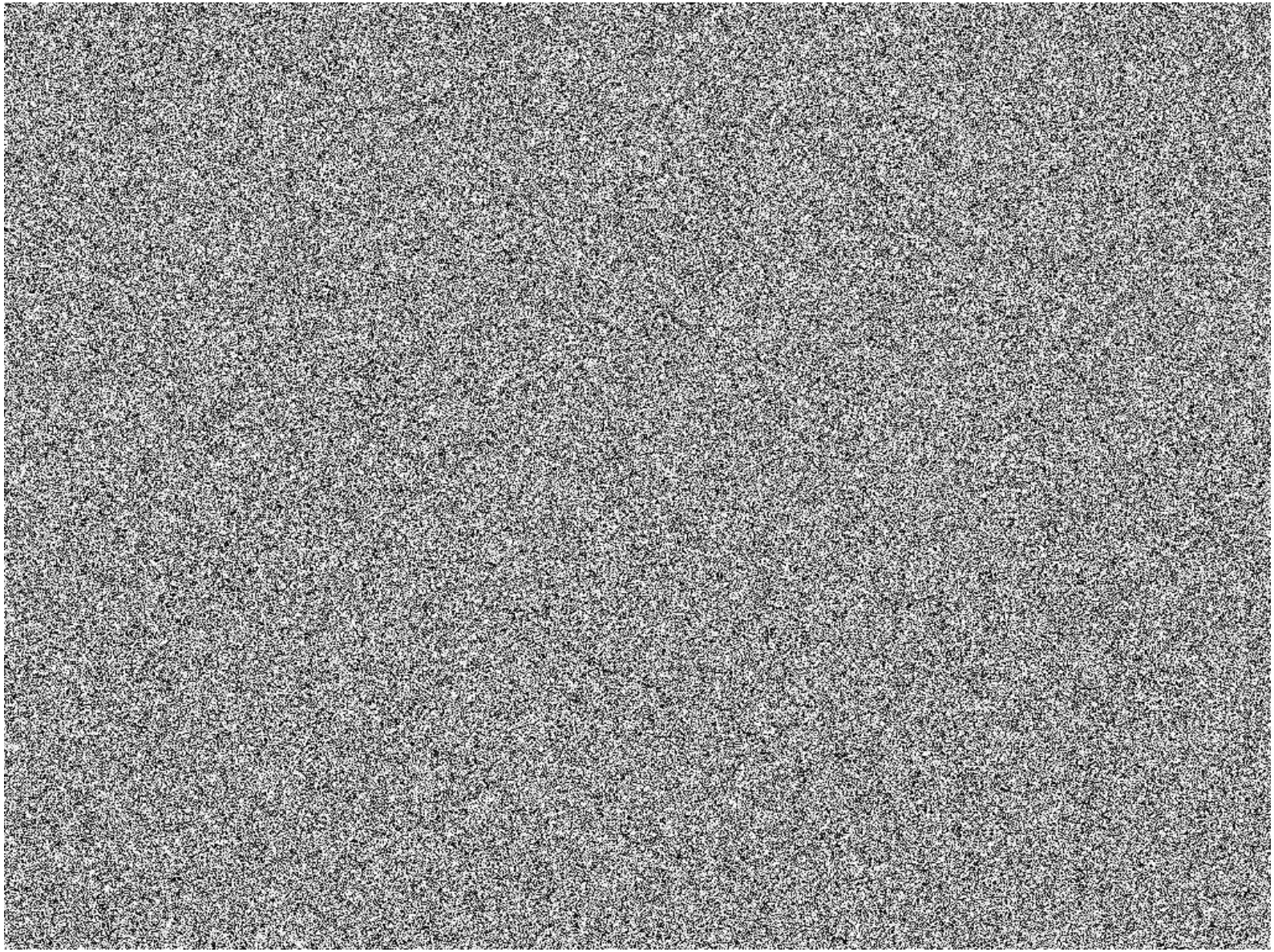
Source: Alex Berg

The goal of computer vision

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

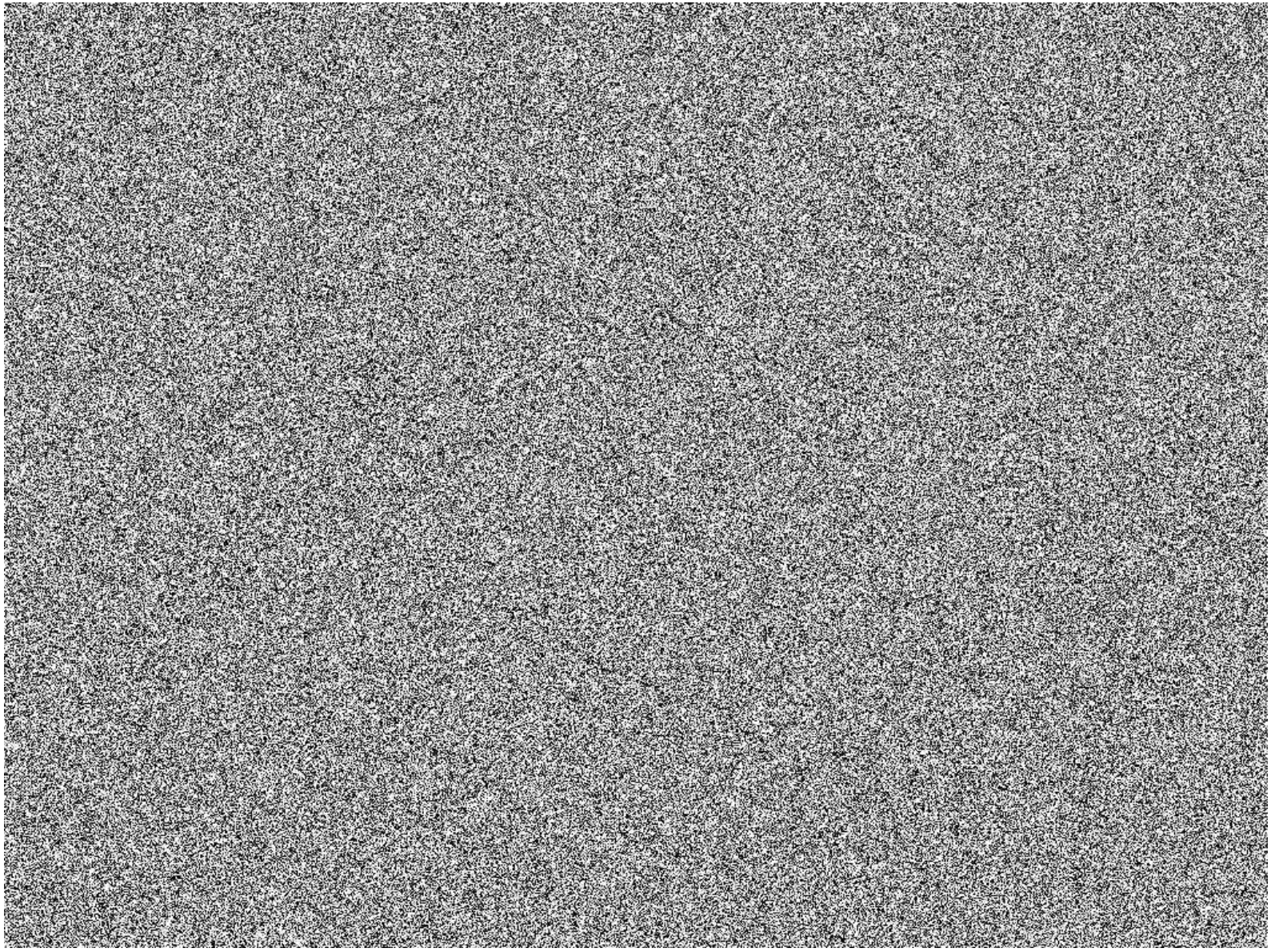
We are remarkably good at this!

