COS 429: Computer Vision

Slide credits: Jianxiong Xiao, Chris Bregler, Svetlana Lazebnik, Jitendra Malik

Image credit: kirkh.deviantart.com

COS 429: Computer Vision

Instructor: Szymon Rusinkiewicz

Guest lecturer: Andras Ferencz, MobilEye

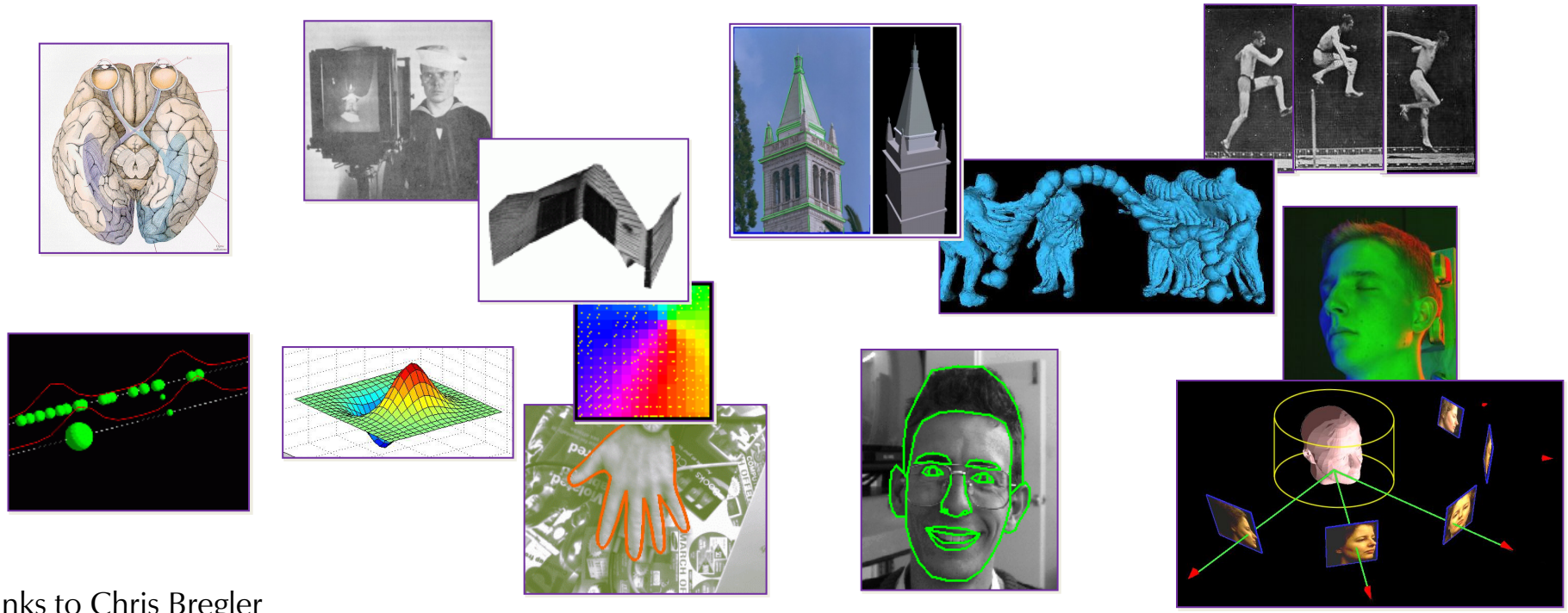
TAs: Karan Kathpalia, Riley Simmons-Elder, Andy Zeng

Web page:

`http://www.cs.princeton.edu/~cos429/`

What is Computer Vision?

- Input: images or video
- Output: description of the world
  - Also: measuring, classifying, interpreting visual information



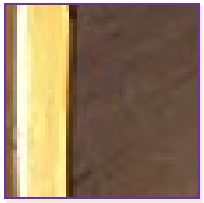
Thanks to Chris Bregler

One Pixel

- Amount of light recorded by a photoreceptor

"Is this the object's color?
Illumination? Noise?
I can't tell!"

Low-Level or “Early” Vision



- Local image/shape properties

“There’s an edge!”

Mid-Level Vision

"There's an object and a background!"

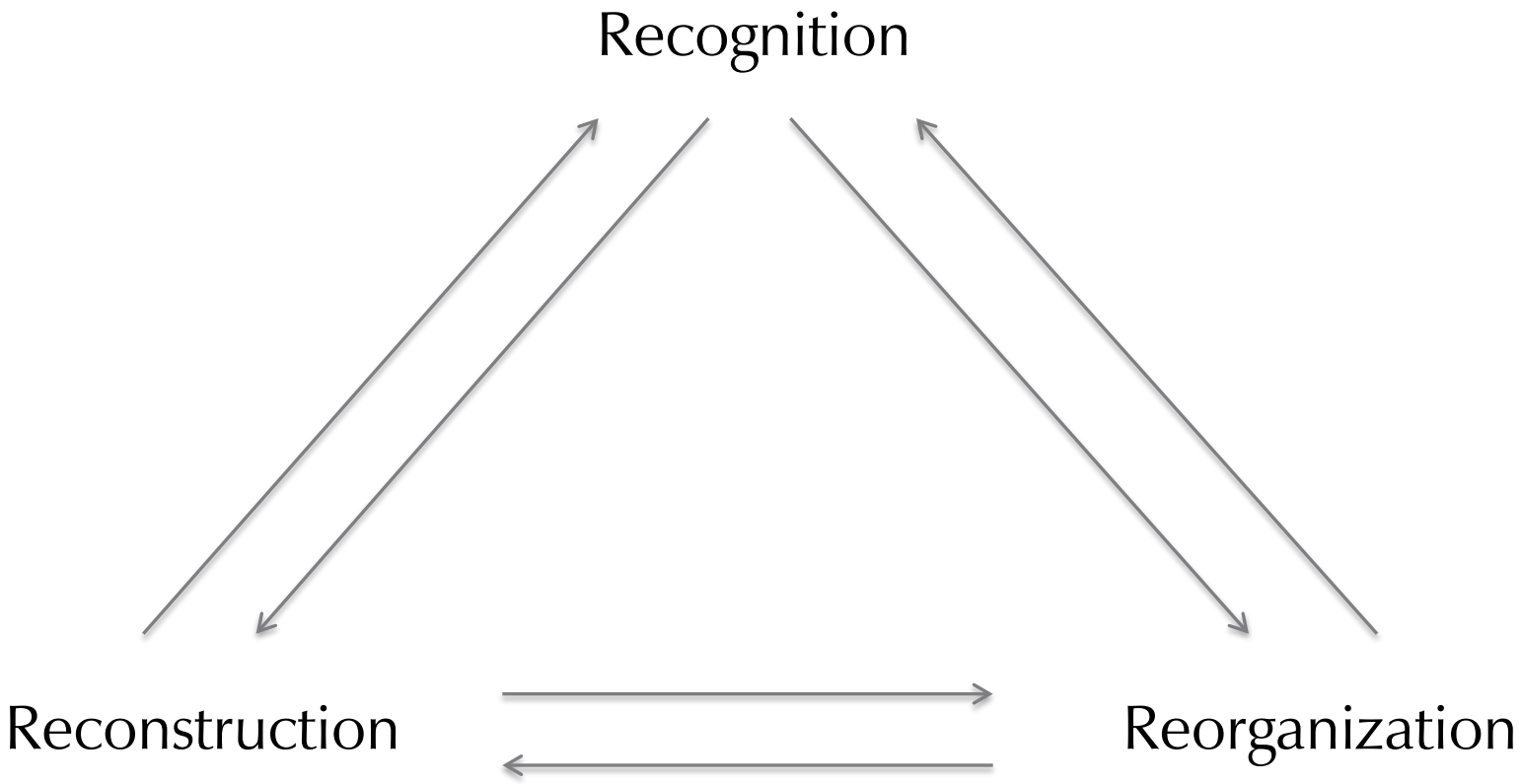
- Grouping and segmentation

High-Level Vision

"It's a chair! It's in a room!"

