CMPSCI 670: Computer Vision

Introduction

University of Massachusetts, Amherst

September 3, 2014

Instructor: Subhransu Maji

Administrivia

- **Lectures**:
  - Monday/Wednesday, 2:30 - 3:45, CS 142
- **Instructor**:
  - Subhransu Maji
  - Office hours: Wednesday, 3:45 - 4:45, CS 274
    - This week: Thursday 3:45 - 4:45, CS 274
- **Course website**
  - News, homework and lectures will be posted here
  - **http://people.cs.umass.edu/~smaji/teaching/fall14/cmpsci670.html**

Administrivia

- **Grading policy**:
  - **10%** class participation (readings, discussions, questions, attendance)
  - **90%** homework (five mini-projects) and a final project
    - final project will be graded as two homework
    - worst grade will be dropped
- Final project (alone or in groups of two)
  - guidelines will be posted in October
- **Course textbooks (recommended)**:
  - Forsyth and Ponce, Computer Vision: A Modern Approach, 2nd edition
  - 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)

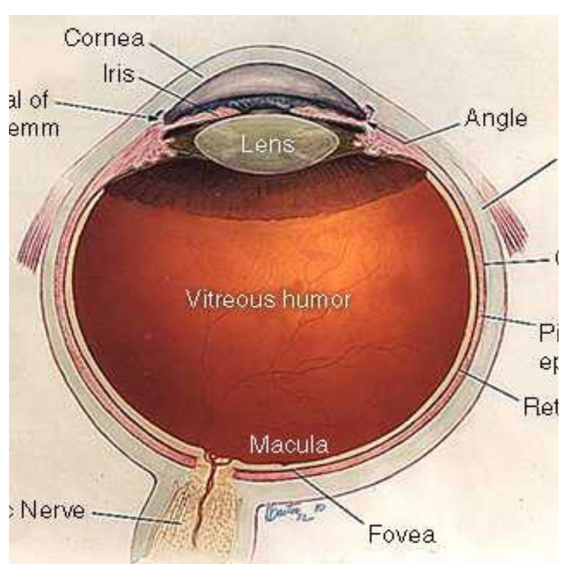
Academic integrity policy

- **Feel free to**:
  - discuss assignments with each other, but coding must be done individually (except for the final project).
  - incorporate code or tips you find on the Web, provided this doesn't make the assignment trivial and you explicitly acknowledge your sources.

Why Vision?

Why Vision? Light!

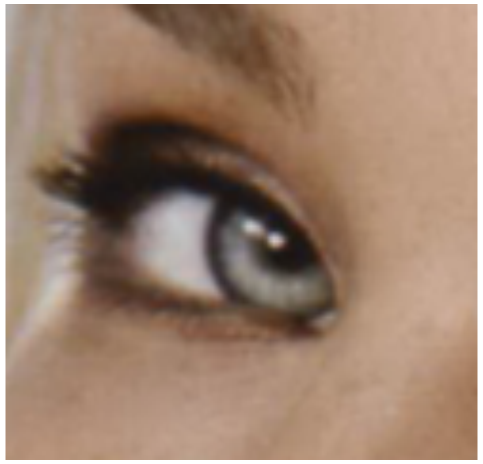
It is how we see other people,
navigate our environment,
communicate ideas, entertain,
and **measure** the world around us.



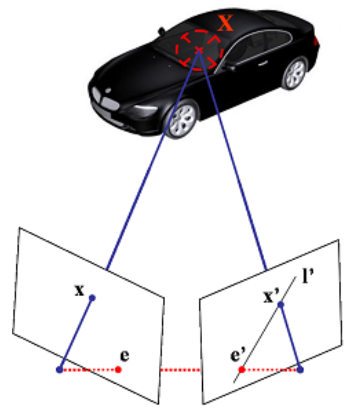
Source: A. Berg

Why is light good for measurement?

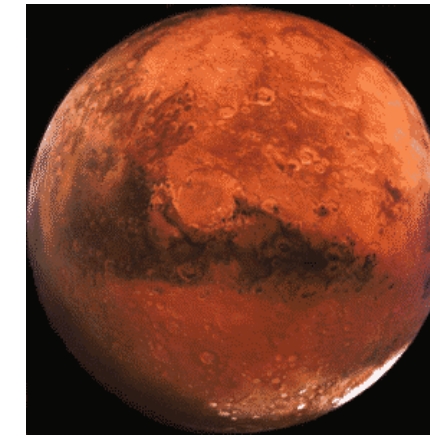
Microscopy



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)

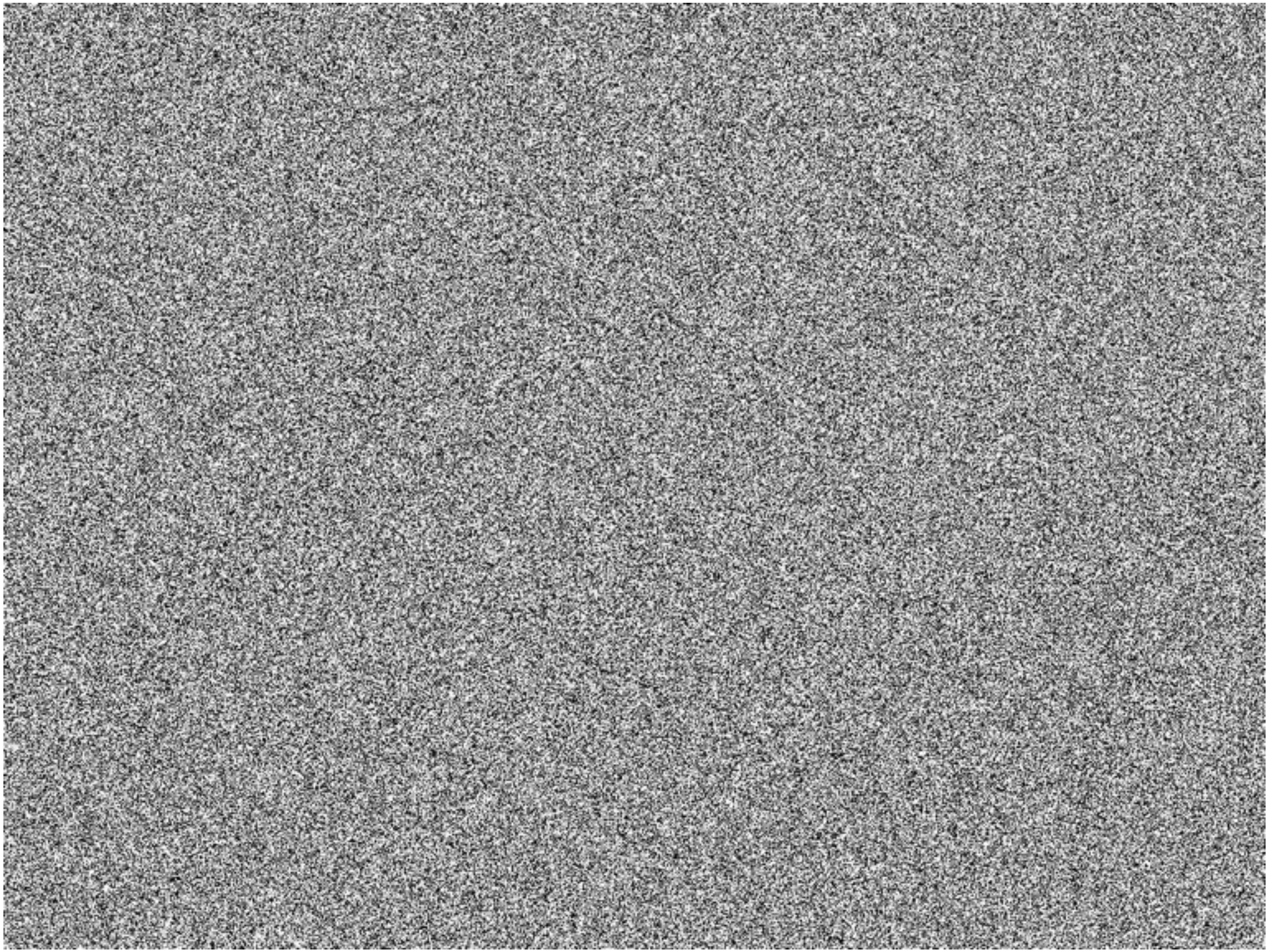
Source: A. Berg

The goal of computer vision

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

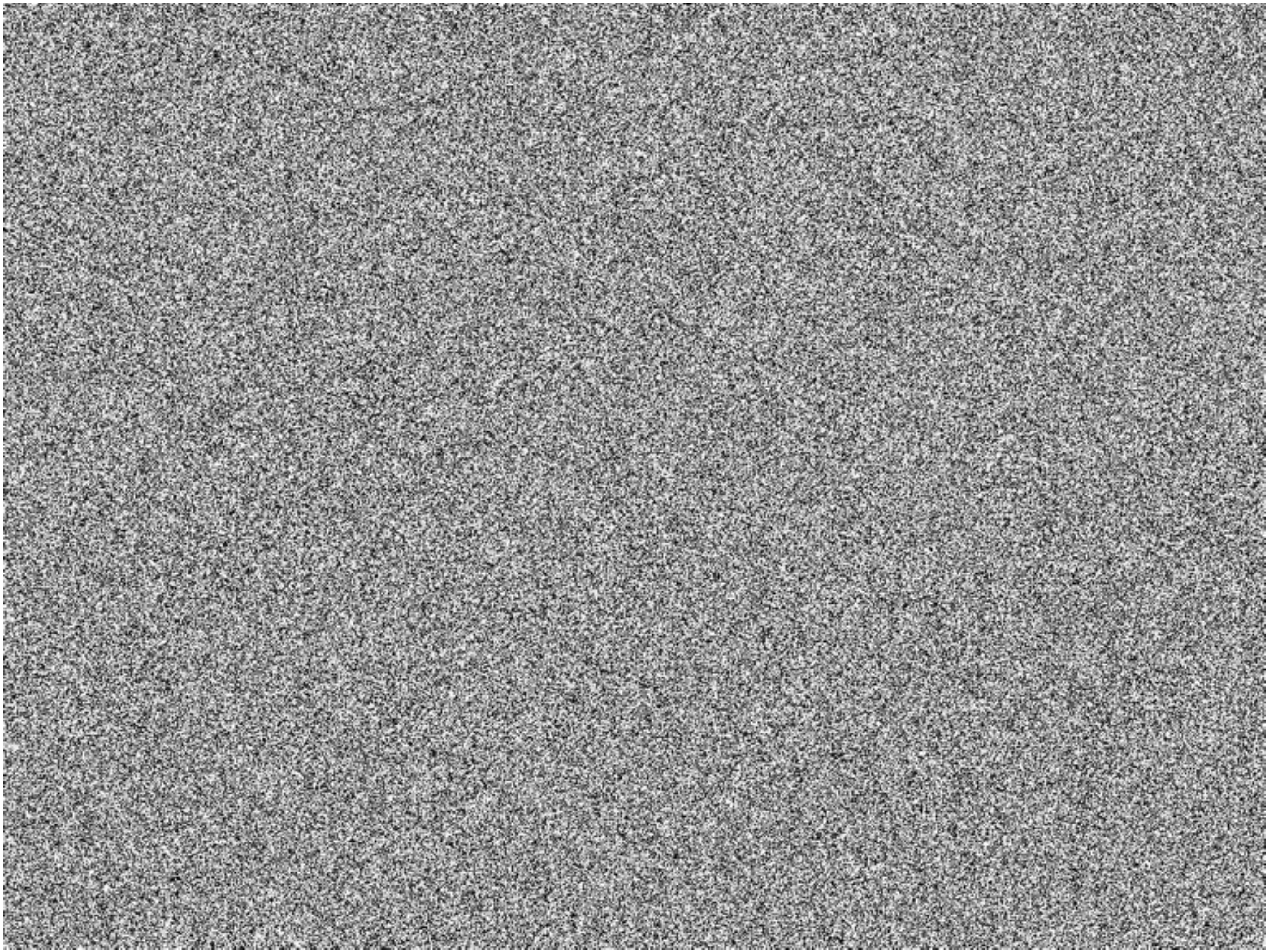
We are remarkably good at this!