An experiment …#1



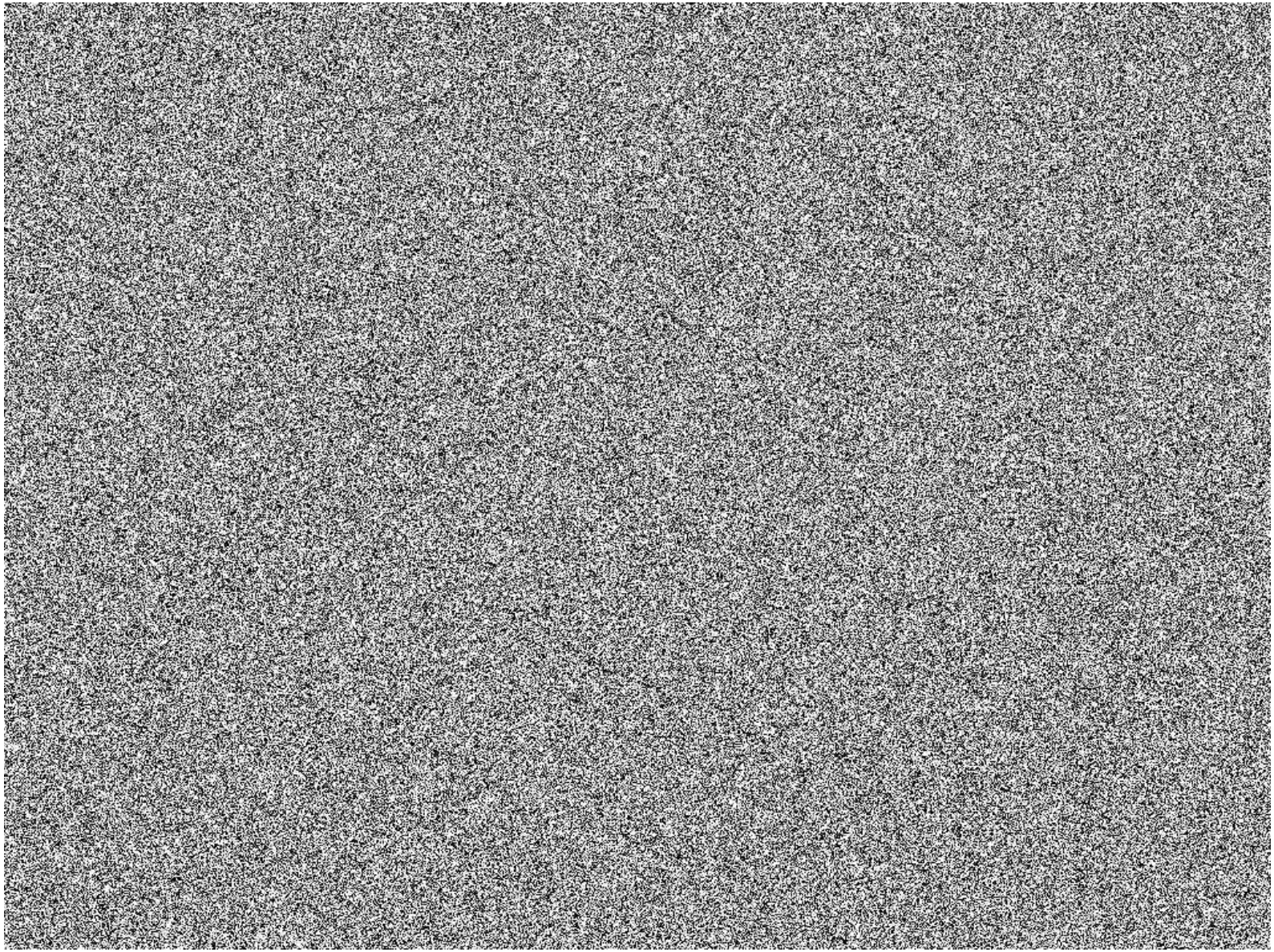
animal or not?

An experiment …#2



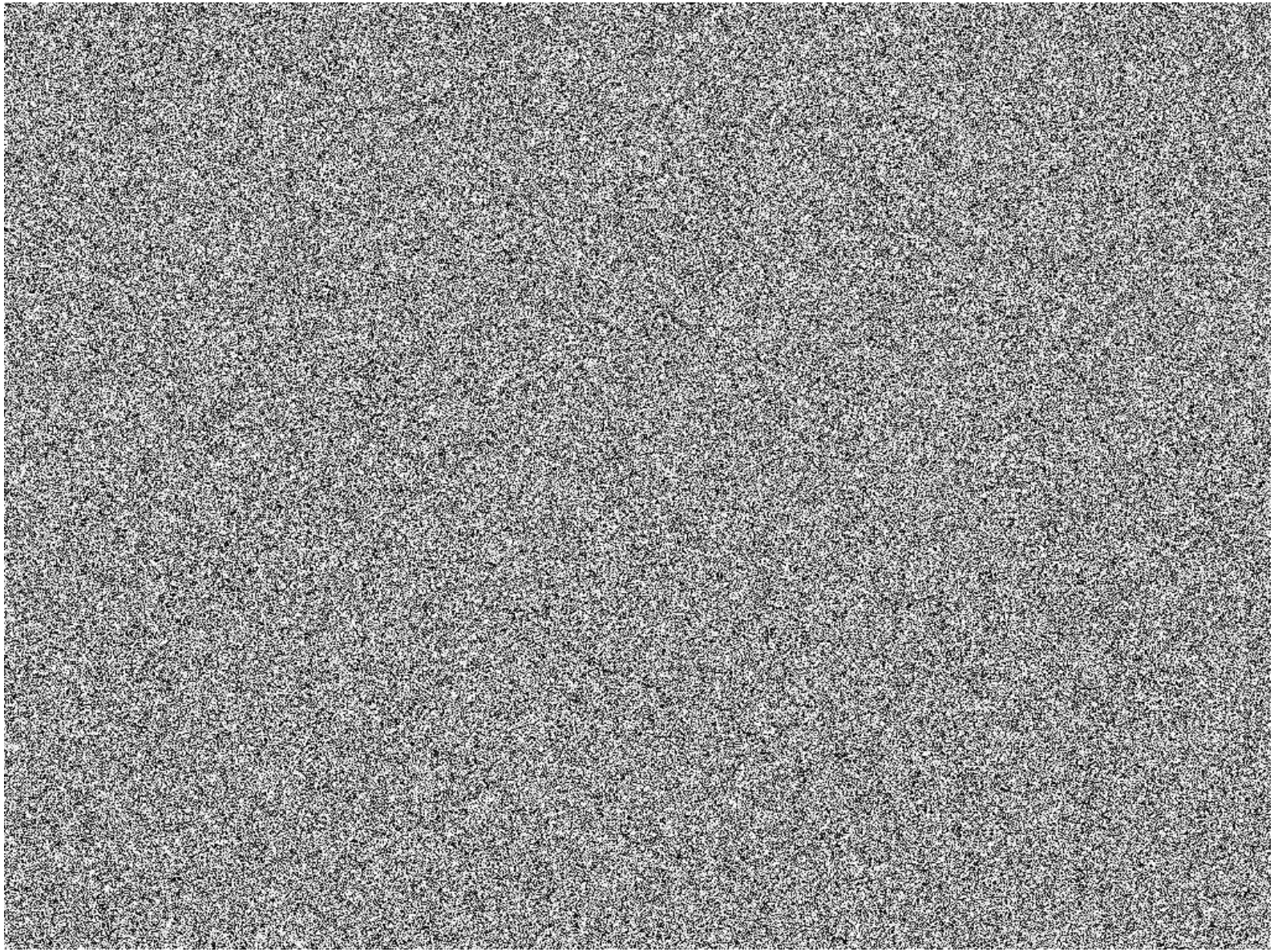
animal or not?