- Recognition
- Classification

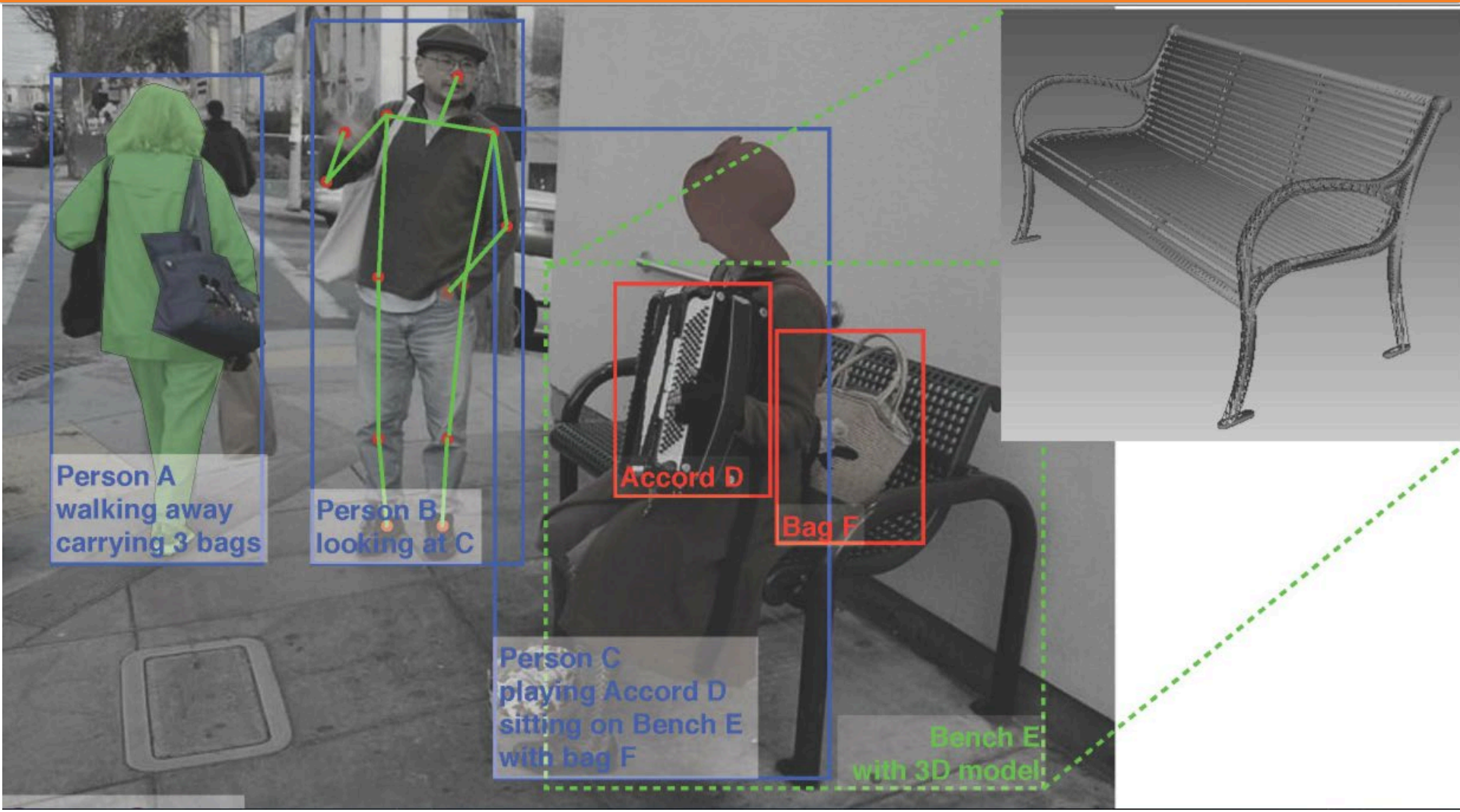
The Three "R"s of Vision



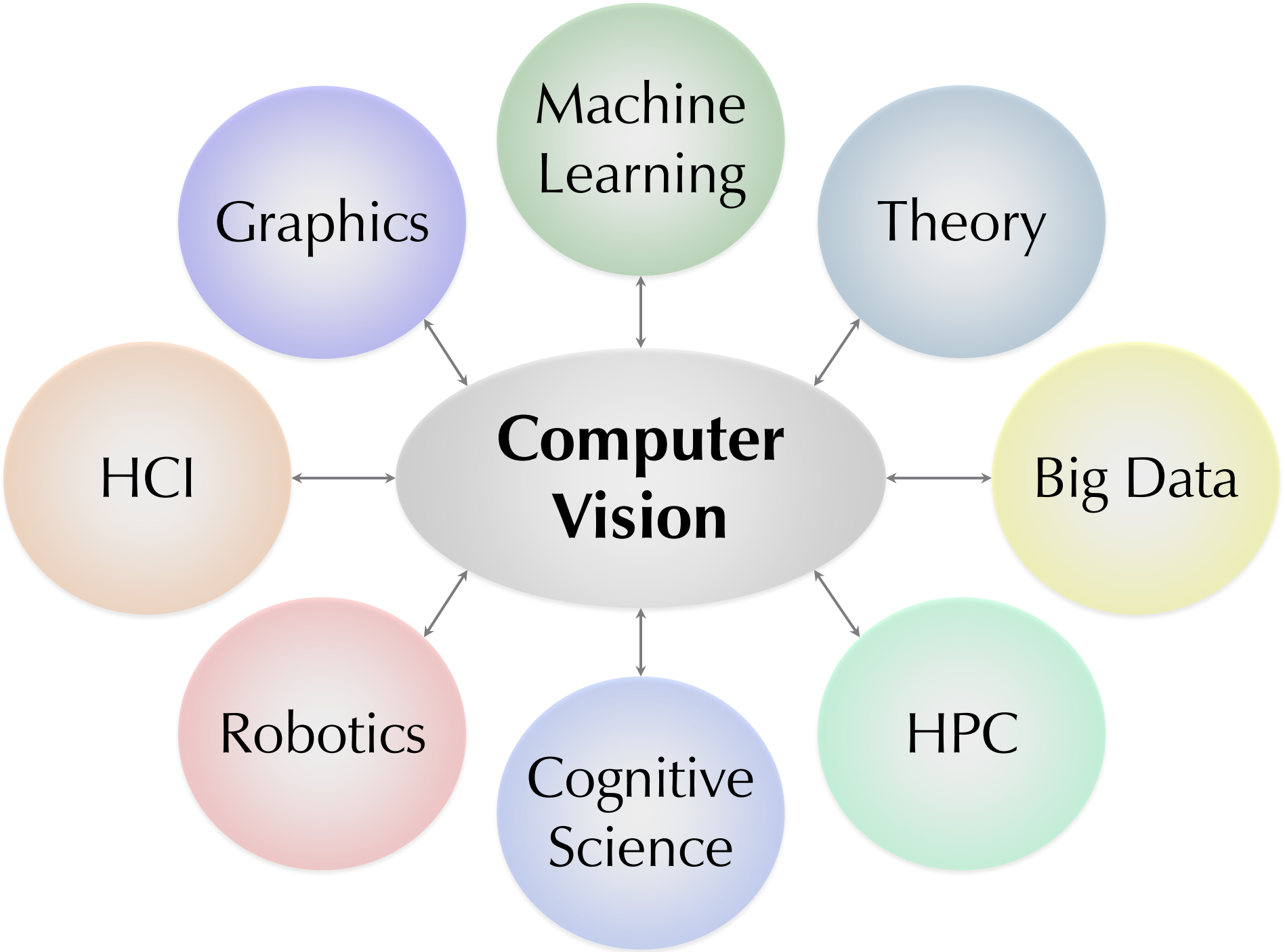
Each of the 6 directed arcs in this diagram is a useful direction of information flow

[Malik]

What We Would Like to Infer…

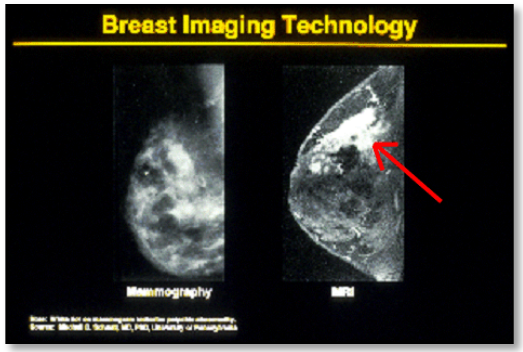


Will person B put some money into Person C's tip bag?