An experiment …#1



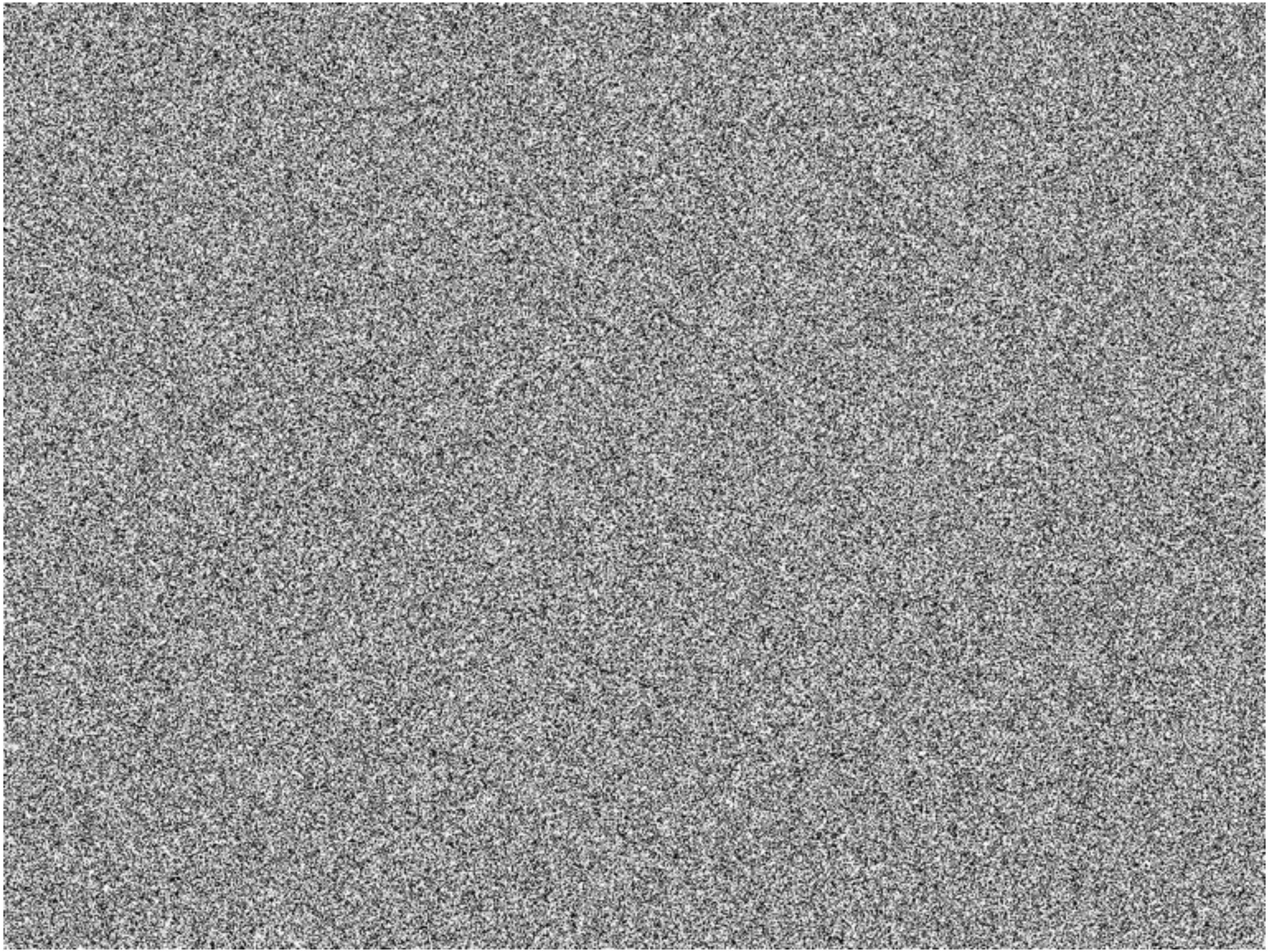
animal or not?

An experiment …#2



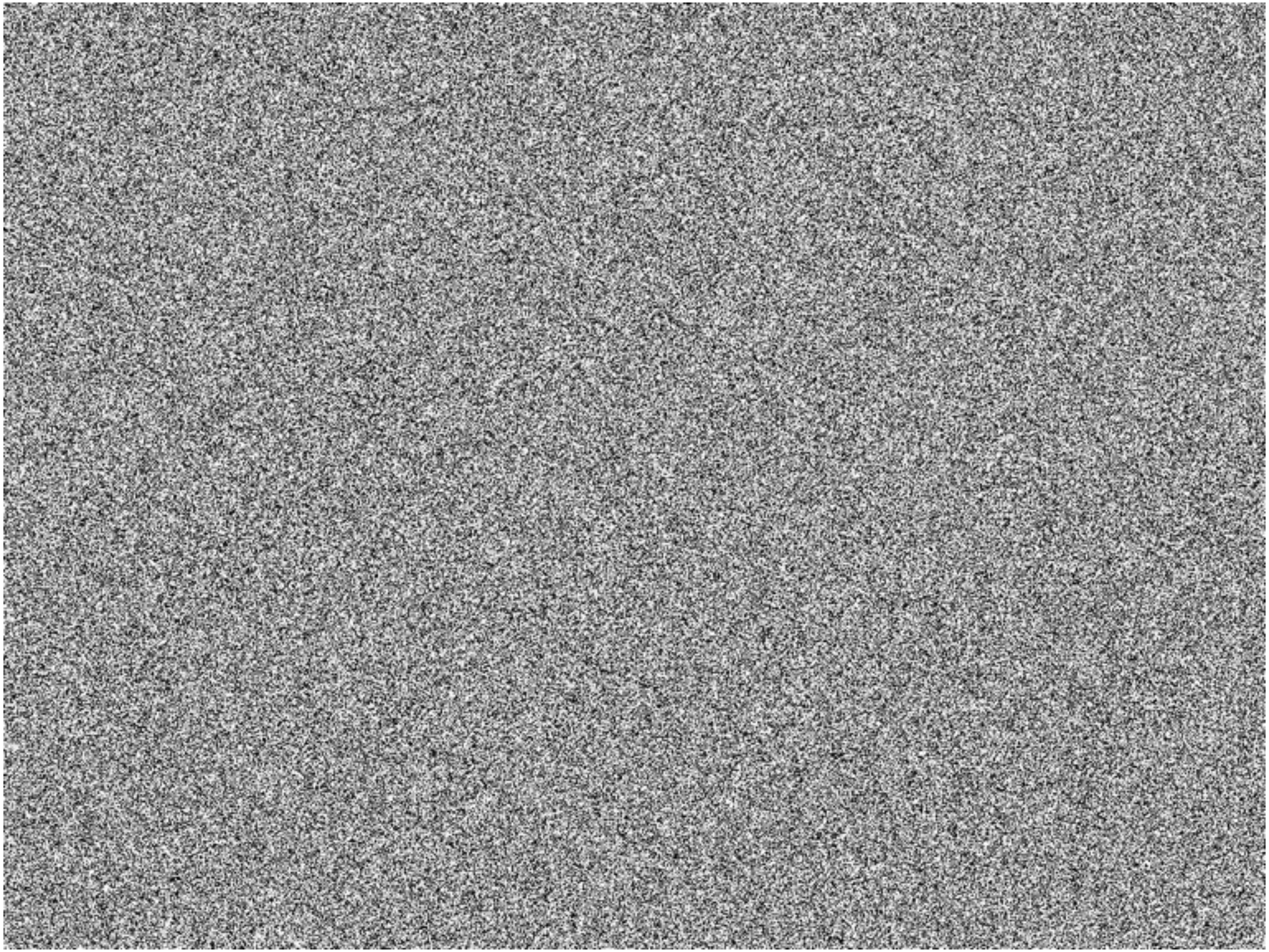
animal or not?

An experiment …#3



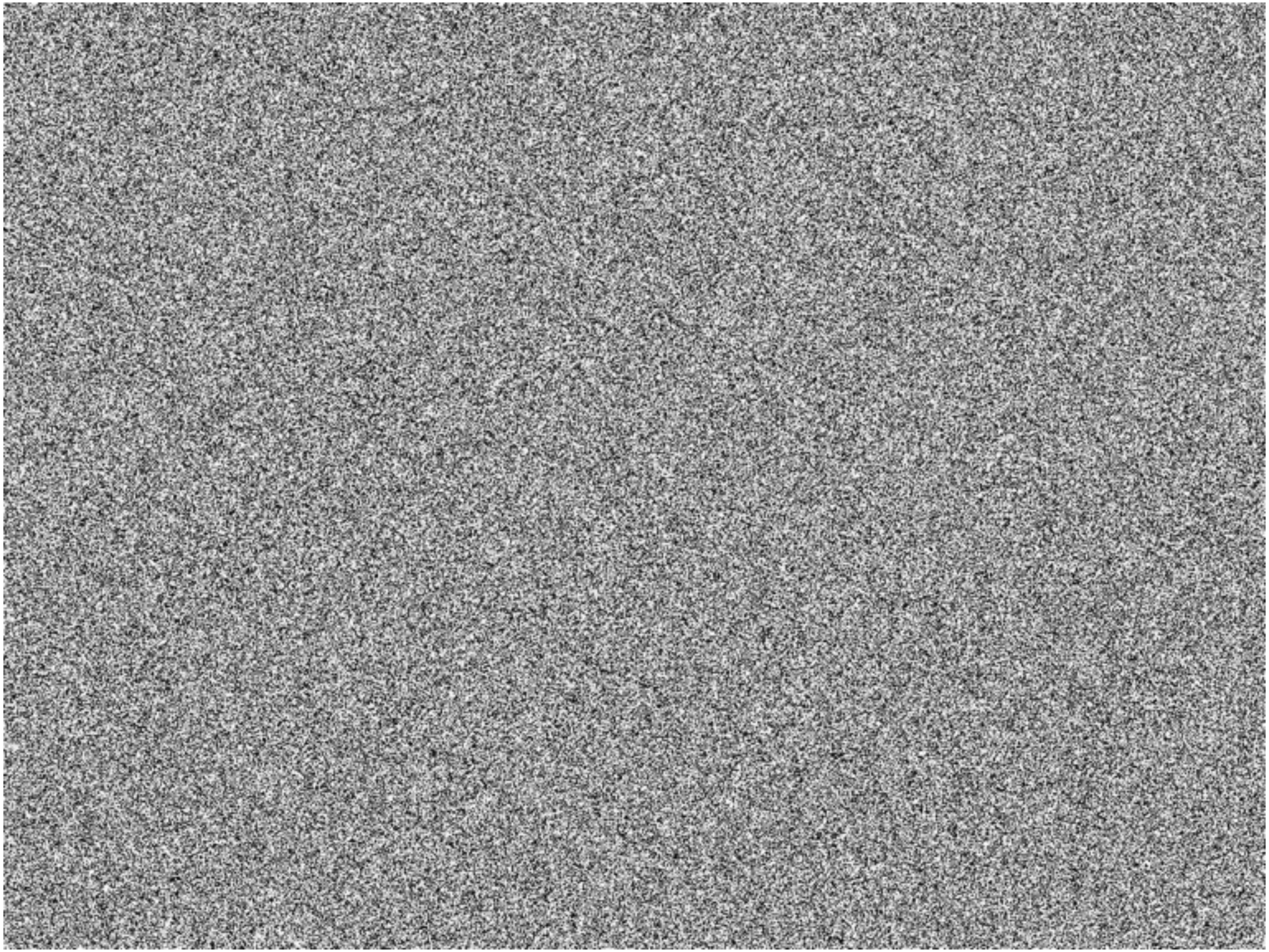
animal or not?