An experiment ….#3



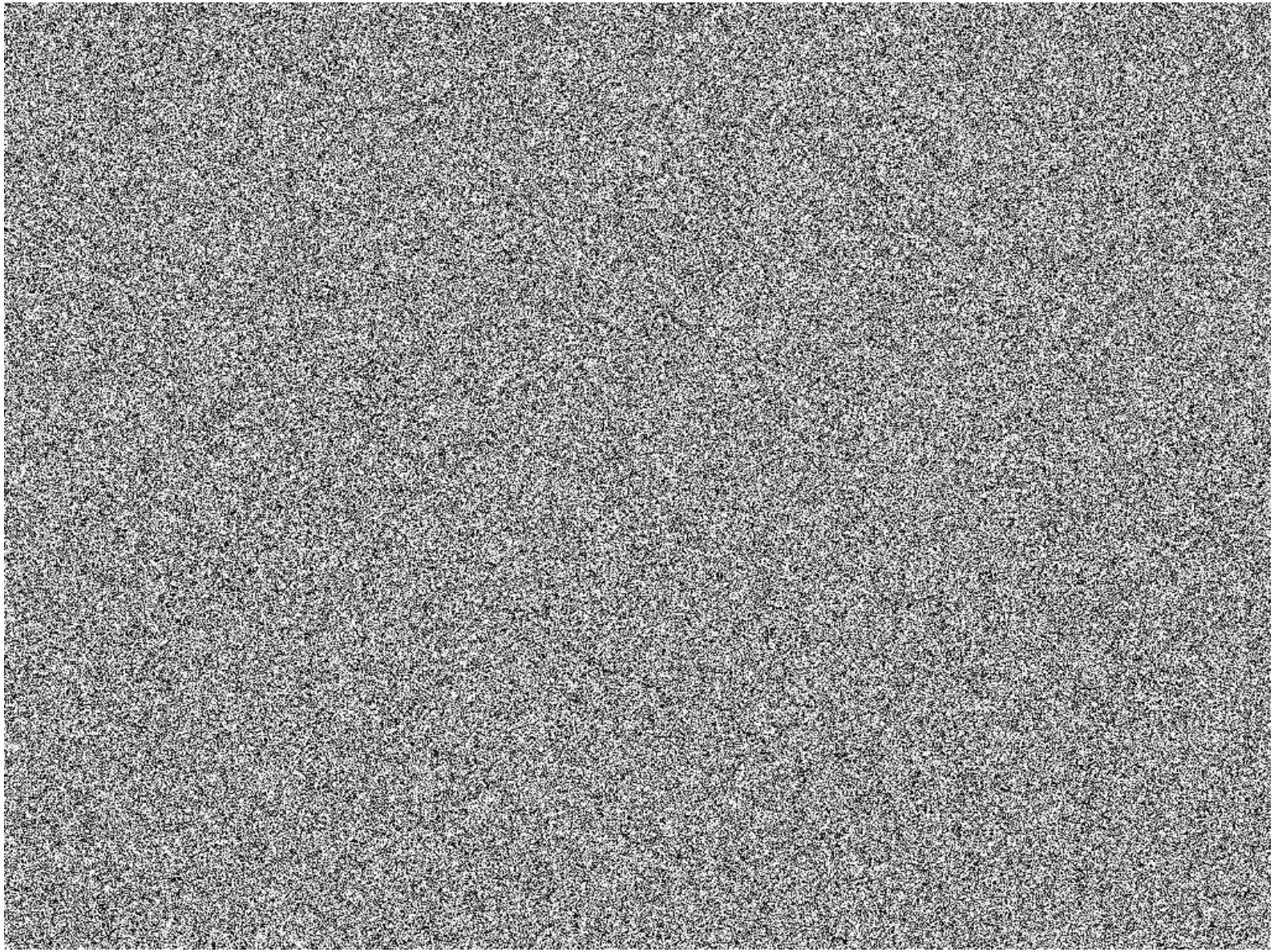
animal or not?

An experiment …#4



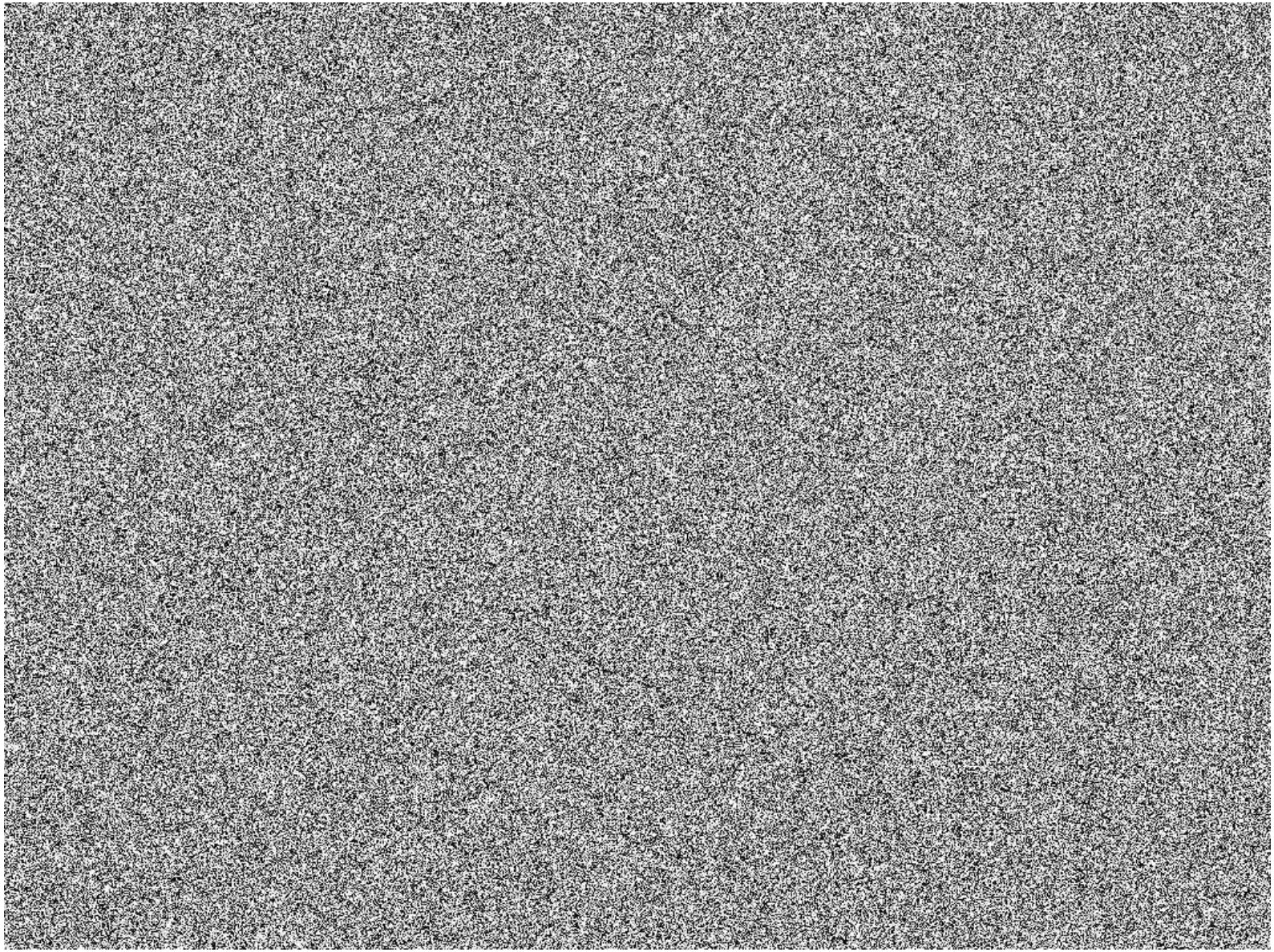
animal or not?

An experiment ….#5



animal or not?

An experiment ….#6



animal or not?