Machine Learning
Graphics
Theory
HCI
Computer Vision
Big Data
Robotics
Cognitive Science
HPC

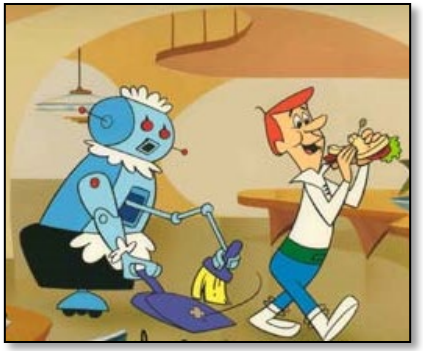
Why Does Computer Vision Matter?

Safety



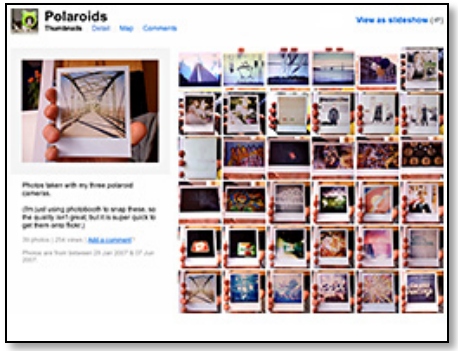
Health

Security



Comfort

Fun



Access

Consumer Applications

BBC NEWS

UK version | International version | About the versions

News Front Page
World
UK
England
Northern Ireland
Scotland
Wales
Business
Market Data
Your Money
E-Commerce
Economy
Companies
Politics
Health
Education

Last Updated: Monday, 6 February 2006, 14:29 GMT

E-mail this to a friend Printable version

Face-hunting cameras boost Nikon

Japanese camera maker Nikon has tripled its profits on the back of strong sales of digital cameras that automatically focus on human faces.

Face recognition cameras like the Coolpix L1 are popular

Object Recognition

[Google Goggles]

[Bing Vision]

Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com

3D Shape Capture for Special Effects

The Matrix movies, ESC Entertainment, XYZRGB, NRC

3D Maps

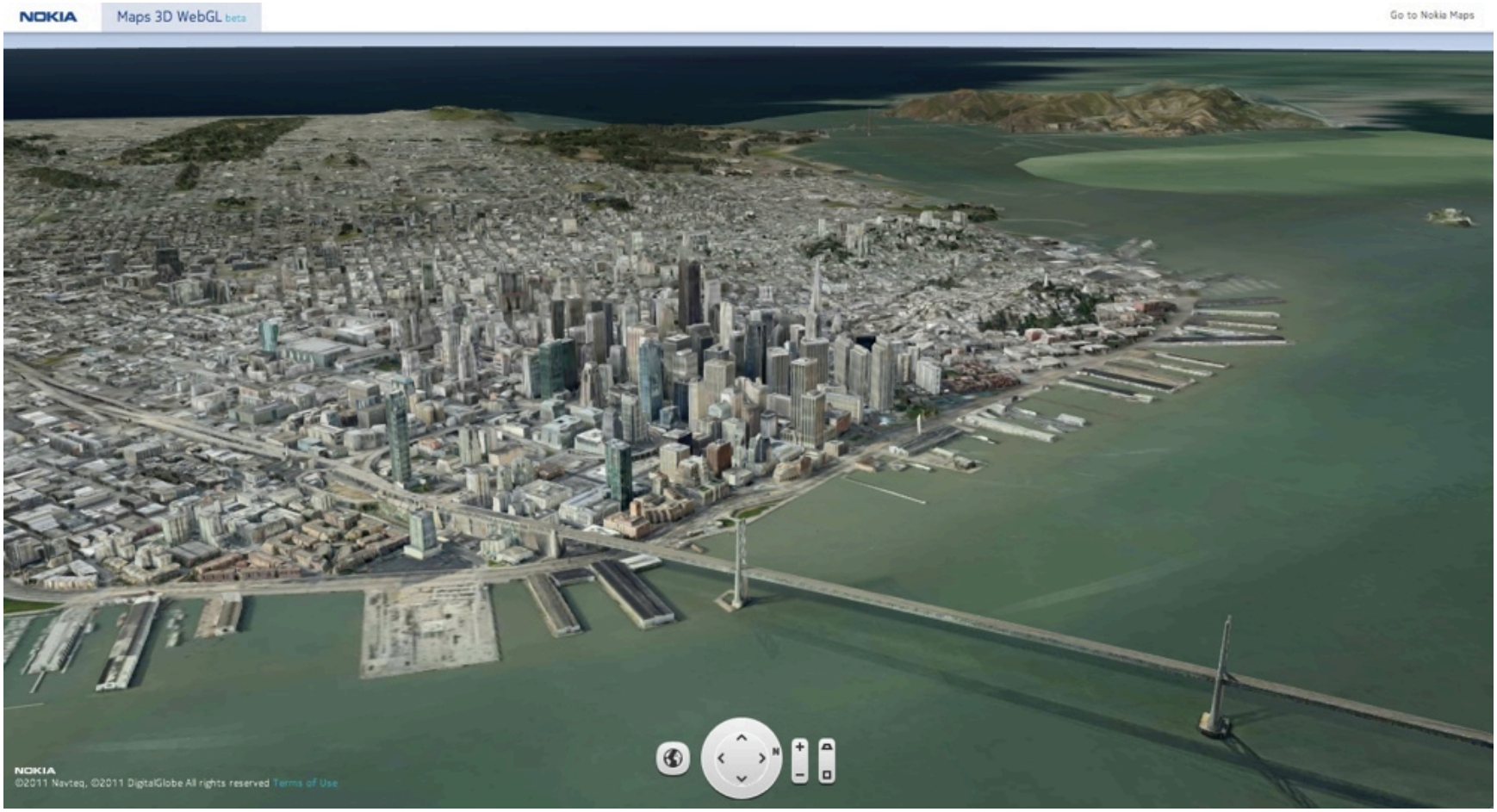
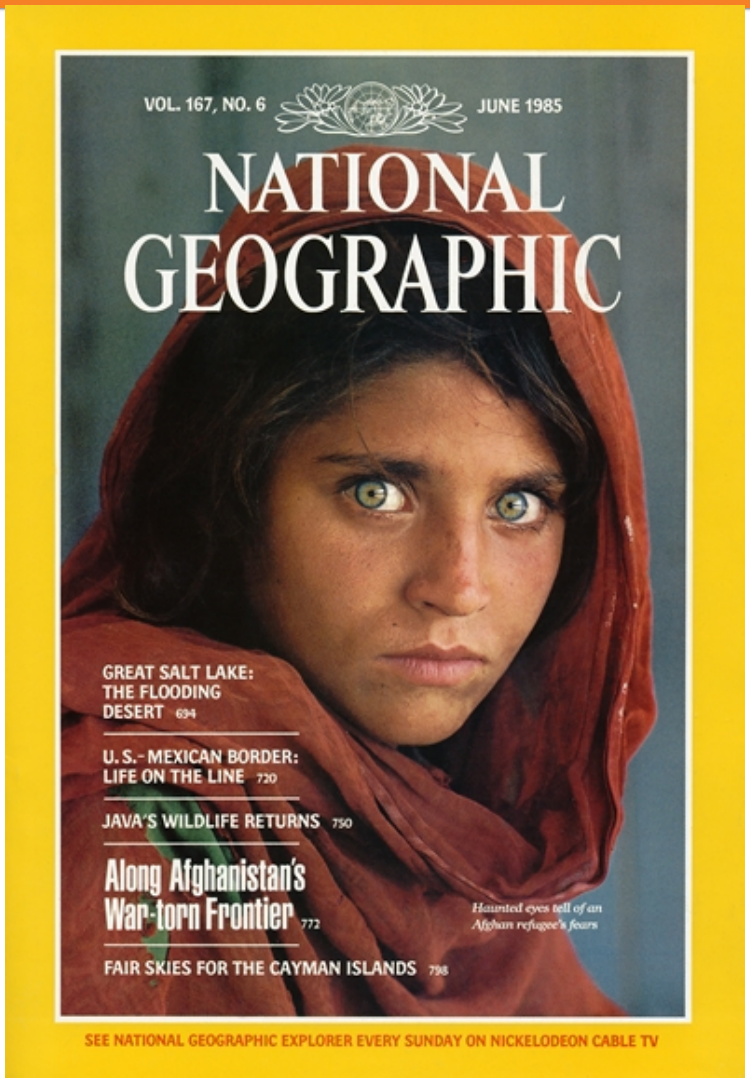