An experiment …#4



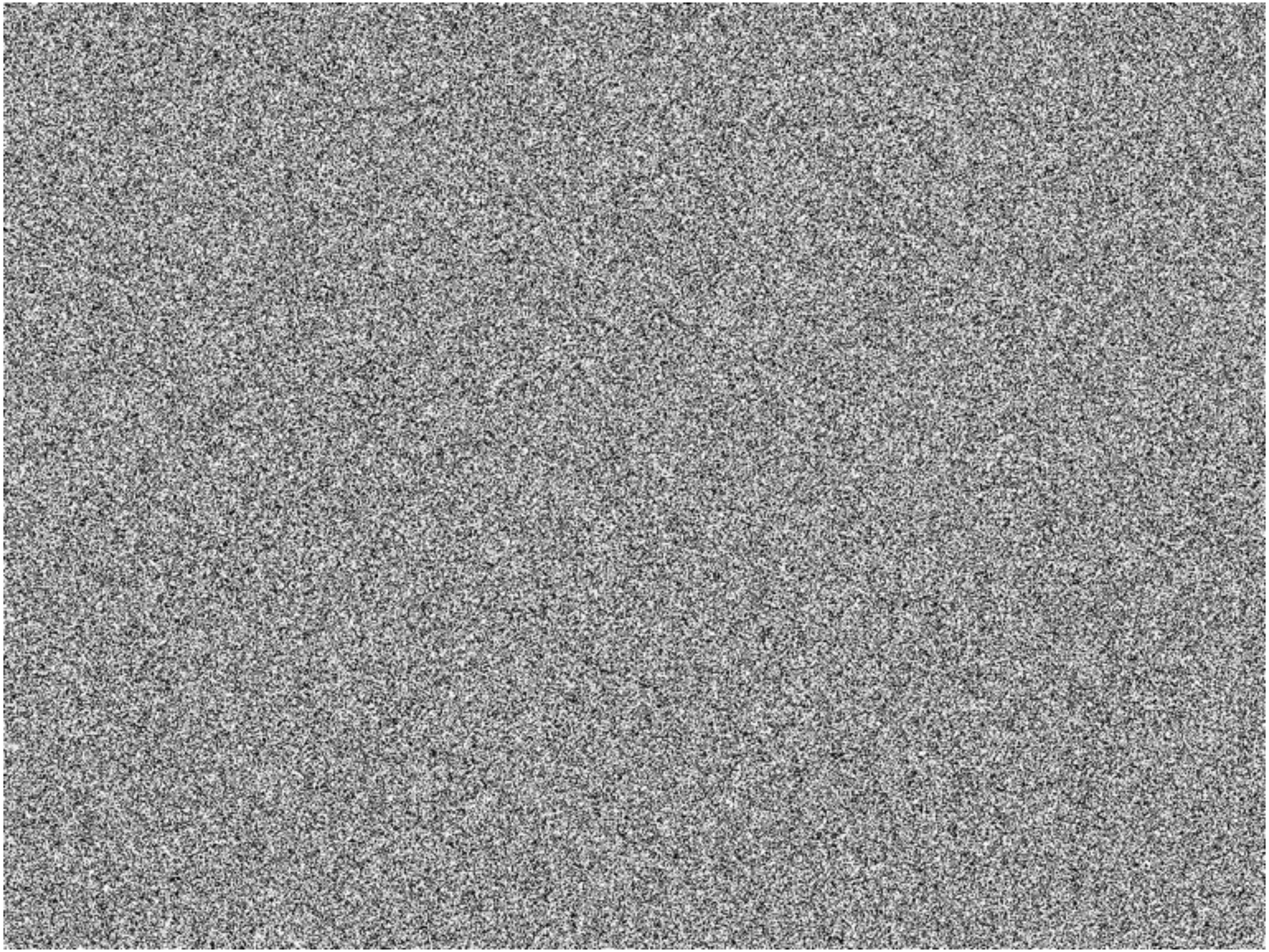
animal or not?

An experiment …#5



animal or not?

An experiment …#6



animal or not?

The images …

#1

#2

#3

#4

#5

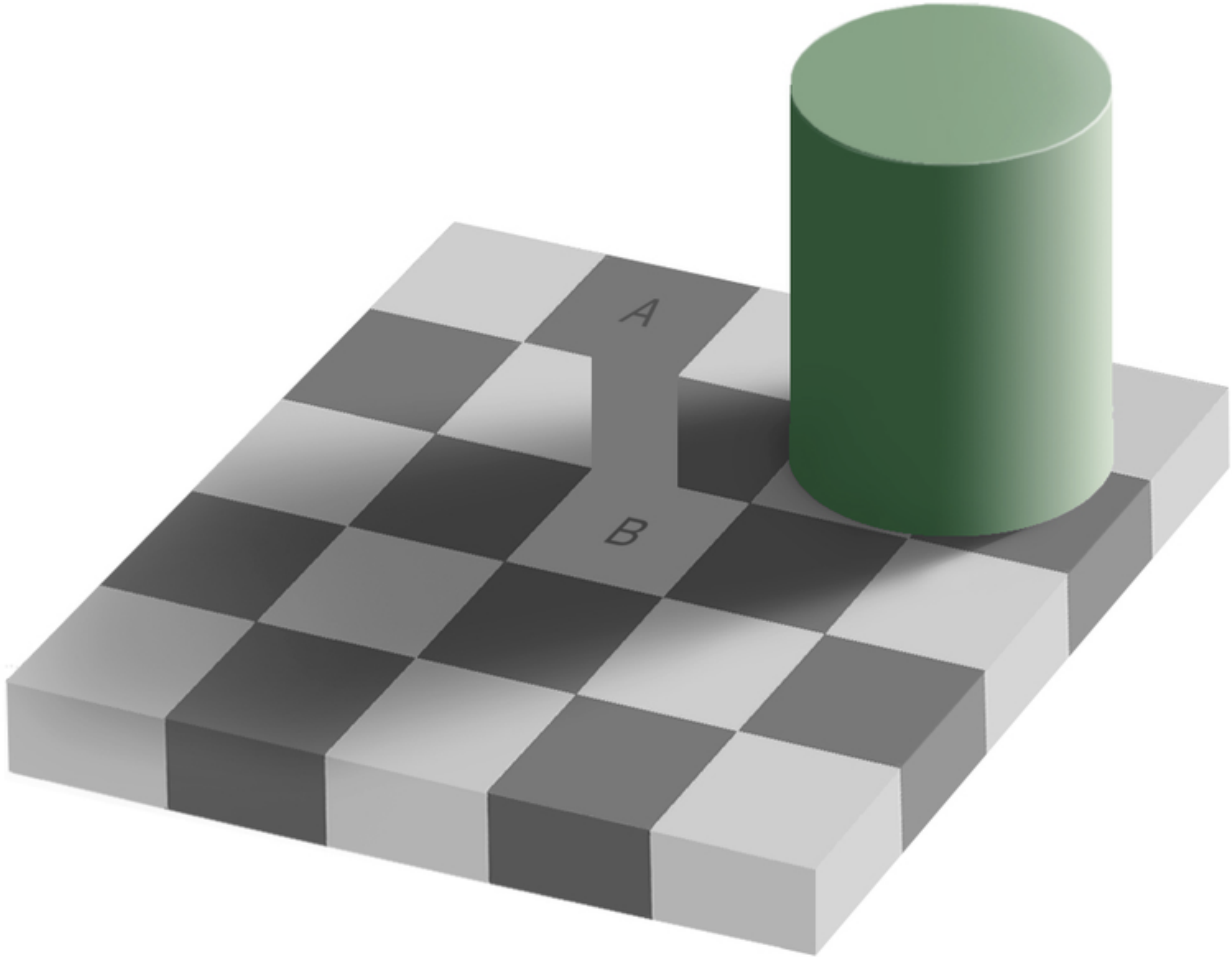
#6

Human vision

- Amazingly good, fast and accurate
- 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!