The images …

#1

#2

#3

#4

#5

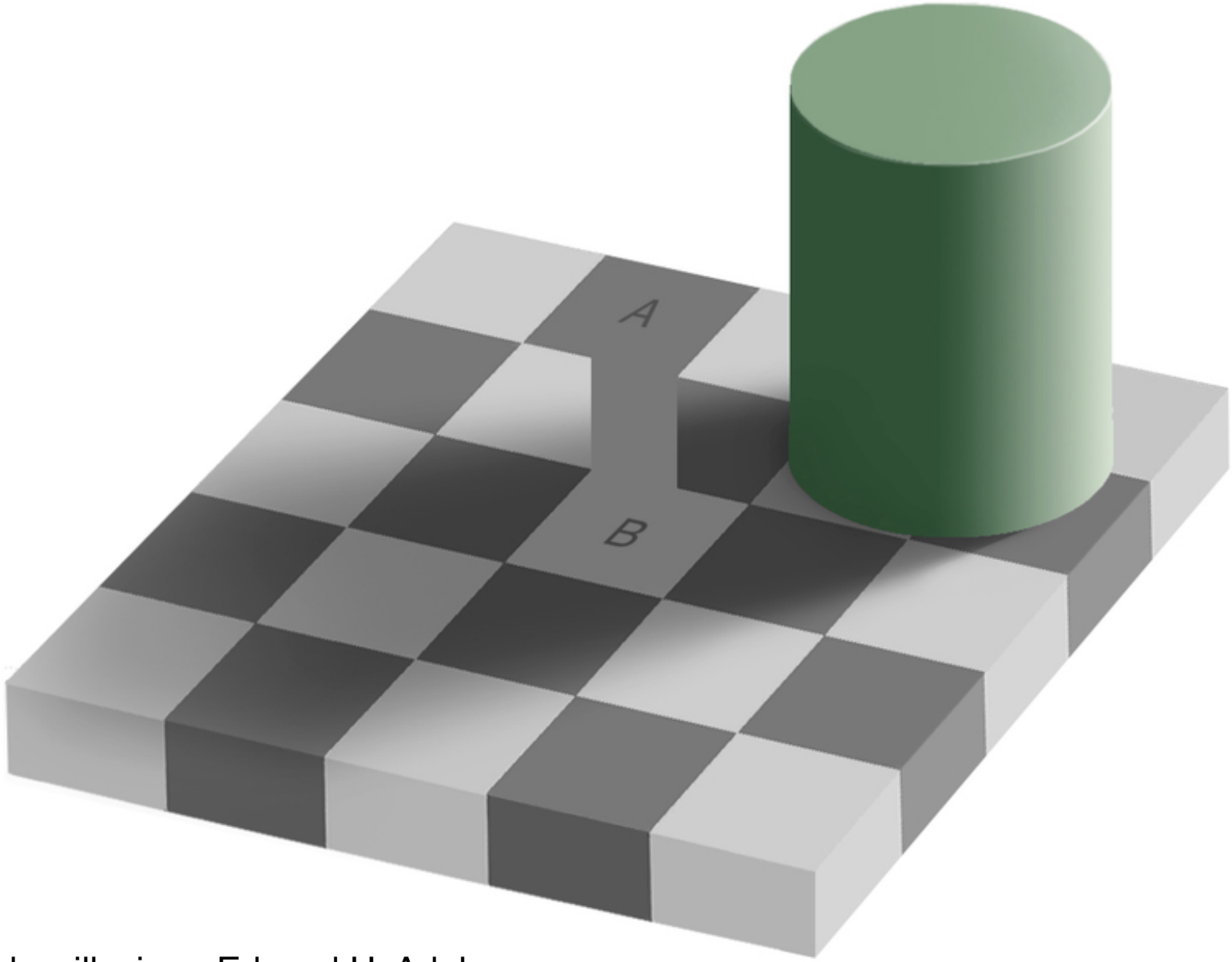
#6

Human vision

- Amazingly good, fast and accurate
- 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

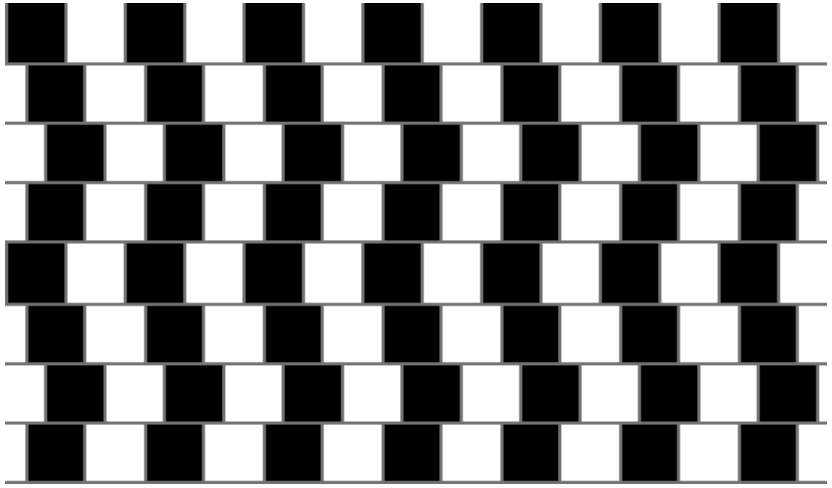
But we make mistakes …



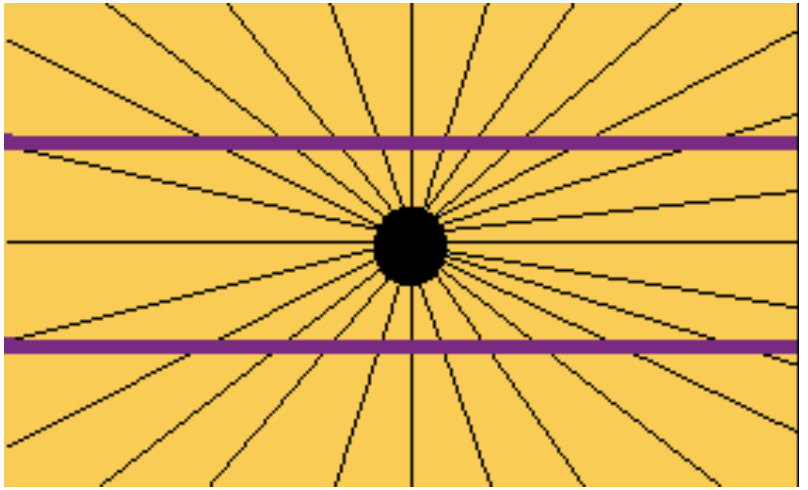
Checker shadow illusion - Edward H. Adelson

Other optical illusions

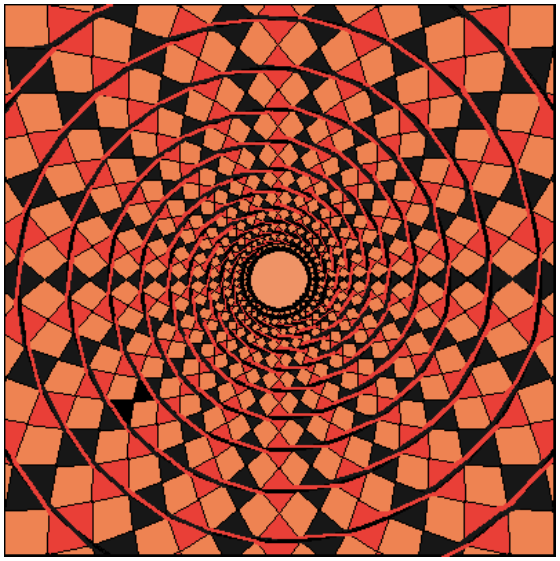
http://www.illusions.org



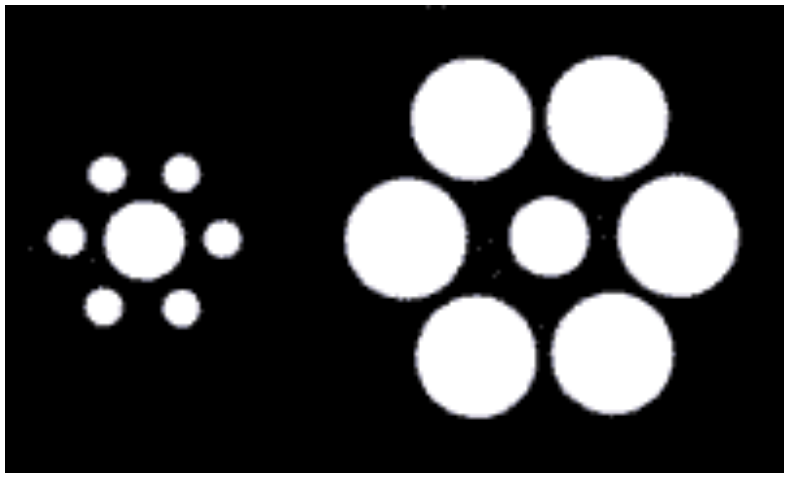
Are the horizontal lines parallel?