Image from Nokia's Maps 3D WebGL
(see also: Google Maps GL, Google Earth)

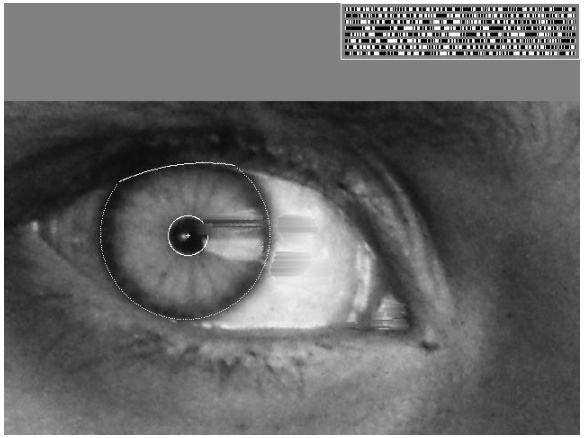
Face Recognition

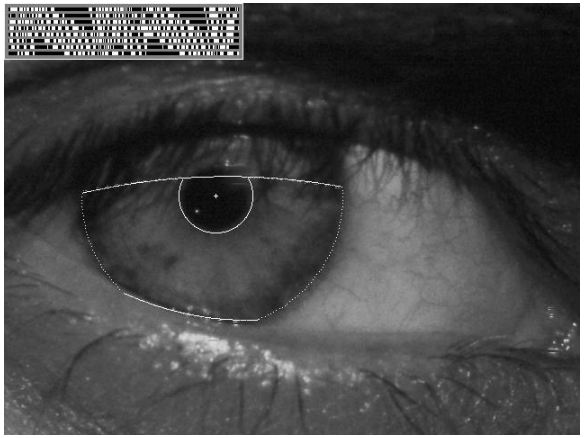


Who is she?

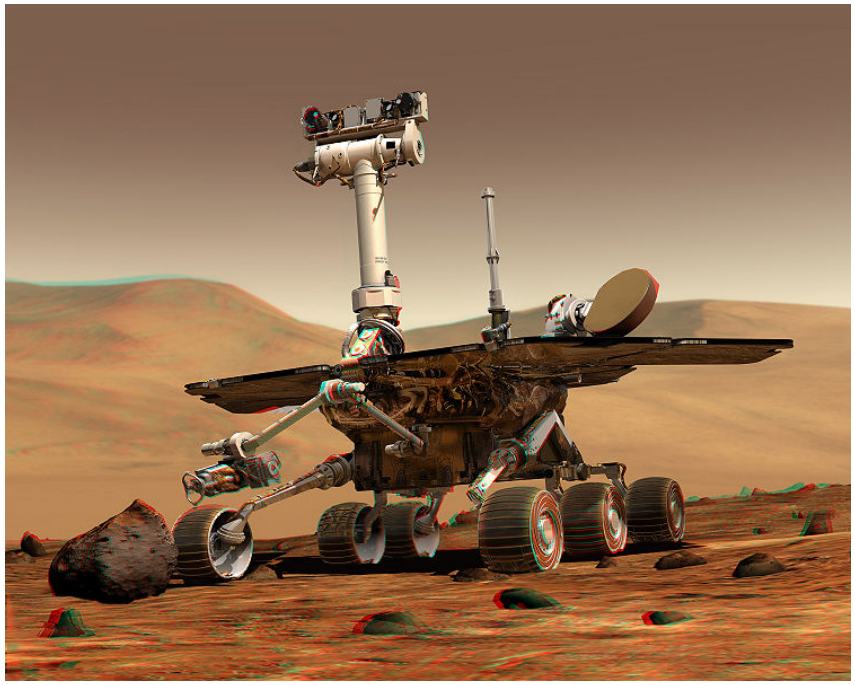
Vision-Based Biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story

