CMSC 733: Computer Processing of Pictorial Information (Computer Vision)

David Jacobs

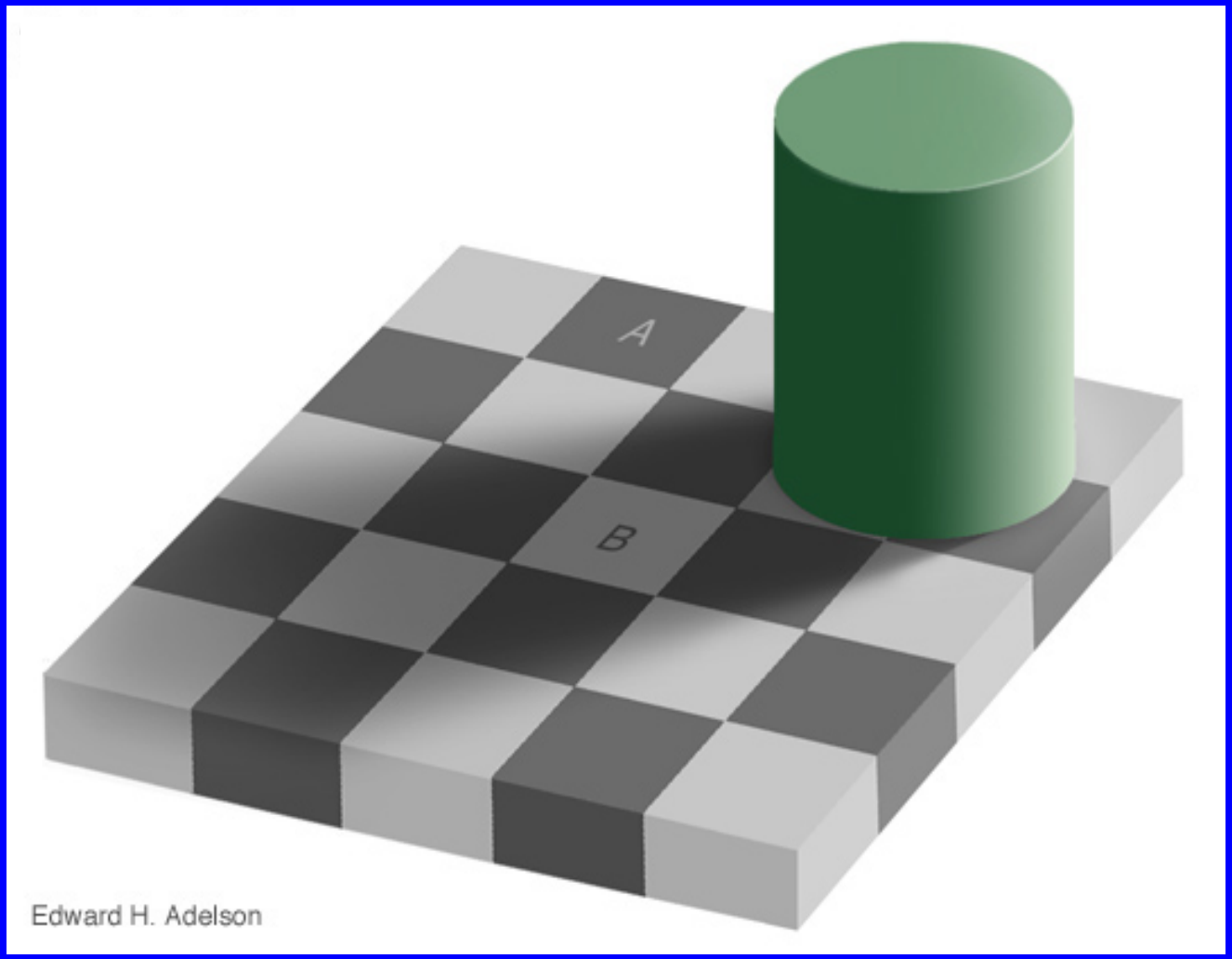
Vision

- ``to know what is where, by looking.'' (Marr).
- Where
- What

Why is Vision Interesting?