Are the purple lines straight?



Is this a spiral?



is the left circle (in the center) bigger?

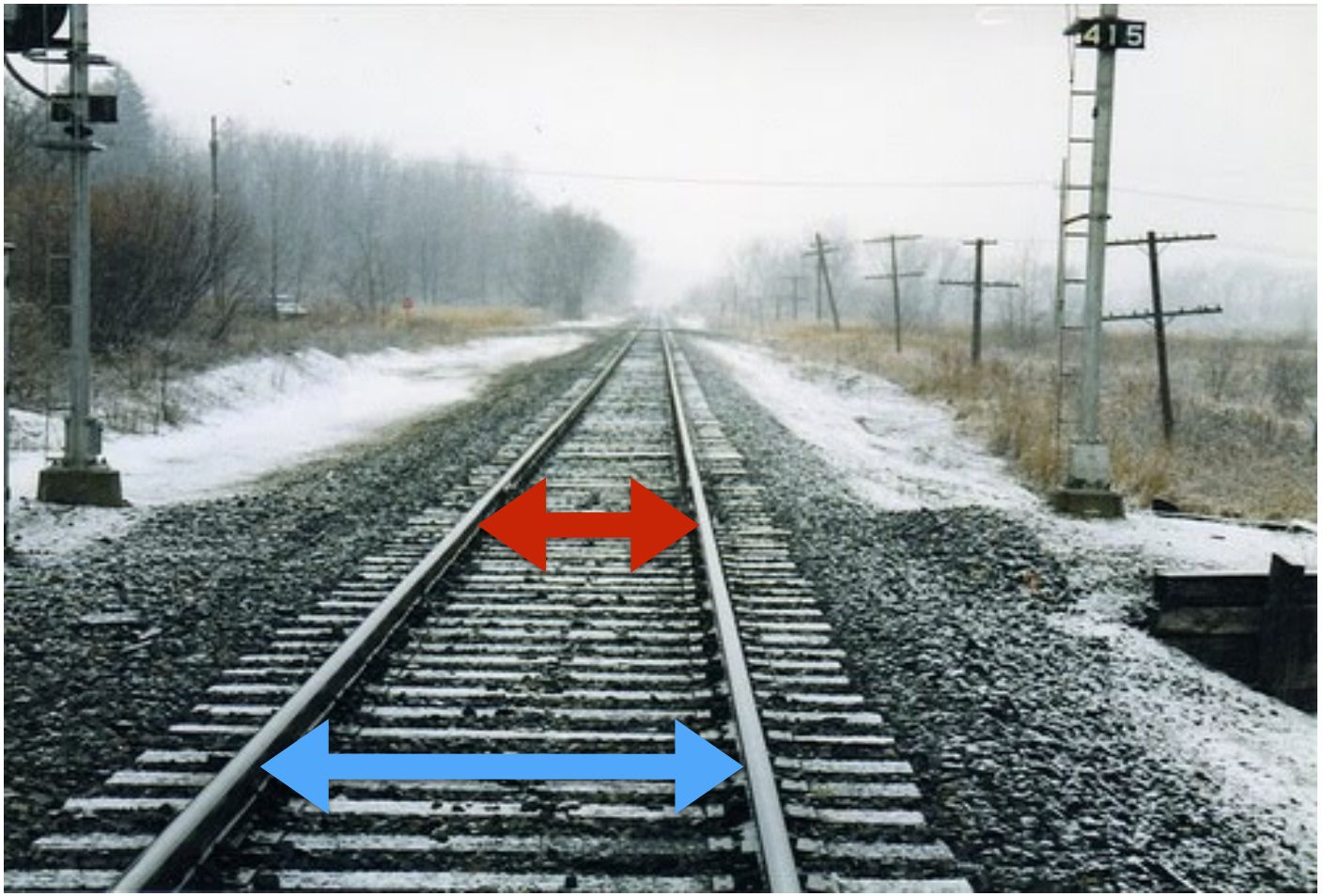
Are these failures of our vision system?

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!

Slide credit: J. Koenderink

Cues: Linear perspective



http://kalisdigitalphotos.blogspot.com

Parallel lines merge at the horizon

Analyzing parallel lines to estimate space

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*.

Cues: Occlusion ordering

Chicago loop, image source: **wikipedia**

Cues: texture gradient

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

Cues: shading and lighting

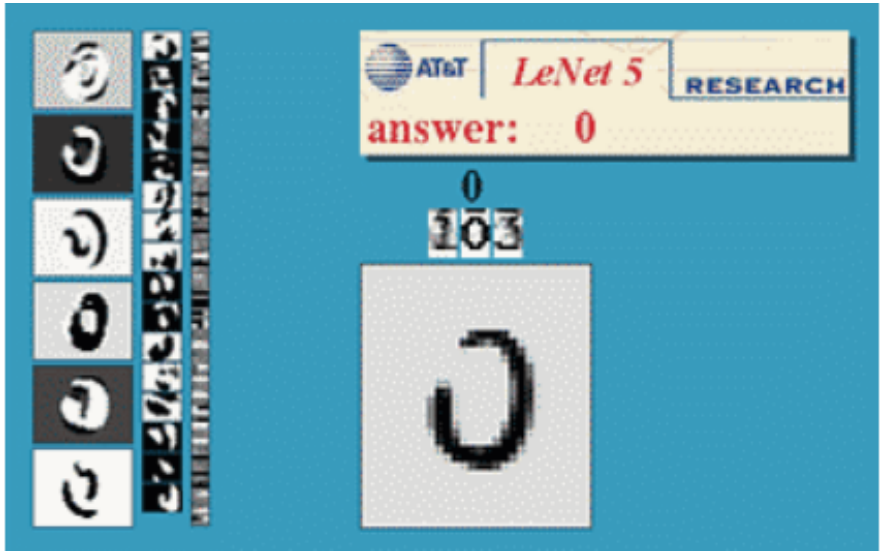
“The four seasons” sculpture set

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.