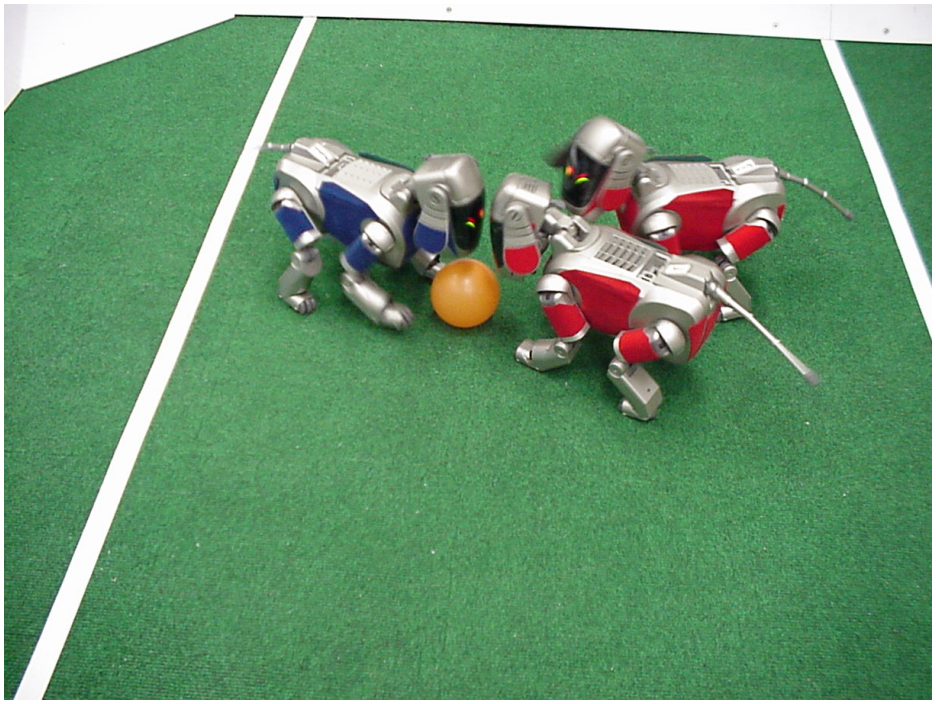




Robotics

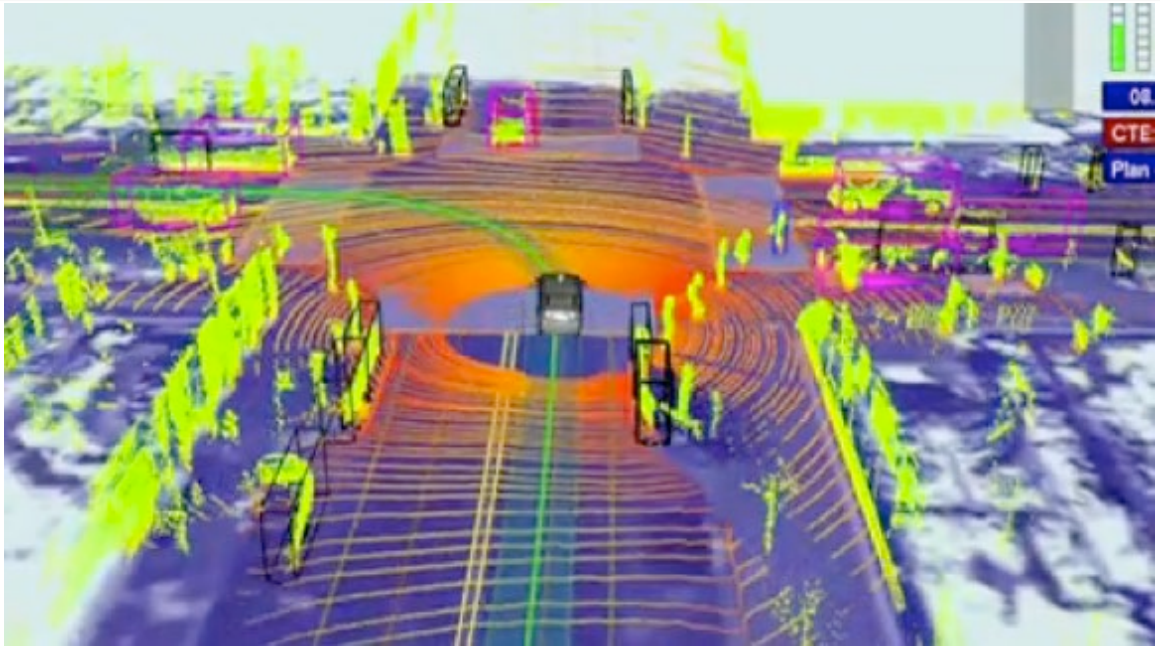


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover



http://www.robocup.org/

Self-Driving Cars

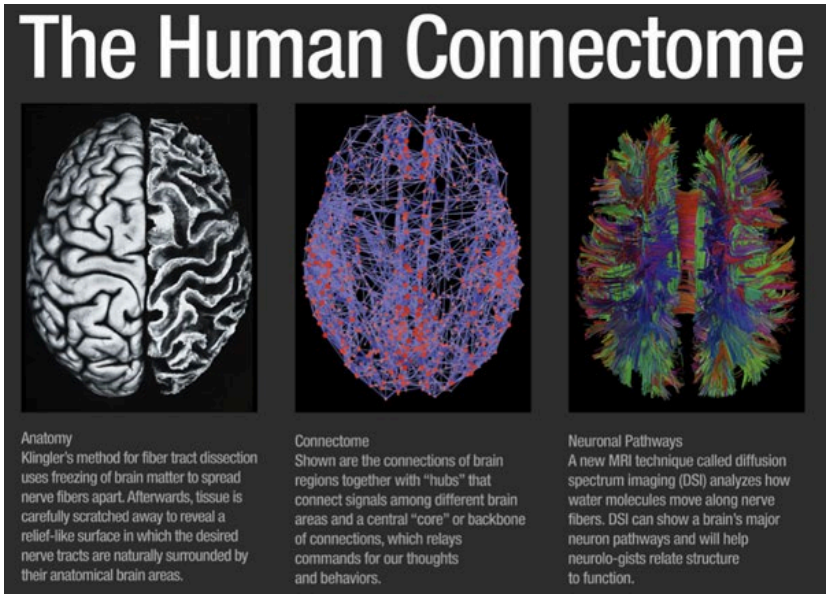


"Our self-driving cars have now traveled nearly 200,000 miles on public highways in California and Nevada, 100 percent safely. They have driven from San Francisco to Los Angeles and around Lake Tahoe, and have even descended crooked Lombard Street in San Francisco. They drive anywhere a car can legally drive."

- Sebastian Thrun, Google

Bad News: Vision is Really, Really Hard

- Vision is an amazing feat of natural intelligence
  - Visual cortex occupies about 50% of Macaque brain
  - More human brain devoted to vision than anything else



Prof. Sebastian Seung

Challenge: Viewpoint Variation

Challenge: Illumination

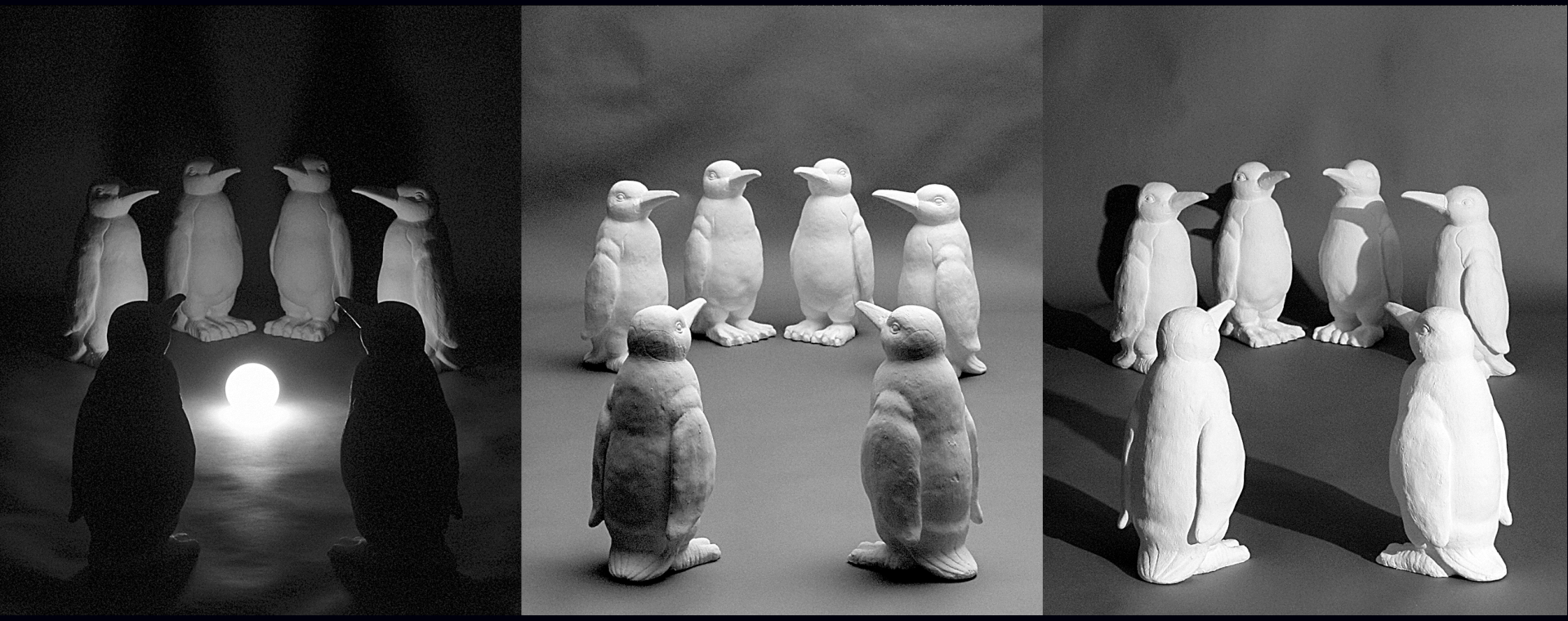
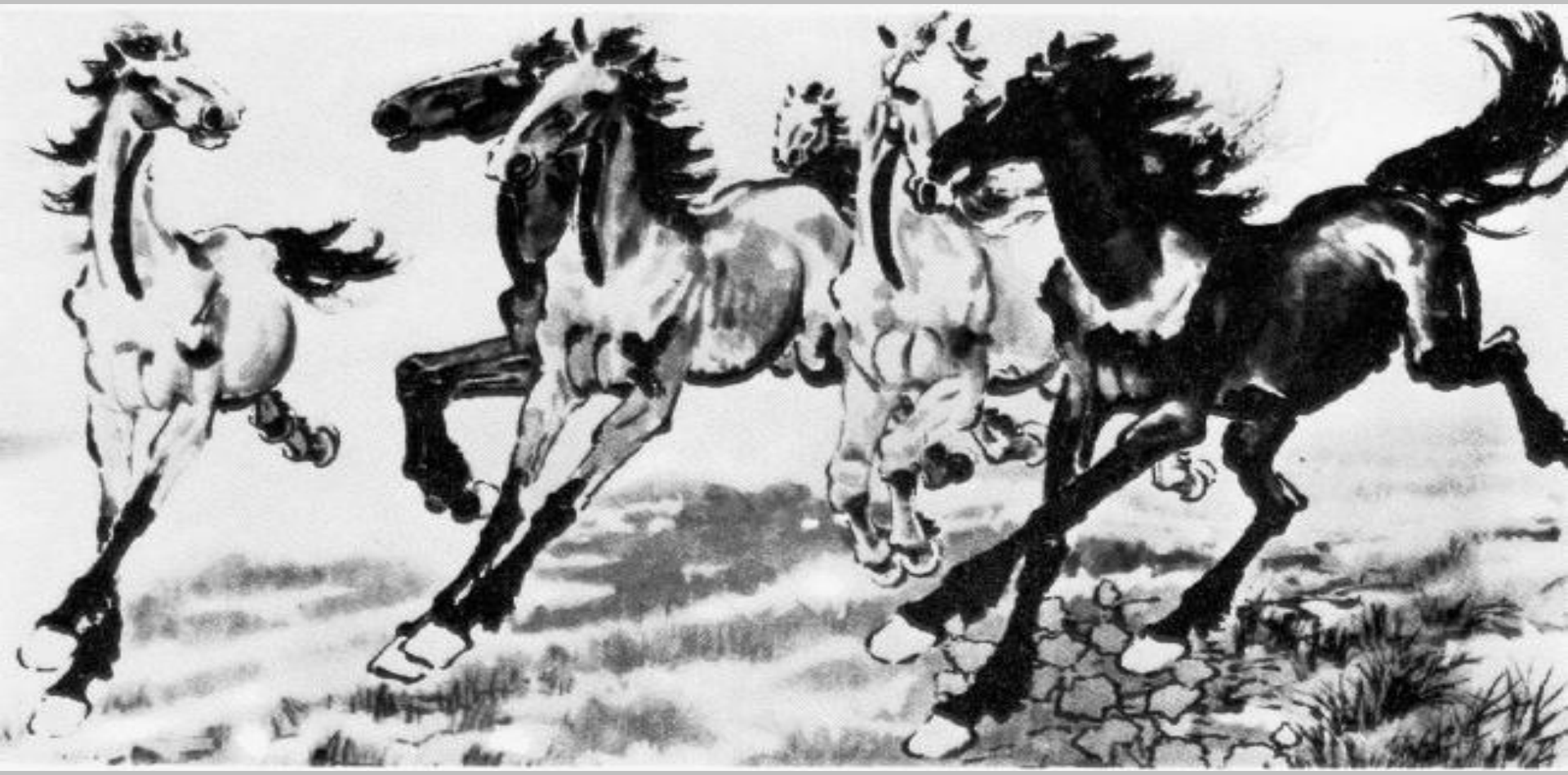