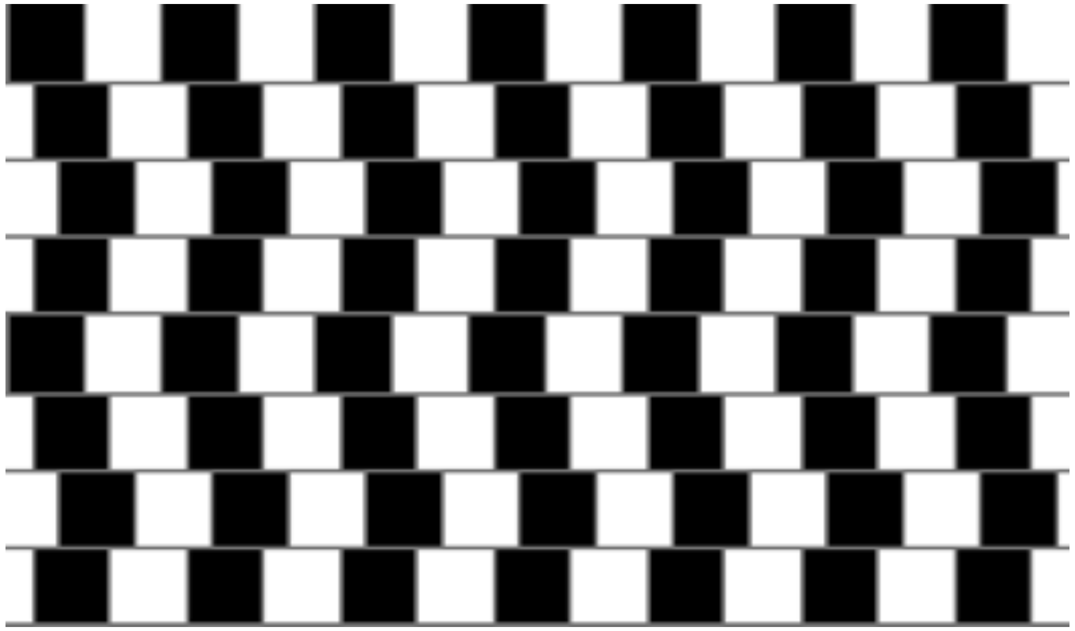
Source: A. Berg

But we make mistakes …

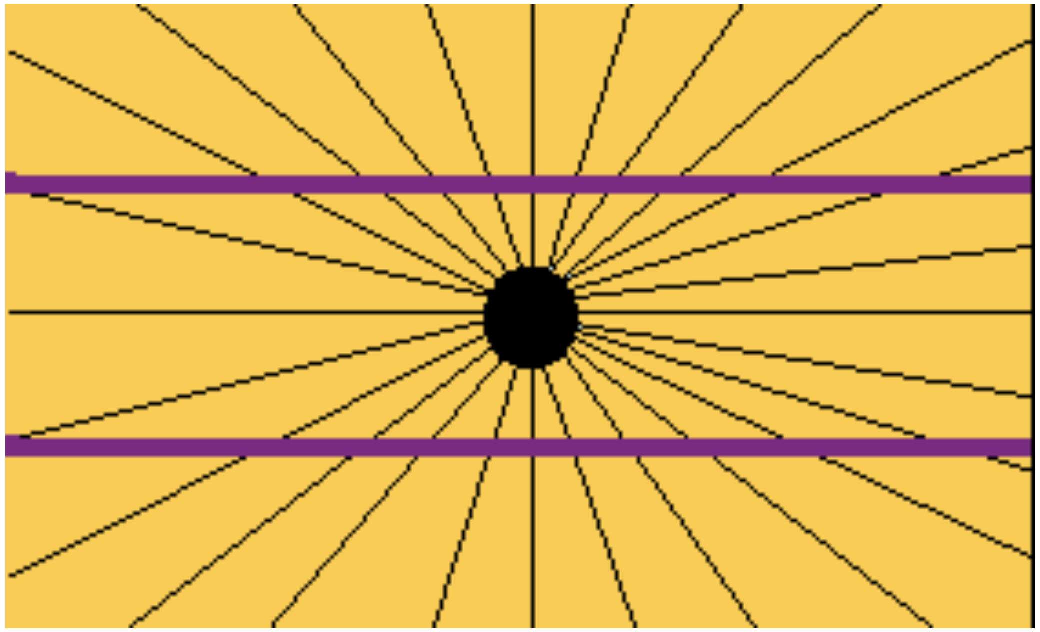


Checker shadow illusion - Edward H. Adelson

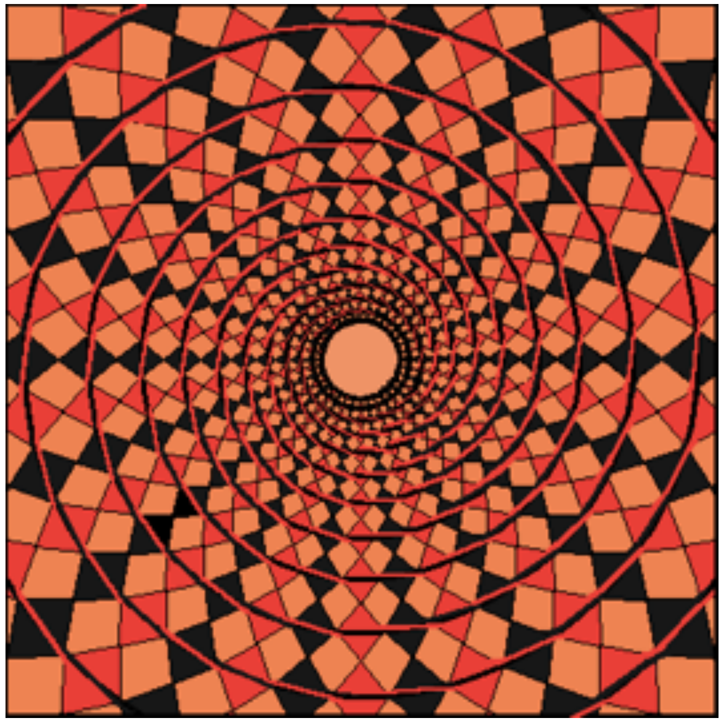
Other optical illusions



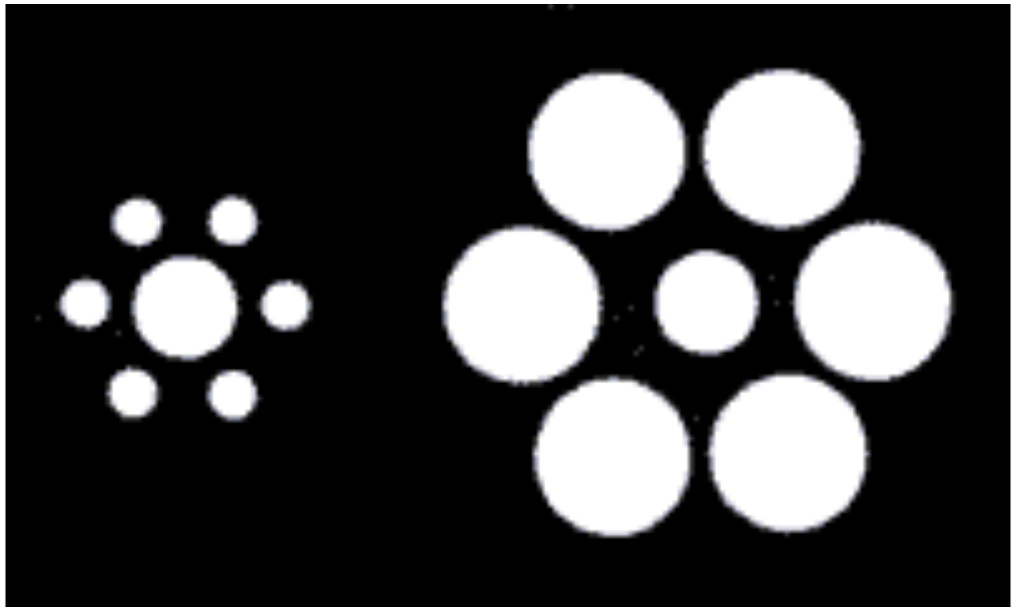
Are the horizontal lines parallel?



Are the purple lines straight?



Is this a spiral?



is the left circle (in the center) bigger?

Are these failures of our vision system?

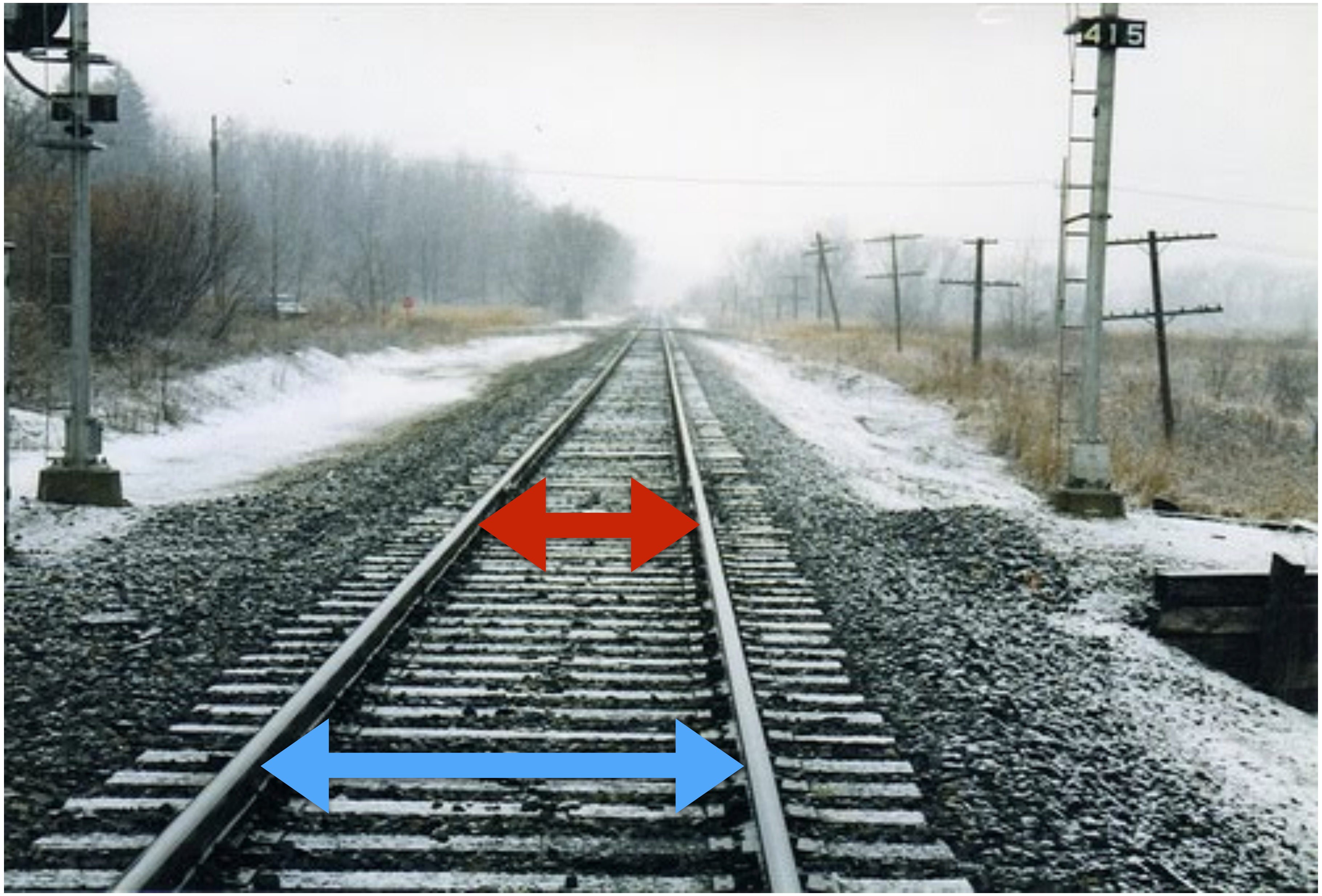
http://www.illusions.org

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)

Cues: Linear perspective



http://kalisdigitalphotos.blogspot.com

Parallel lines merge at the horizon

Analyzing parallel lines to estimate space

Cues: Aerial (Atmospheric) perspective

Scattering of skylight by particles in the air adds to the luminosity

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.

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.

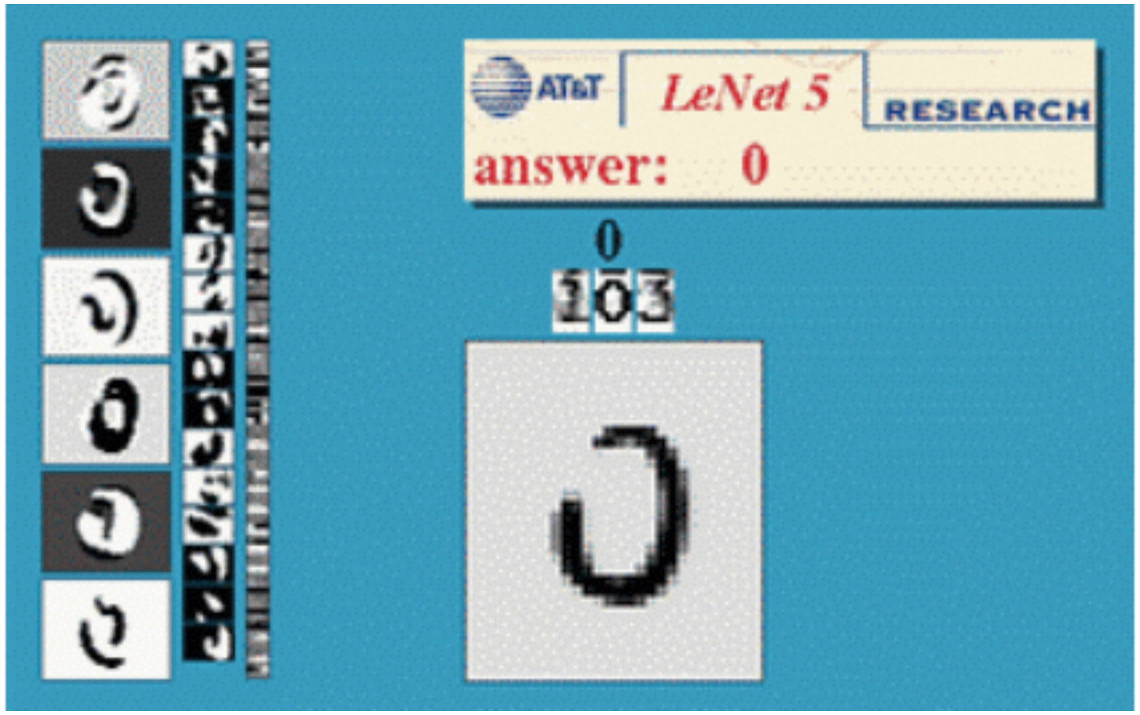
- It is hard, ad-hoc. There are few theorems, but we rely on those from many other areas: optics, geometry, physics, etc.