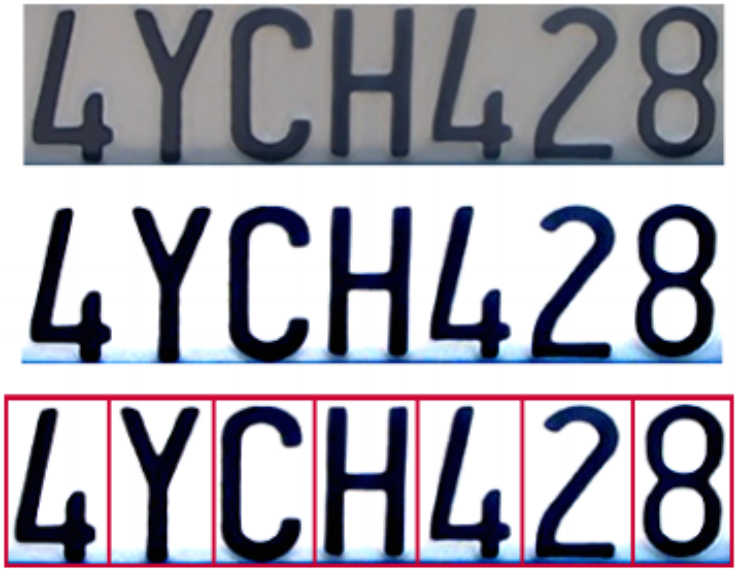
Some examples of successful computer vision applications …

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/

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

Face detection

Face detection is on many cameras these days

Face recognition

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

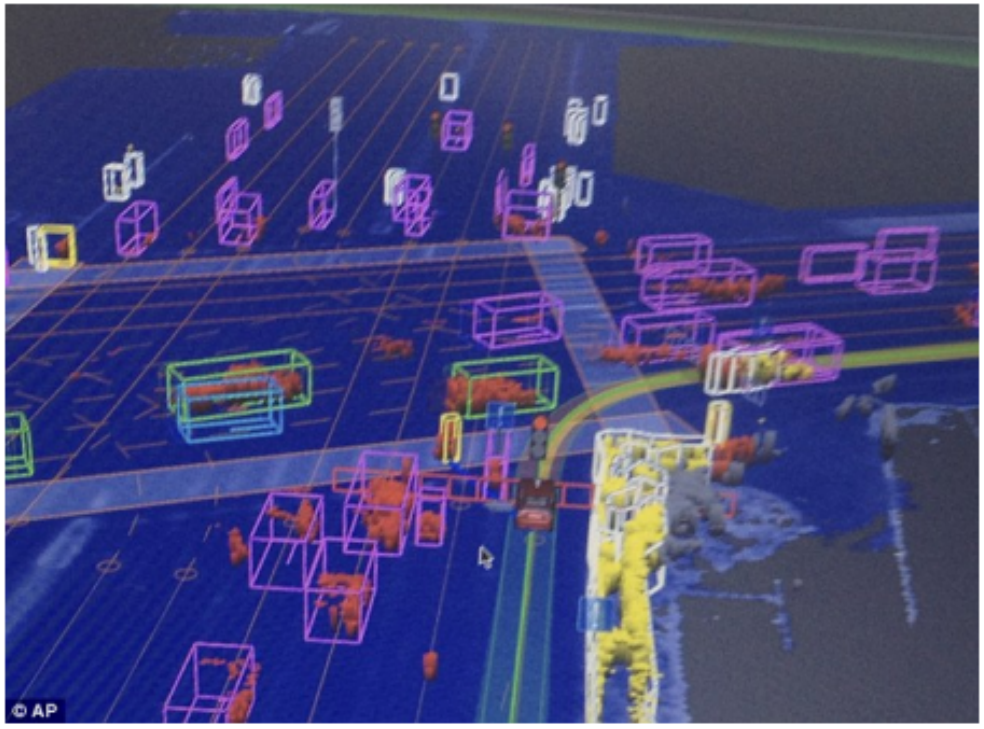
Instance recognition

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

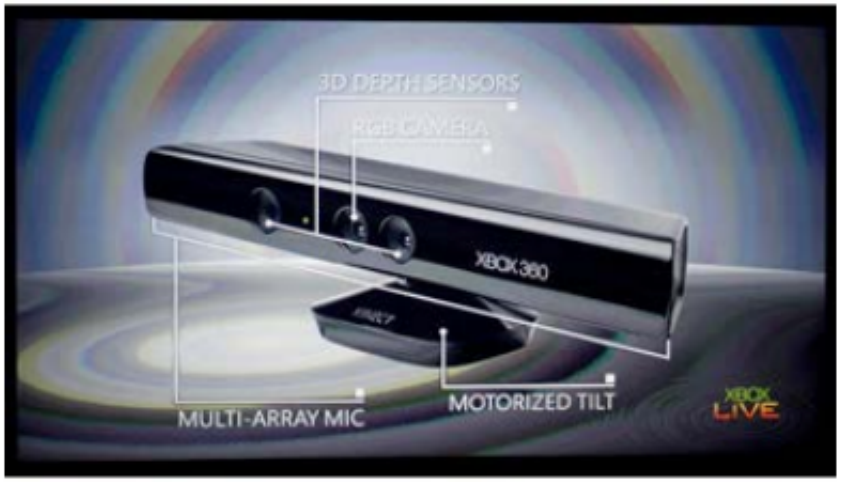
Self-driving cars

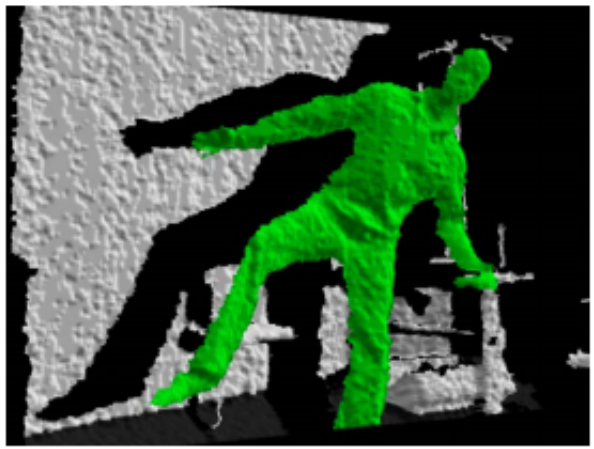


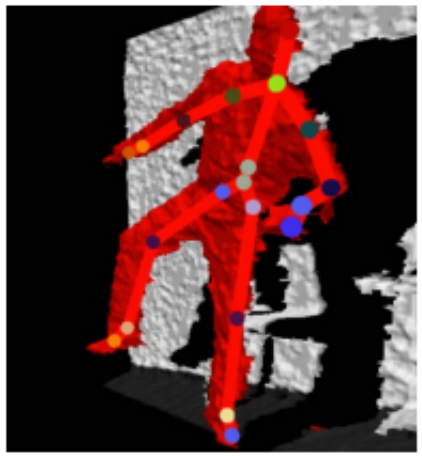
Source: L. Lazebnik

Interactive interfaces

Microsoft Kinect depth sensors





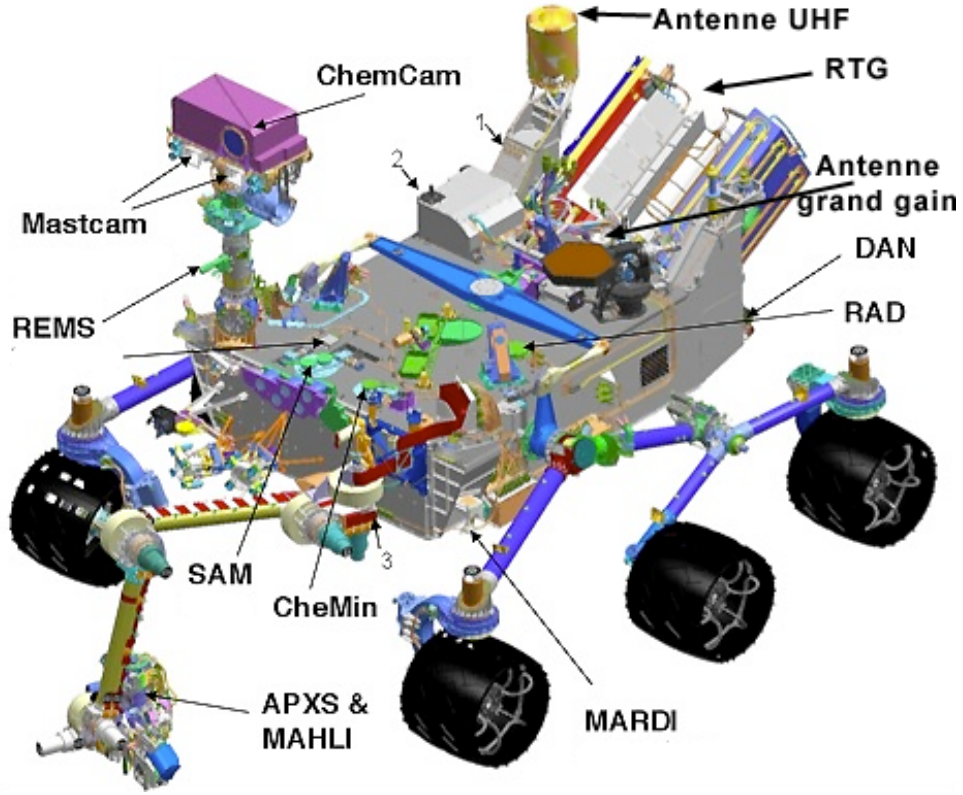


Large-scale 3D reconstruction

Photo Tourism: Exploring Photo Collections in 3D

YouTube link

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)