image credit: J. Koenderink

Challenge: Scale

slide credit: Fei-Fei, Fergus & Torralba

Challenge: Deformation



Xu, Beihong 1943

slide credit: Fei-Fei, Fergus & Torralba

Challenge: Intra-Class Variation

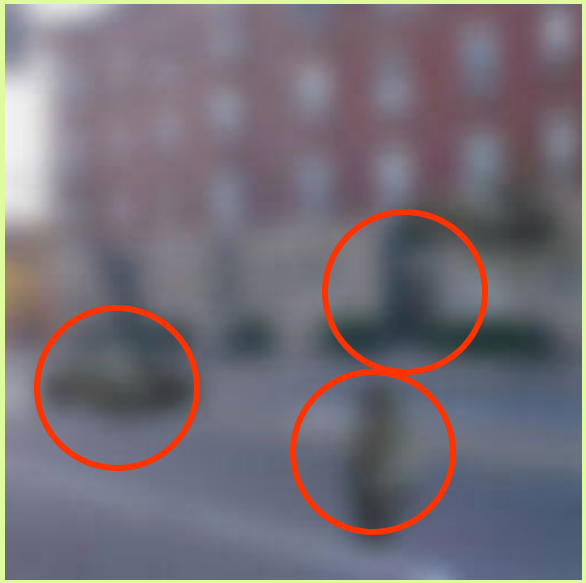
slide credit: Fei-Fei, Fergus & Torralba