- 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)

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

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

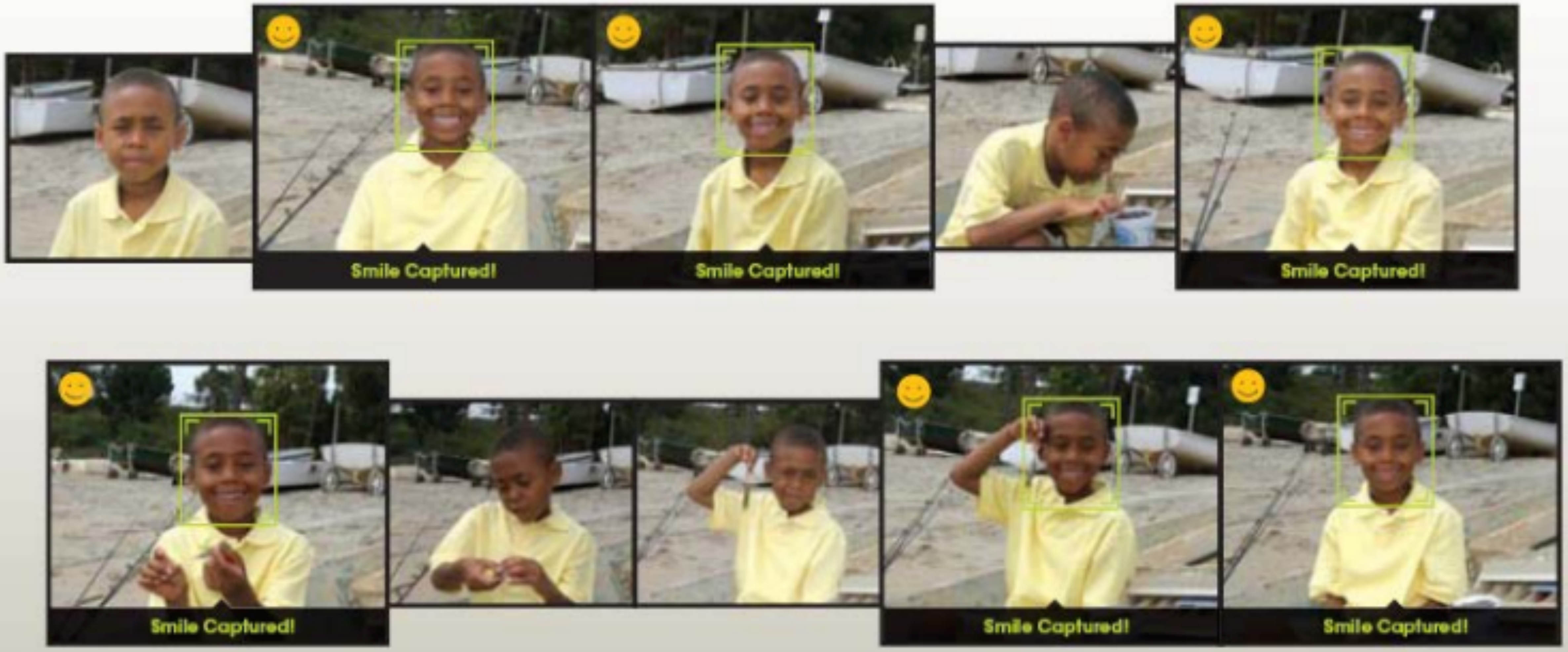
Face detection

Face detection is on many cameras these days

Smile detection

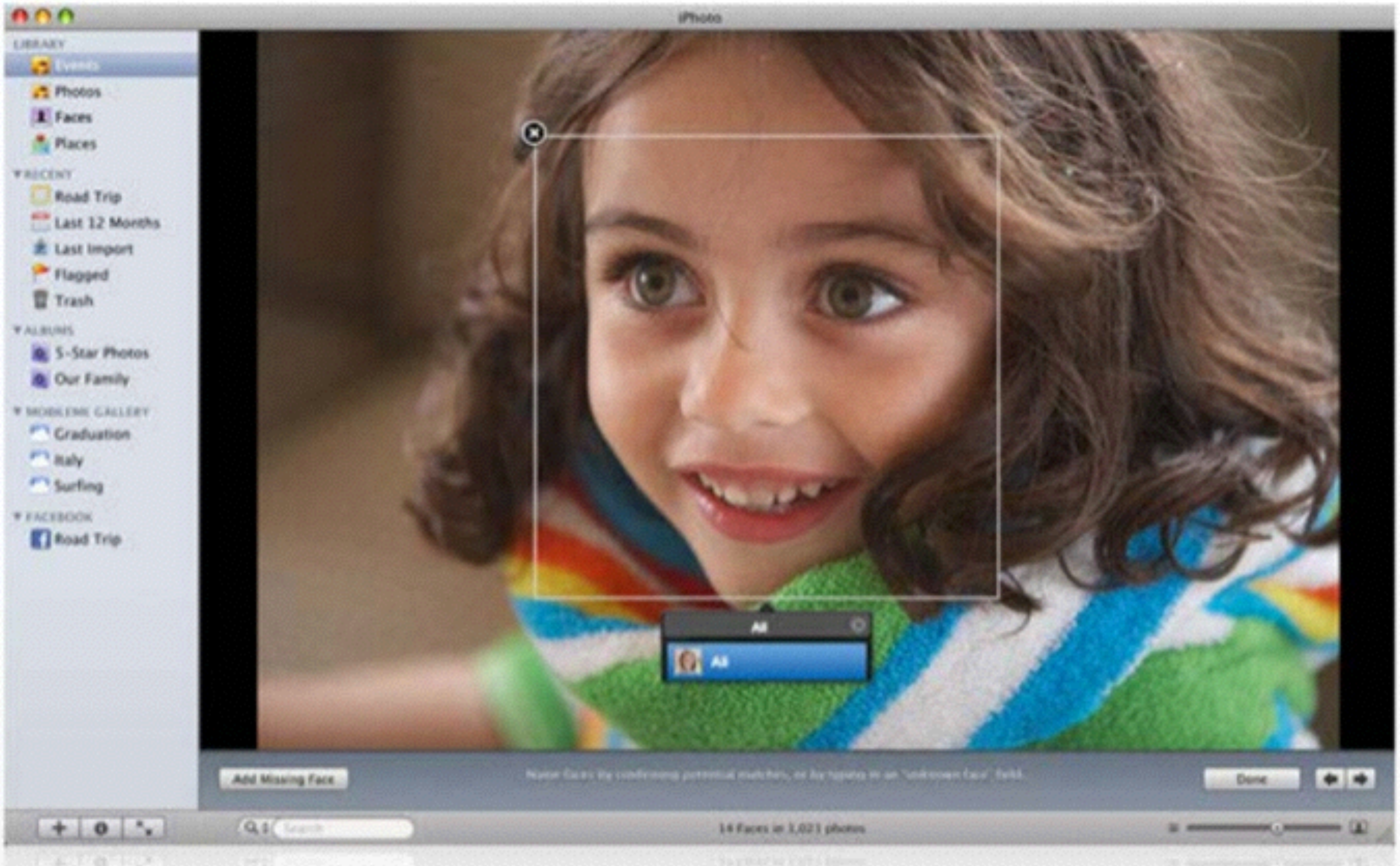
The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