Course details

- Course overview
  I. **Early vision:** image formation, sensing, light and shading
  II. **Basic image processing:** digitizing images, linear filtering and applications such as line, corner and blob detection
  III. **Recognition:** model fitting, image representations, simple classifiers, convolutional neural networks, applications
  IV. **Additional topics** (time permitting)

- 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

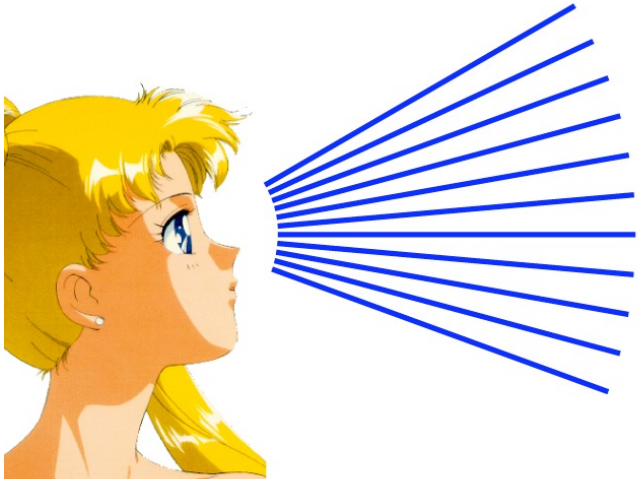
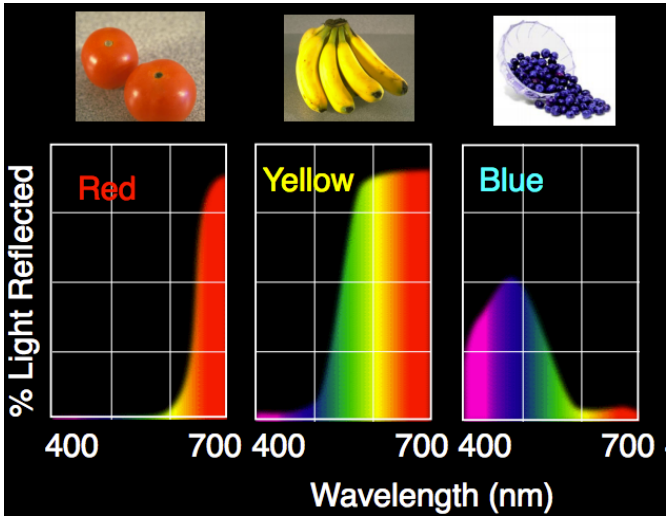


image formation

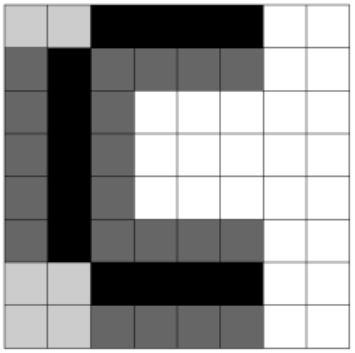


color perception

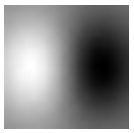
II. Basic image processing

100	100	0	0	0	0	255	255
50	0	50	50	50	50	255	255
50	0	50	255	255	255	255	255
50	0	50	255	255	255	255	255
50	0	50	255	255	255	255	255
50	0	50	50	50	50	255	255
100	100	0	0	0	0	255	255
100	100	50	50	50	50	255	255

(a)



(b)



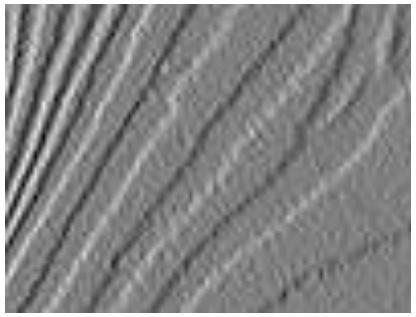
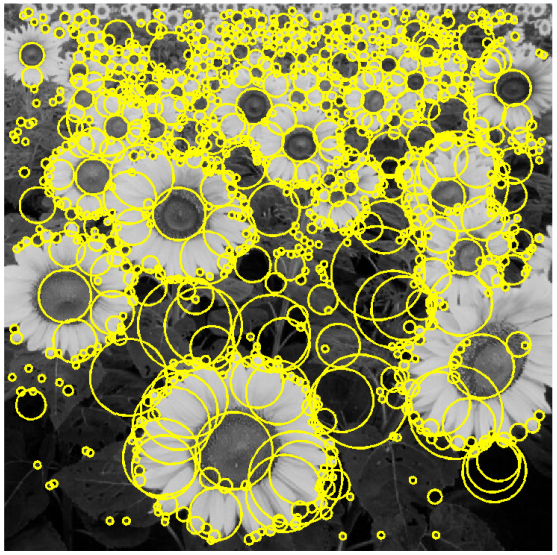