Challenge: Occlusion, Clutter

Image source: National Geographic

Challenge: Motion

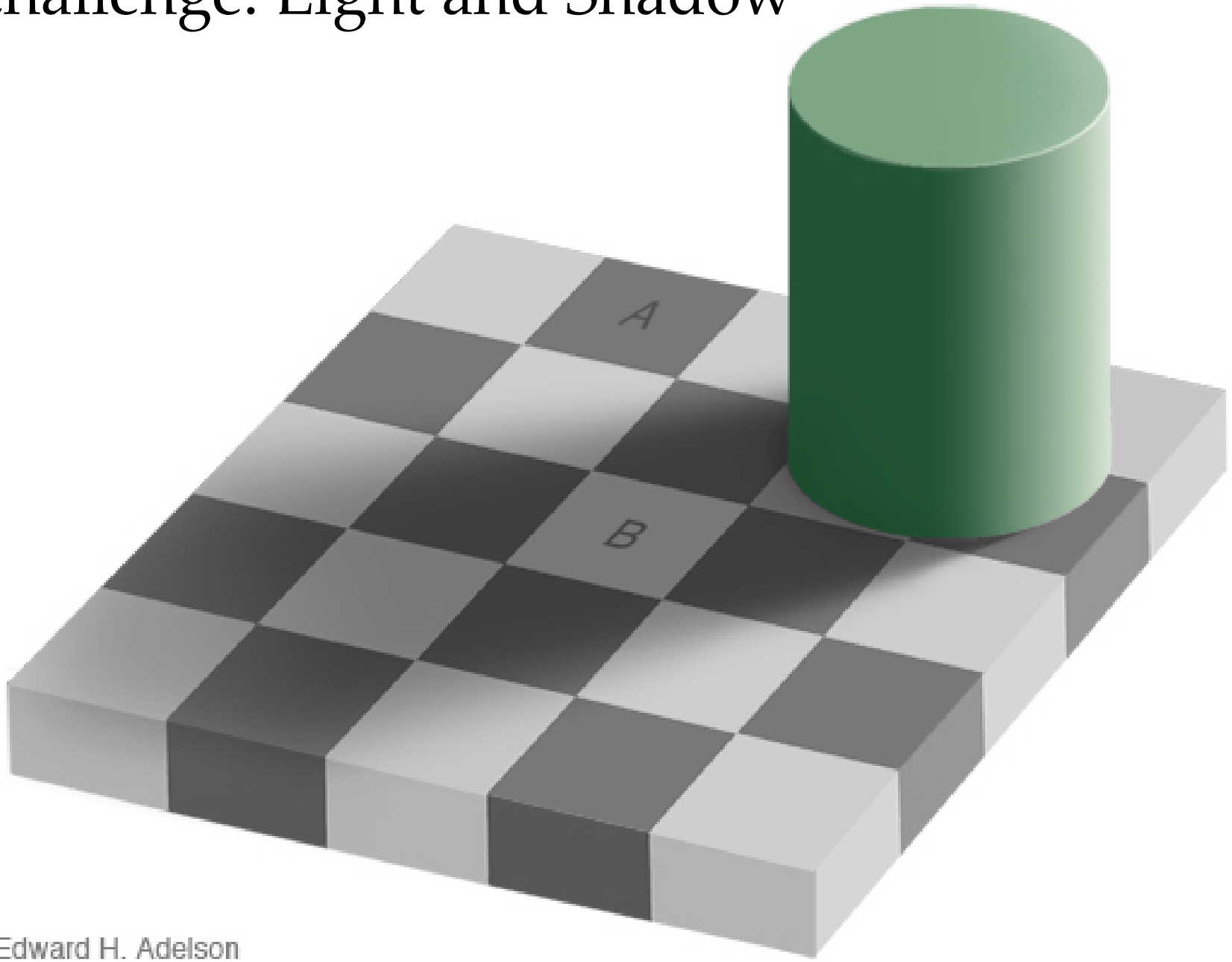
Challenge: Ambiguity



slide credit: Fei-Fei, Fergus & Torralba

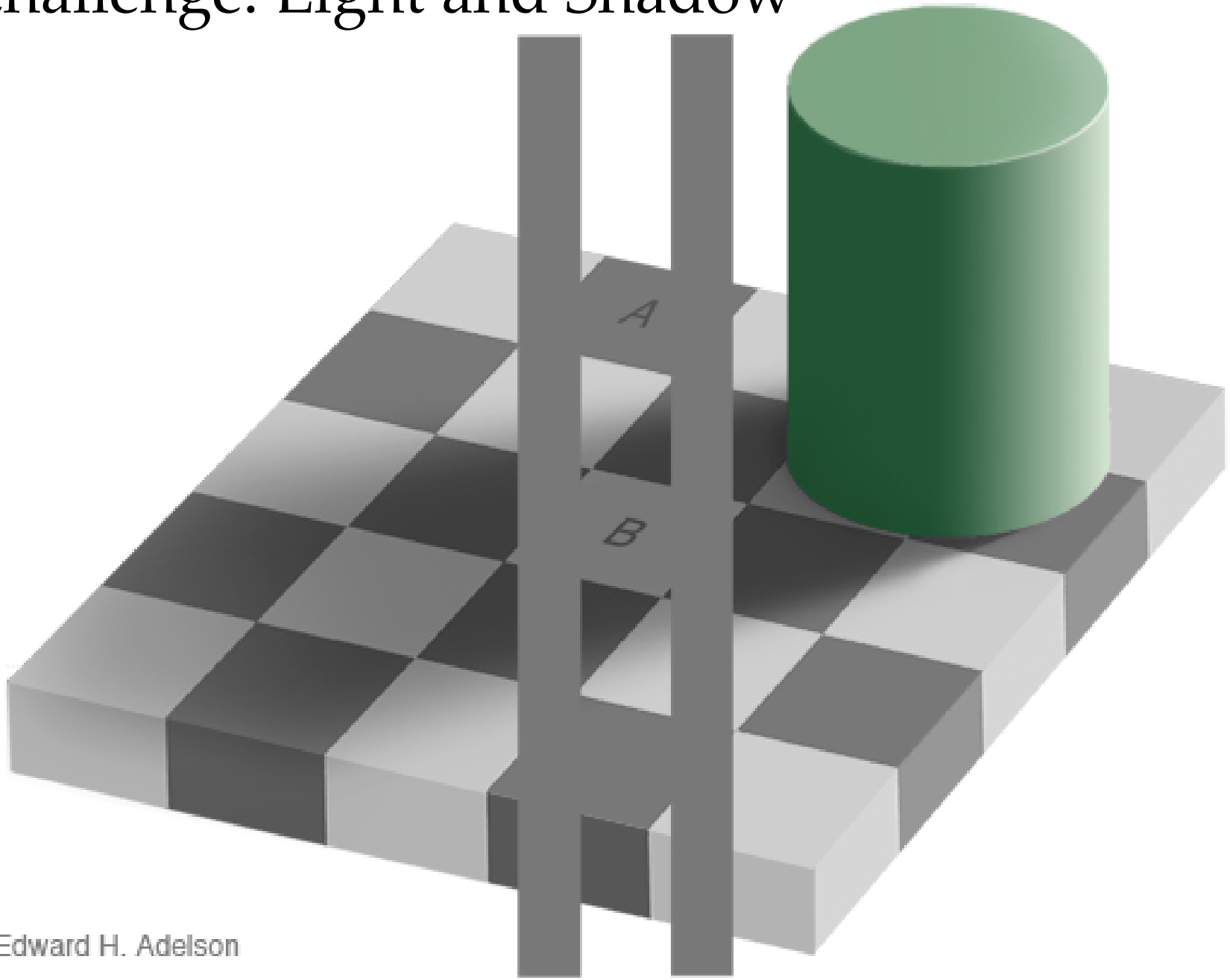
Challenge: Perspective

Challenge: Light and Shadow



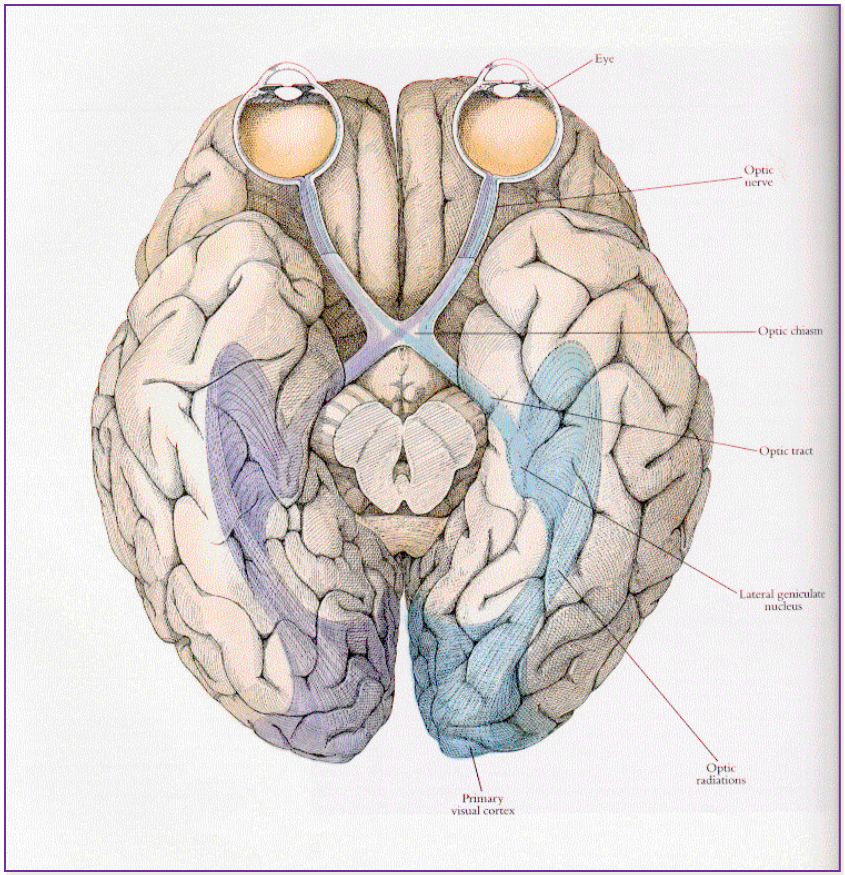
Edward H. Adelson

Challenge: Light and Shadow

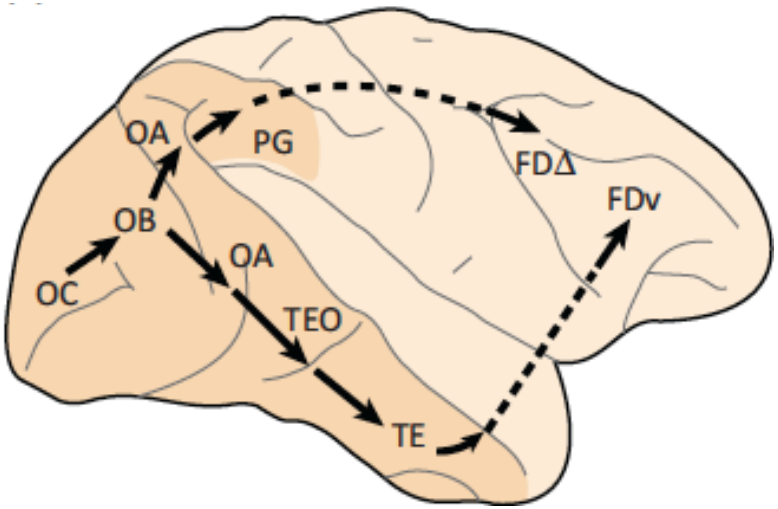


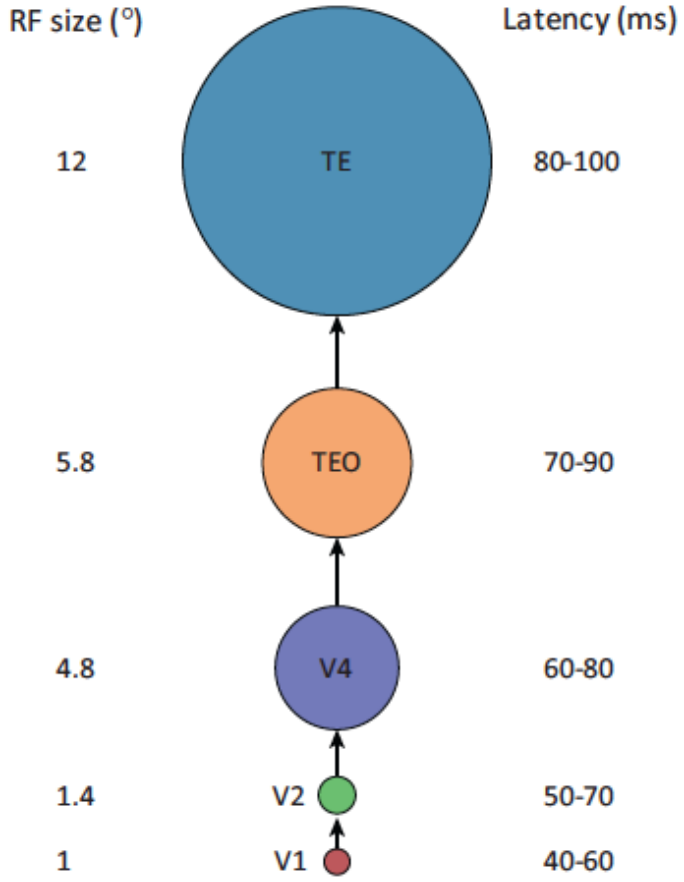
Human Visual System

- Lens, photoreceptors
- Retinal ganglion cells
- Lateral Geniculate Nucleus – visual adaptation?
- Primary Visual Cortex
  - Simple cells: orientational sensitivity
  - Complex cells: directional sensitivity
- Further processing
  - Temporal cortex: what is the object?
  - Parietal cortex: where is the object? How do I get it?