Source: S. Seitz

Face recognition



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

Source: S. Seitz

Instance recognition

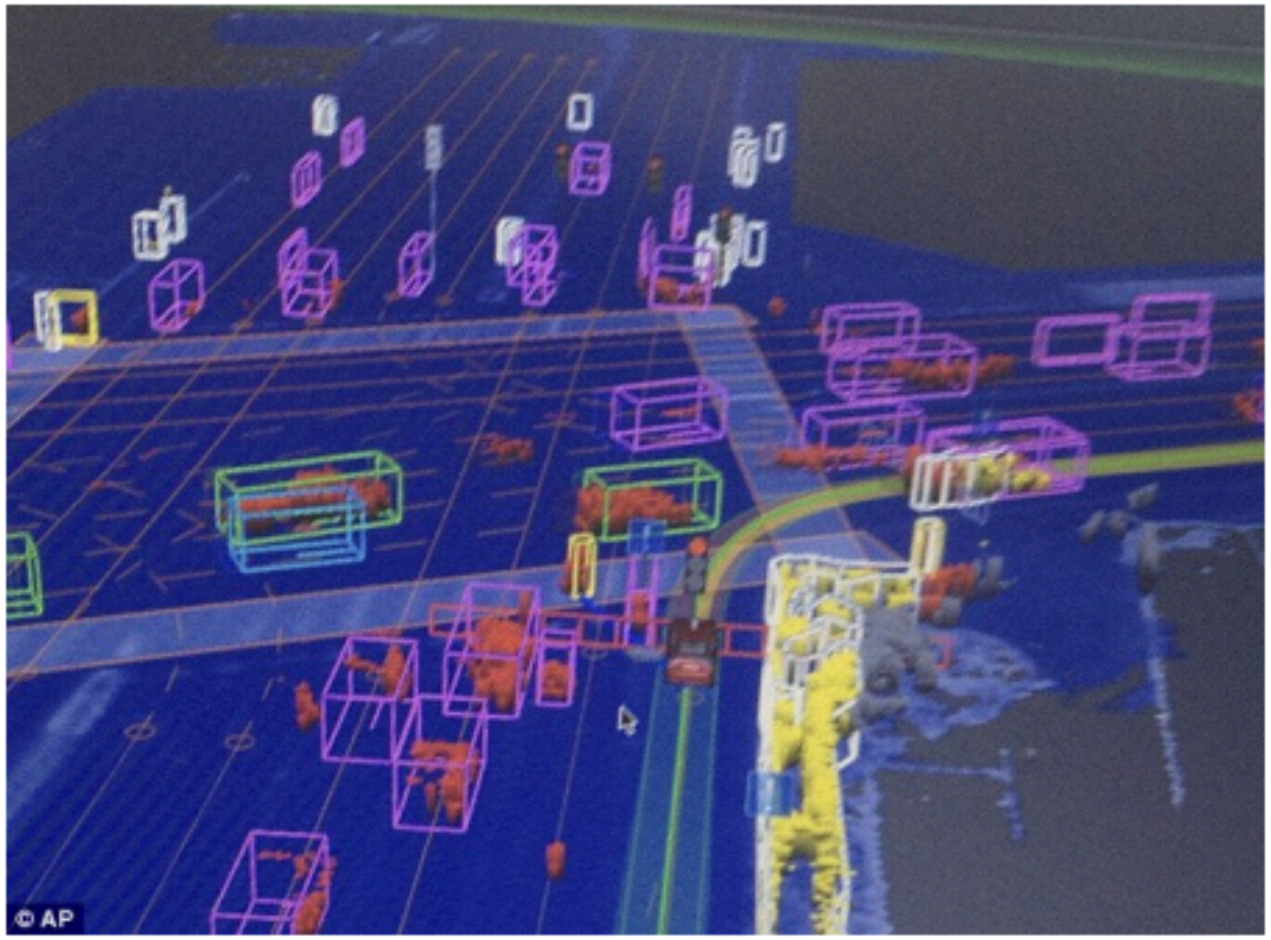
Source: S. Seitz

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

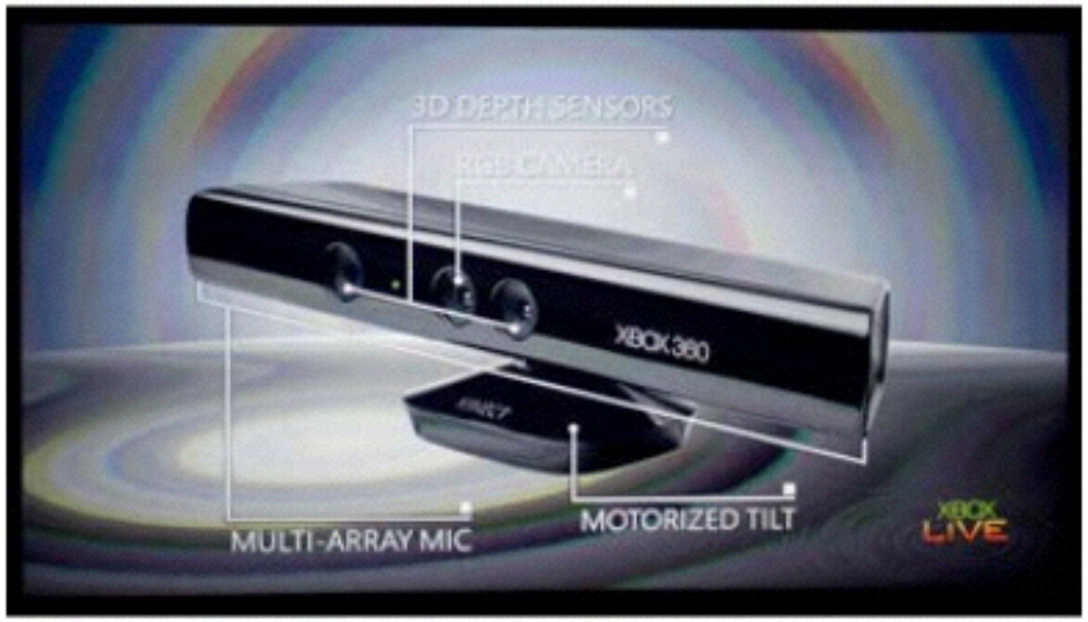
Self-driving cars

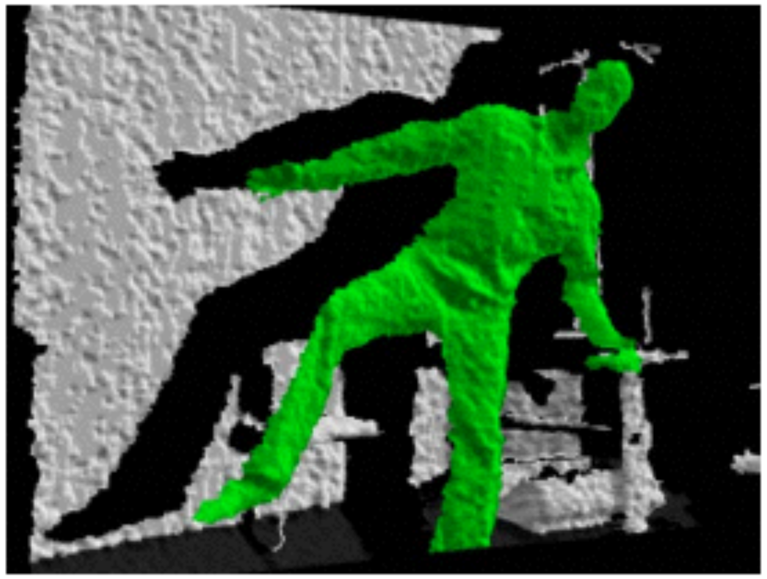
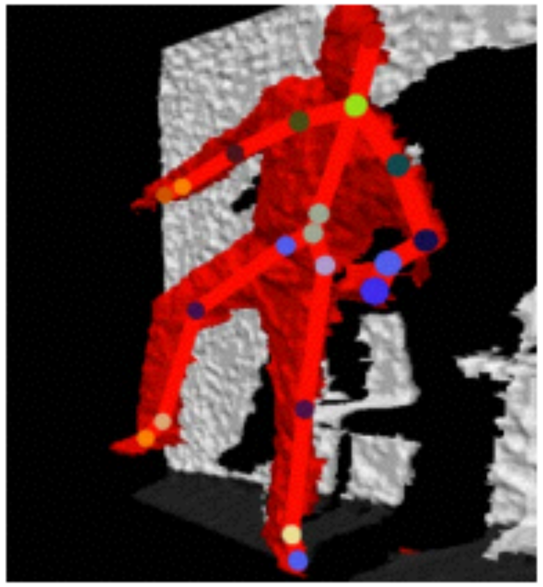


Source: L. Lazebnik

Interactive interfaces

Microsoft Kinect depth sensors

Source: L. Lazebnik

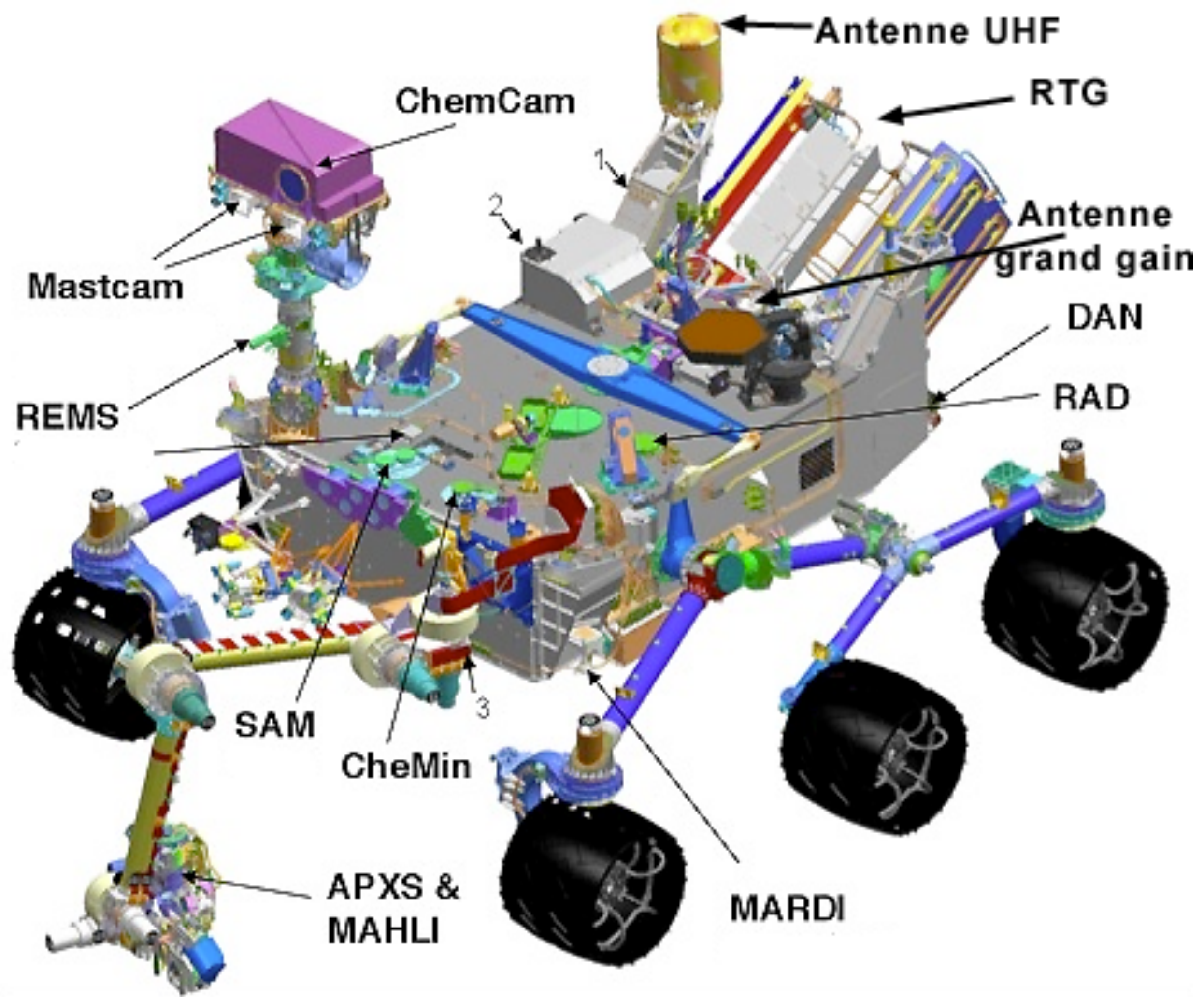
Large-scale 3D reconstruction

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

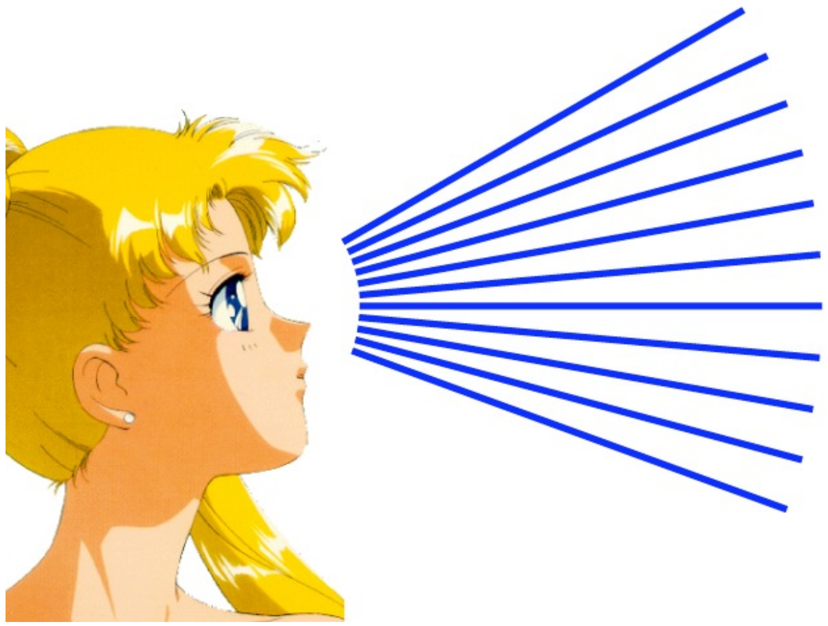


image formation

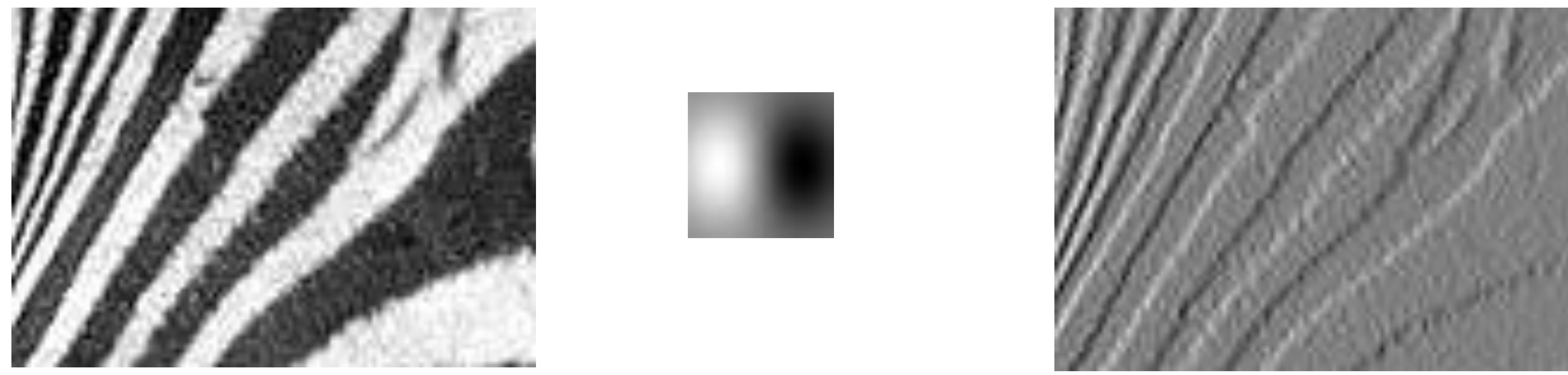
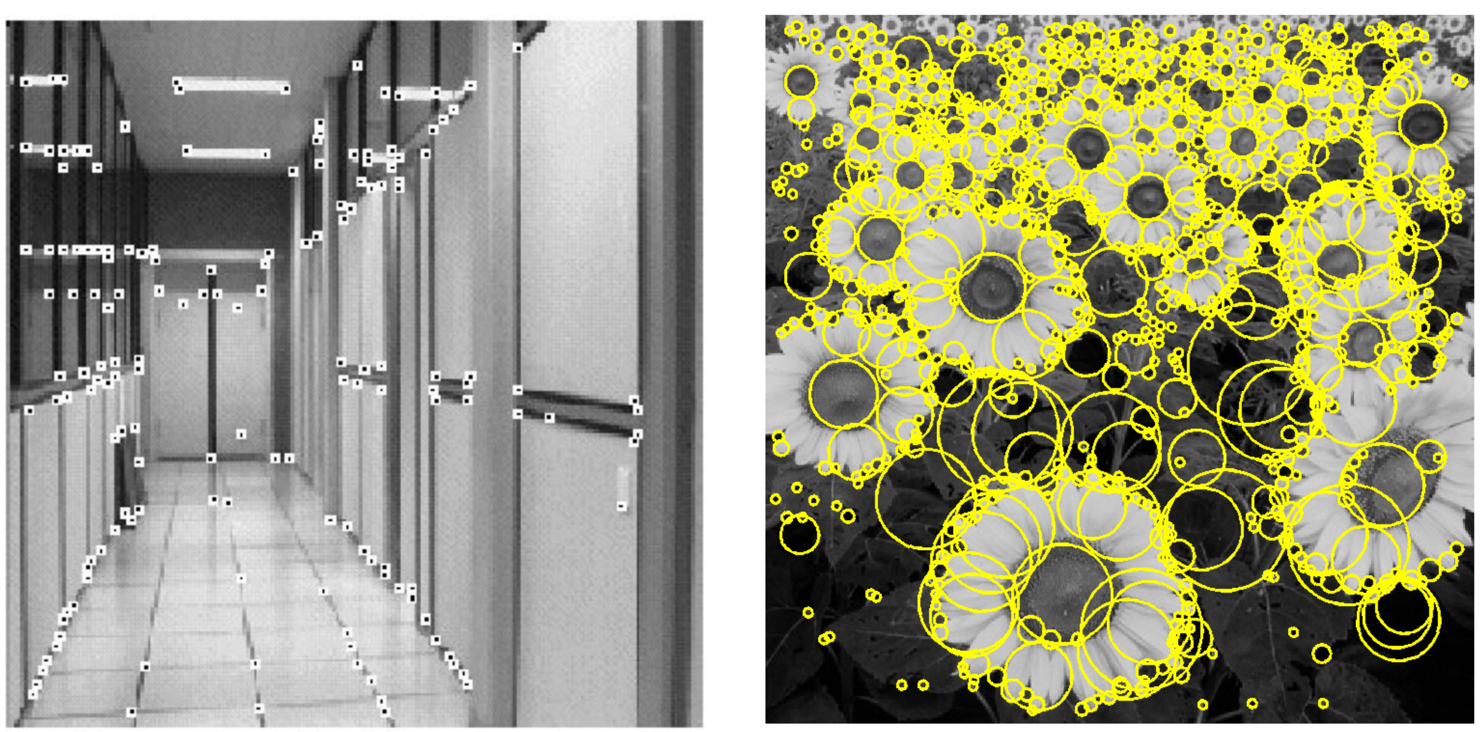