image representation

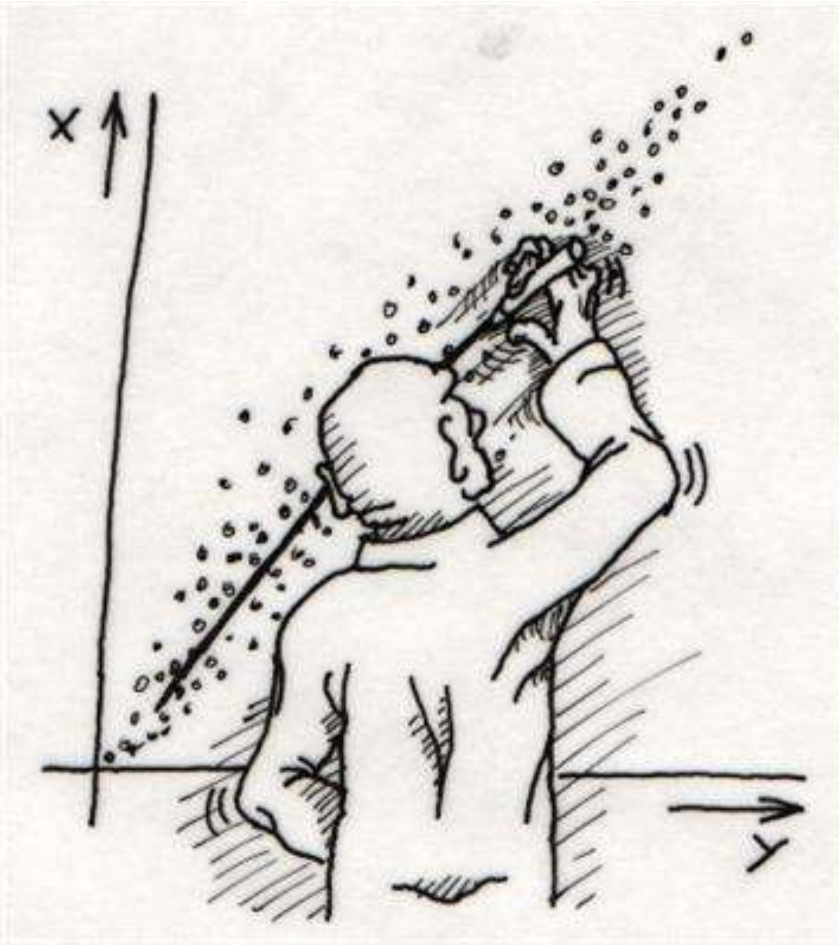
image filtering



corner and blob detection

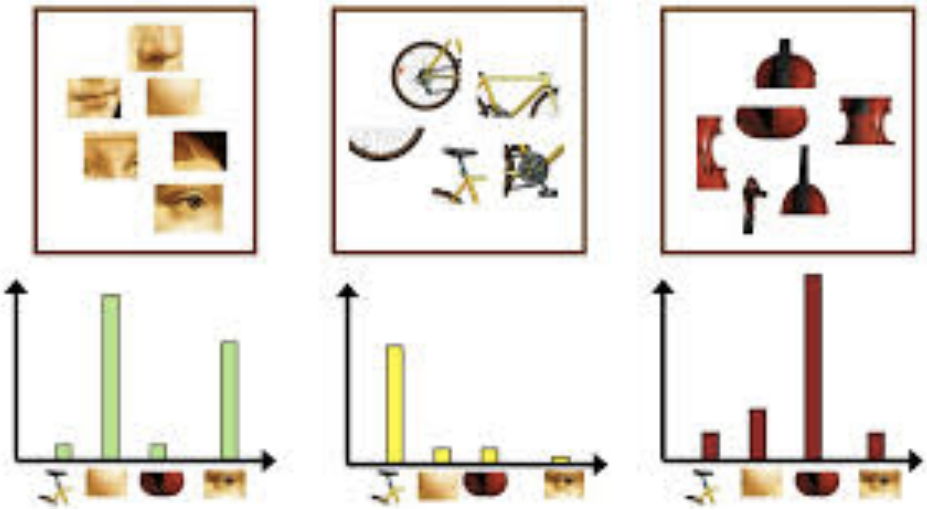
Source: L. Lazebnik

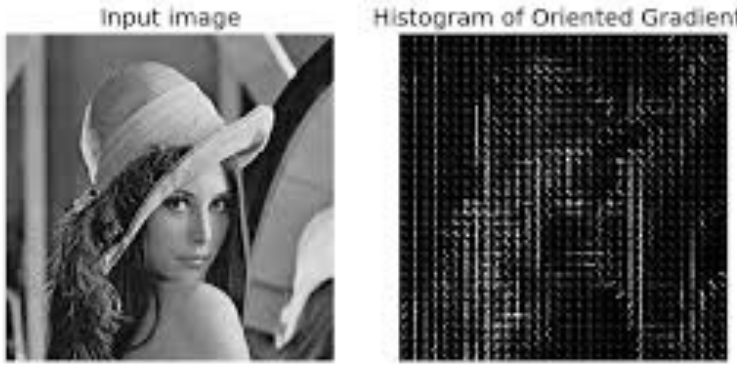
III. Recognition



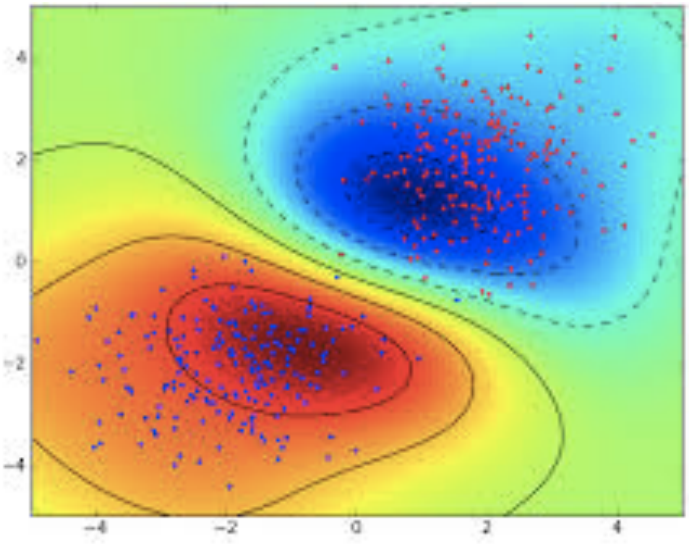
model fitting

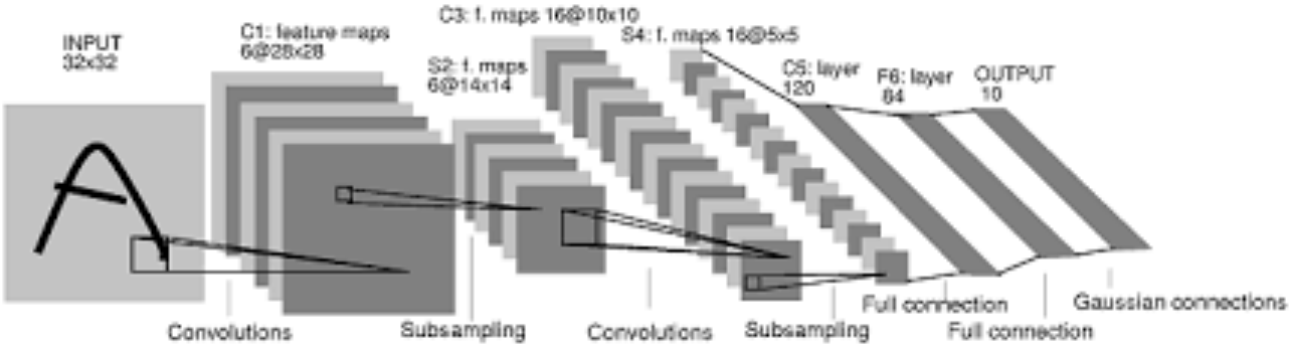
III. Recognition





designing image representations

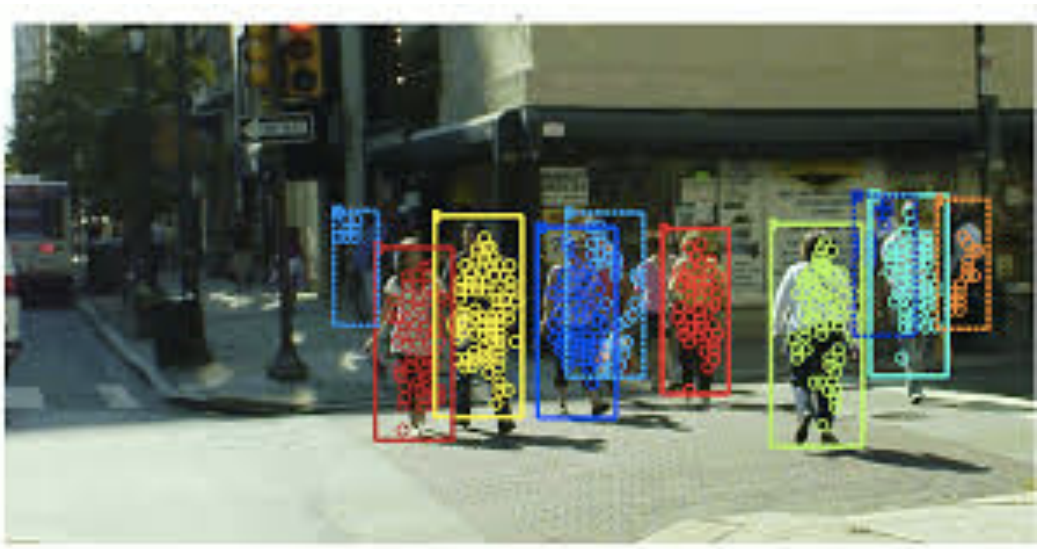




machine learning

V. Additional topics

Optical flow



Tracking

Or something else?

For next class …

- Finish and submit homework 00
- Readings:
  - **The speed of processing in the human visual system**, Thorpe et al., Letters to Nature, 1996
  - Chapter 1 in Richard Szeliski (RS) textbook