Feed-Forward Model of the Ventral Stream

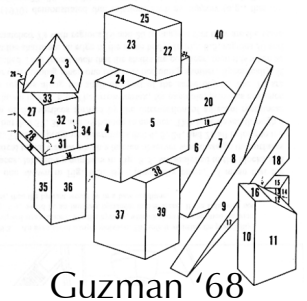




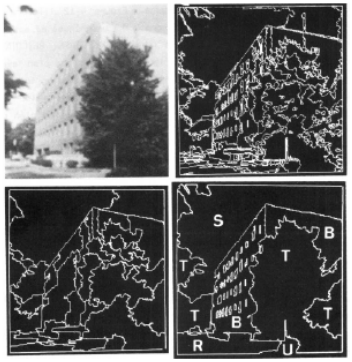
Kravitz et al, Trends in Cognitive Science 2013

A Brief History of Computer Vision

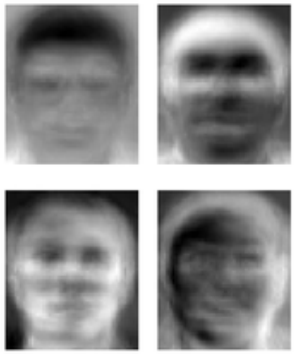
- 1966: Marvin Minsky assigns computer vision as an undergrad summer project
- 1960s: interpretation of synthetic worlds
- 1970s: interpretation of carefully selected images
- 1980s: NNs come and go; shift towards geometry and increased mathematical rigor
- 1990s: face recognition; statistical analysis
- 2000s: broader recognition; large annotated datasets available; video processing
- 2030s: robot uprising?



Guzman '68



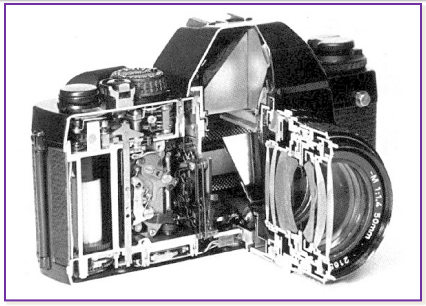
Ohta Kanade '78

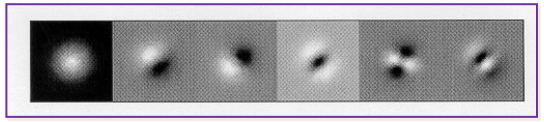


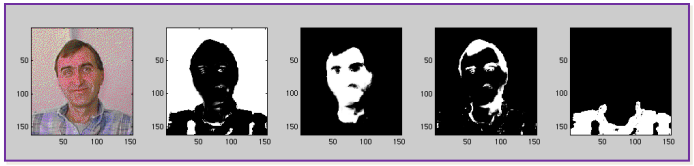
Turk and Pentland '91

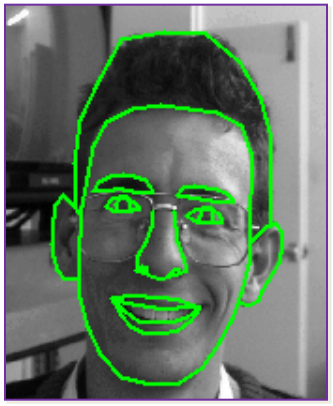
Computer Vision: Implementing the Stages

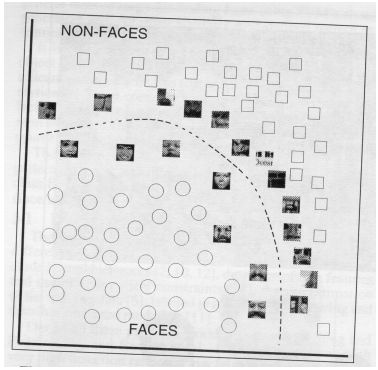
- Lens, image sensors
- Filter banks
- Shape from stereo, etc.
- Clustering, segmentation
- Object detection
- Classification

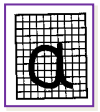




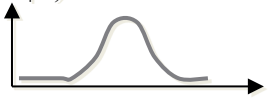




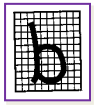




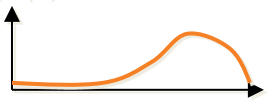
$P(X \mid a)$



black pixels

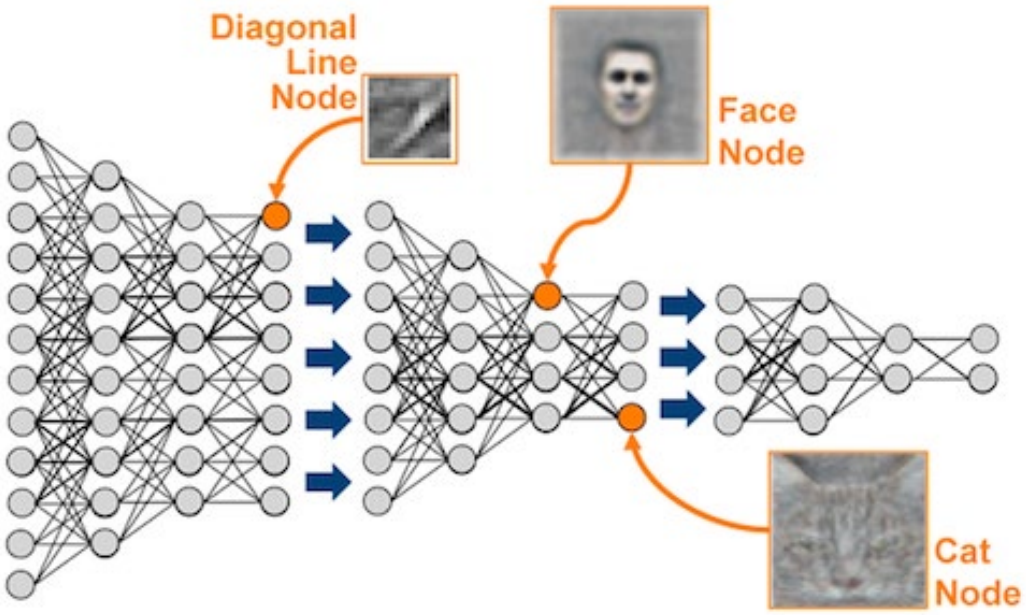


$P(X \mid b)$



black pixels

Deep Learning: End-to-End Optimization



Course Outline

- Image formation and capture
- Filtering and feature detection
- Segmentation and clustering
- Recognition and classification
- Motion estimation and tracking
- 3D shape reconstruction
- Convolutional neural nets / deep learning

Course Mechanics

- Recommended book:

Computer Vision: Algorithms and Applications

© 2010 Richard Szeliski, Microsoft Research

  - Also available online: http://szeliski.org/Book/

- Assigned papers / other readings

Course Mechanics

- 70%: 4 written / programming assignments
  - Individual: all submitted work must be your own
  - 3 free late days
- 30%: Final project
  - Small groups – 2-3 people
  - Presentation / demo in January
  - Writeup due on Dean's date

Q&A

- We will use piazza for Q&A. Please direct all non-private questions there.

- Feel free to answer each others' questions (we will monitor and endorse students' answers) but keep in mind collaboration policy

MATLAB

- The assignments use the MATLAB language.
- Easy to learn for most students – first assignment will walk you through the basics.

- School of Engineering is running a short course, Monday Sep 19 – Tuesday Sep 20
- Registration required: https://fs8.formsite.com/kellercenter/matlabR2016/index.html

COS 429: Computer Vision

Instructor: Szymon Rusinkiewicz

Guest lecturer: Andras Ferencz, MobilEye

TAs: Karan Kathpalia, Riley Simmons-Elder, Andy Zeng

Web page:

`http://www.cs.princeton.edu/~cos429/`