image filtering

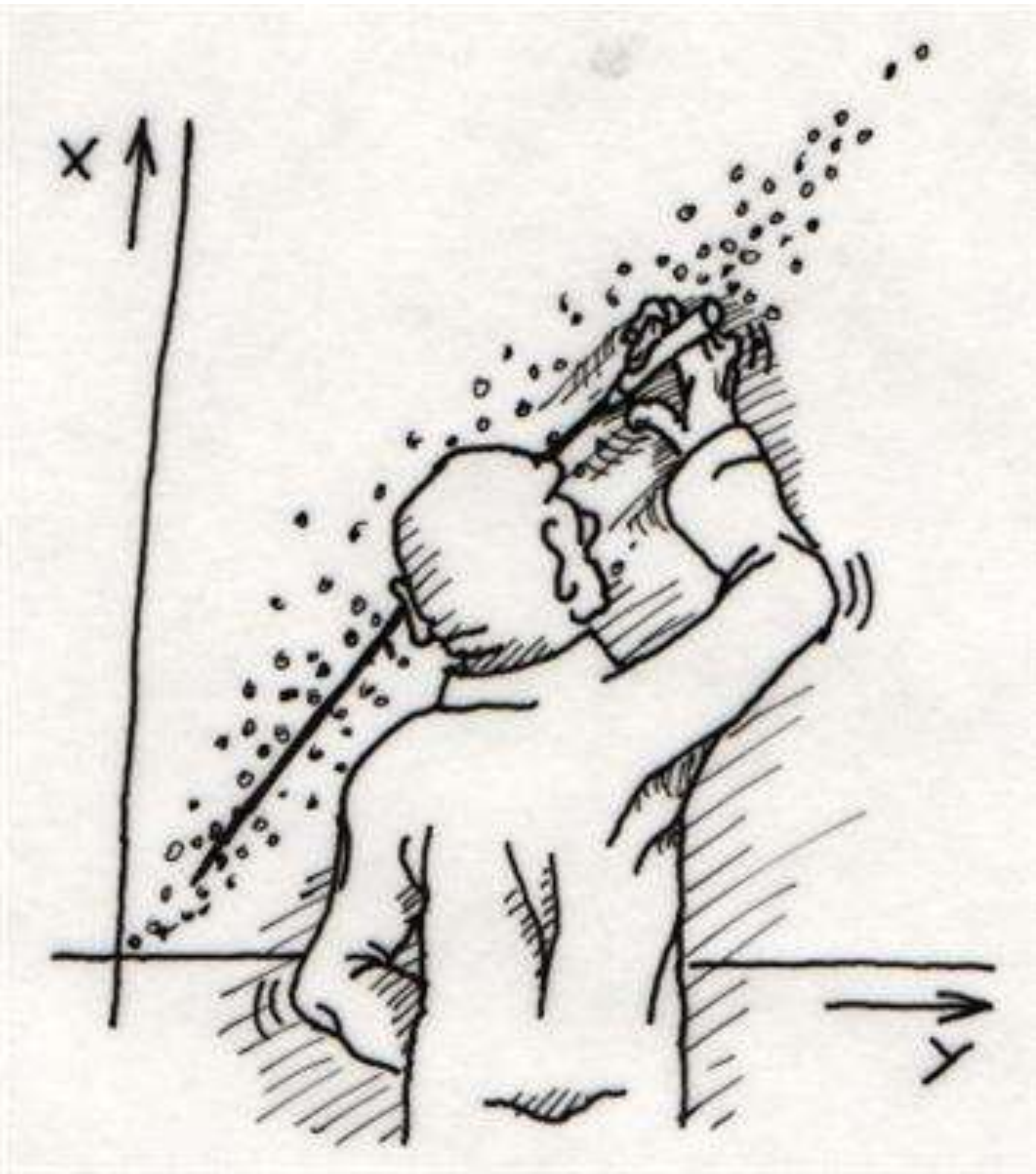


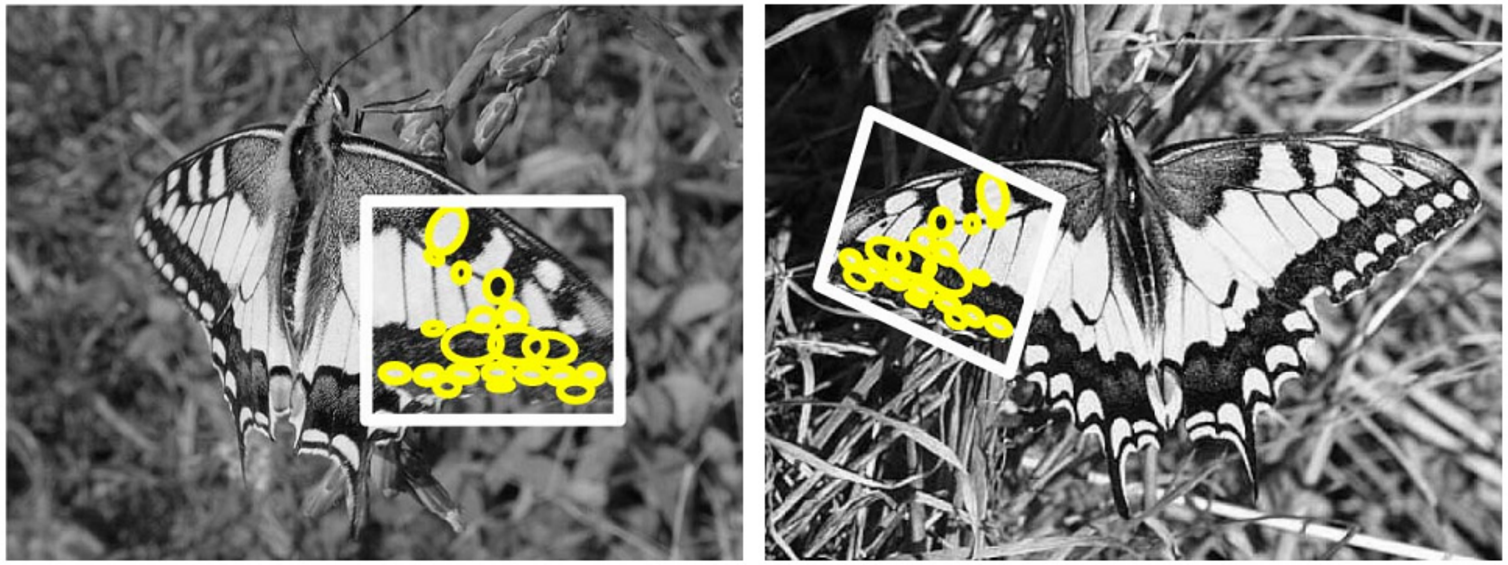
feature extraction, key-point detection

Source: L. Lazebnik

II. Mid-level vision

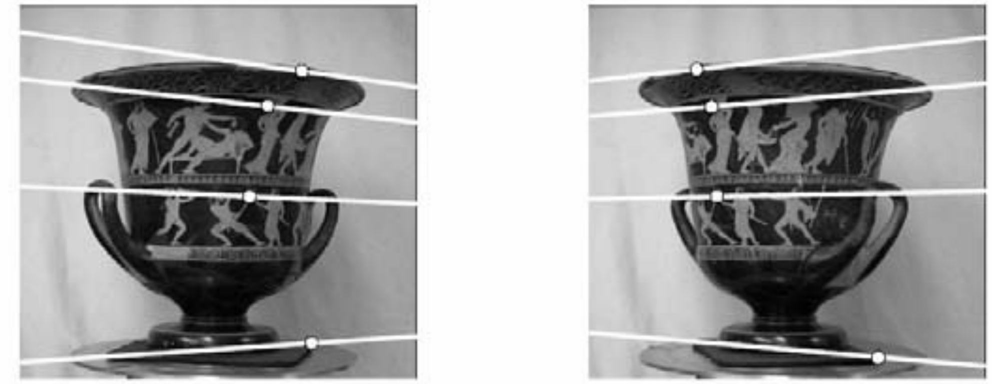
- Model fitting and grouping

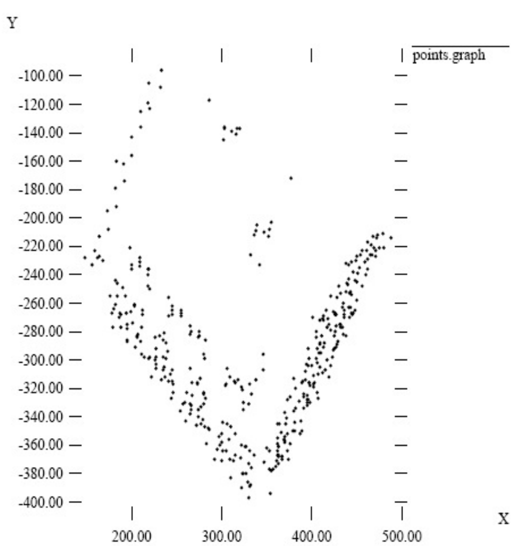


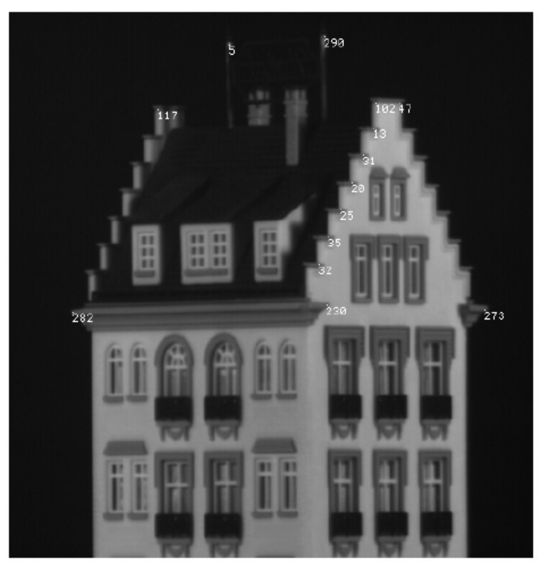


Source: L. Lazebnik

III. Multi-view geometry



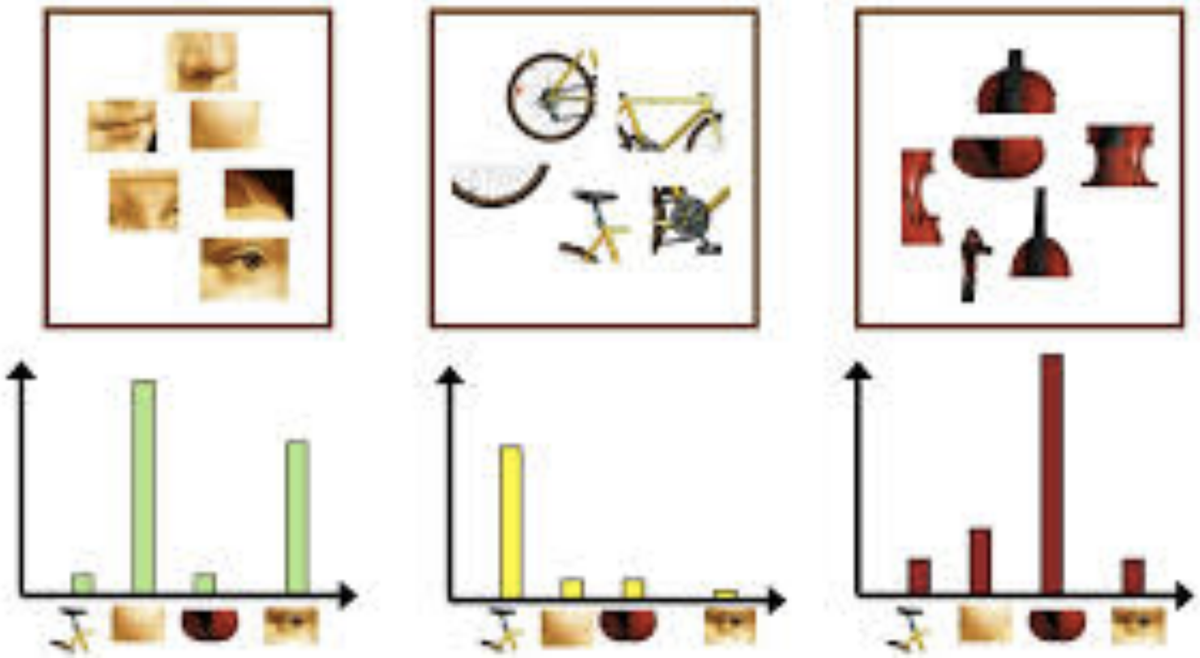




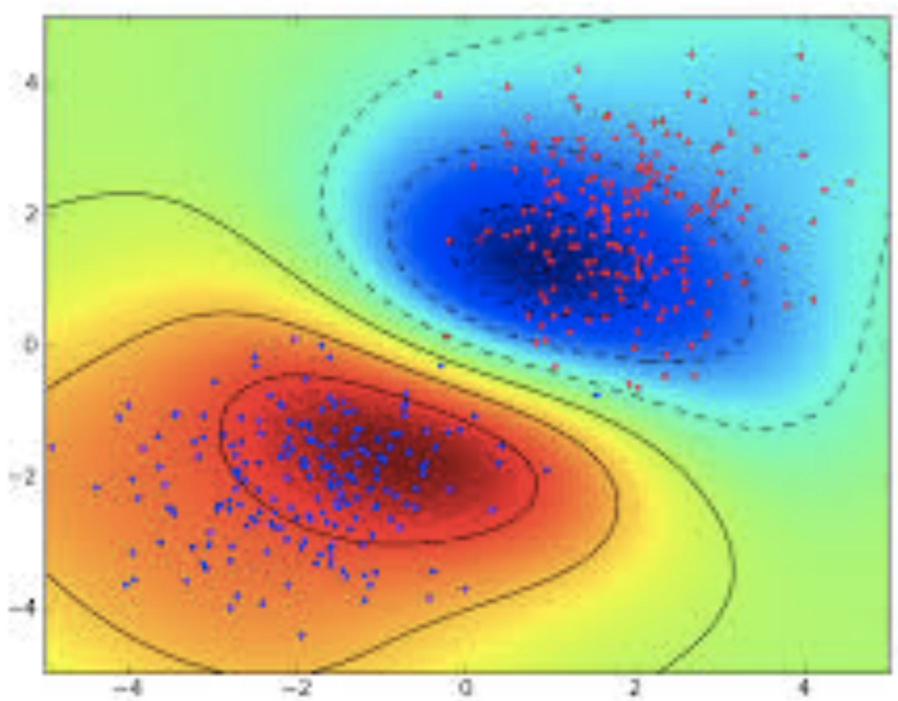
structure from motion

Драконъ, видимый подъ различными углами зрѣнія
По гравюрѣ на мѣди изъ „Oculus artificialis teledioptricus" Цана. 1702 года.

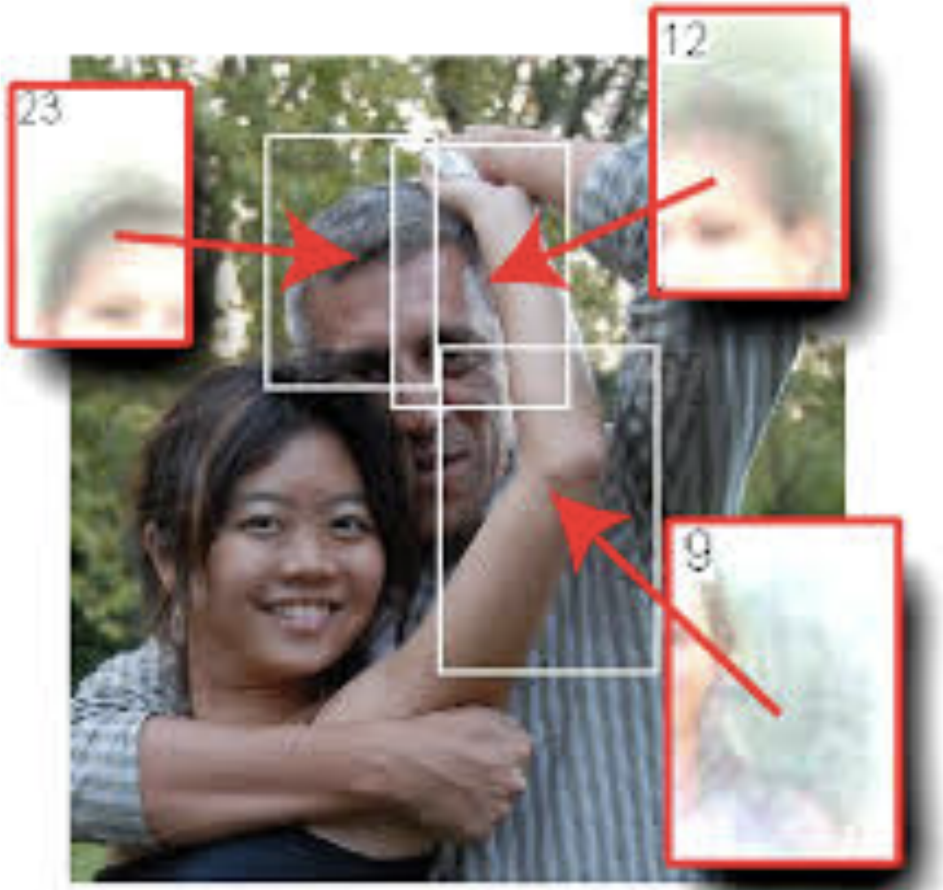
IV. Recognition



bag-of-word models

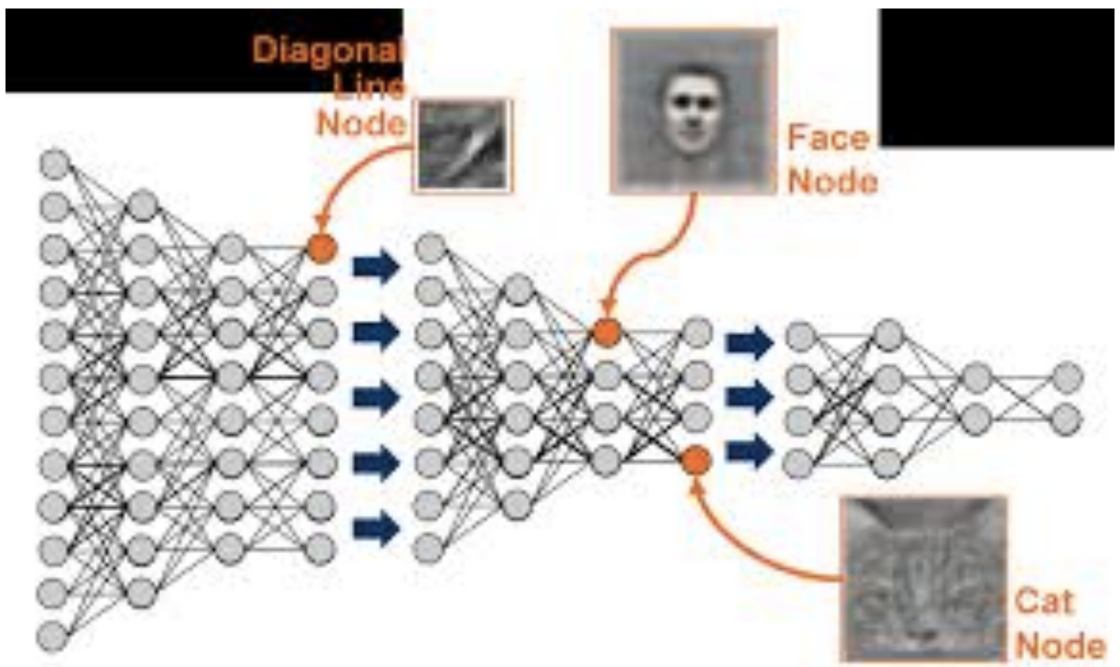


learning



part-based models

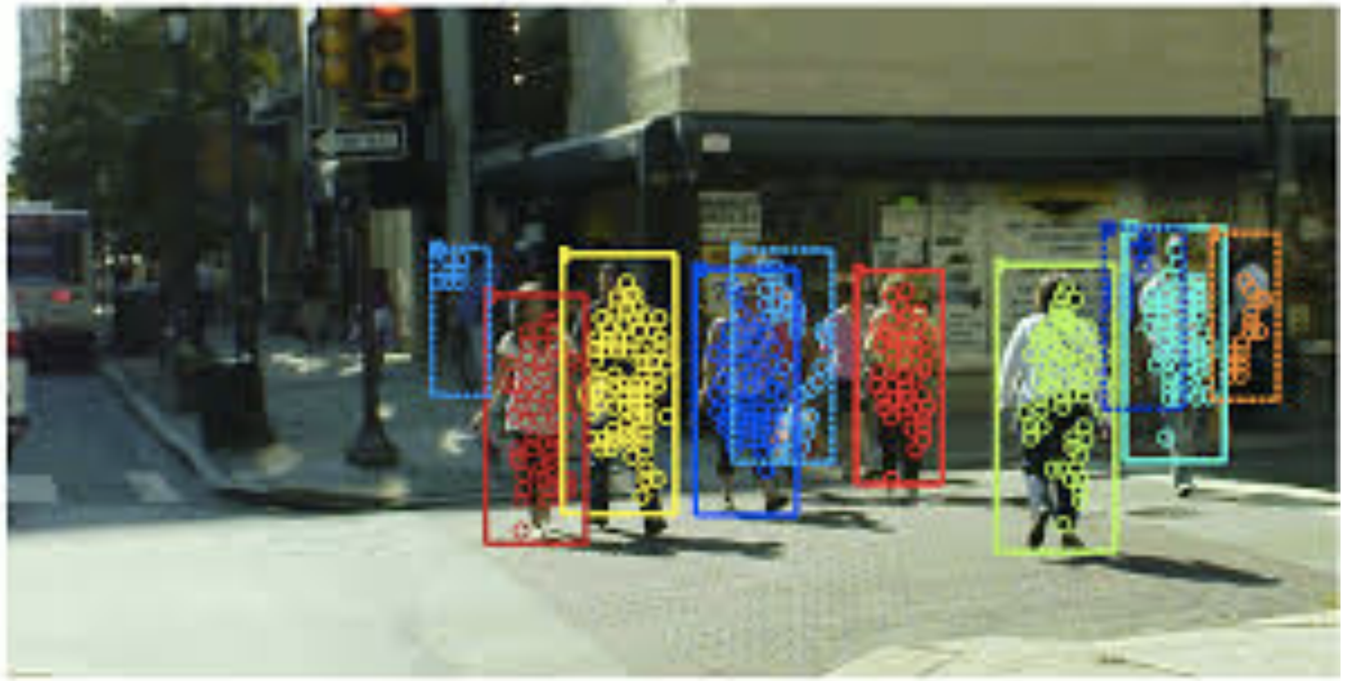
V. Additional topics



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$
  - **EdLab accounts** - each one of you that has registered for the course will get an account. But highly recommend getting a local version for figures.
- Readings:
  - **The speed of processing in the human visual system**, Thorpe et al., Letters to Nature, 1996
  - Chapter 1 in RS textbook