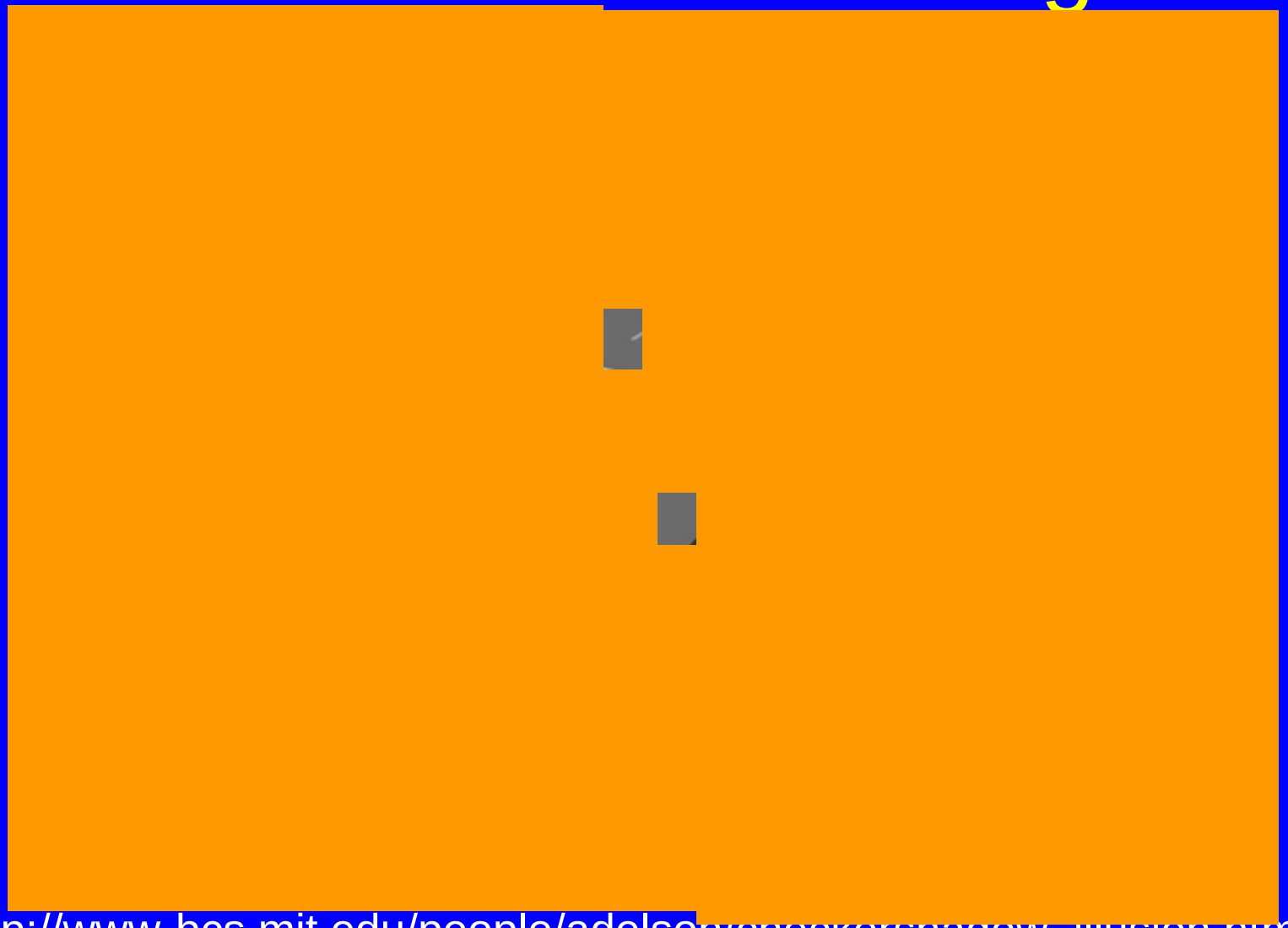
- Psychology
  - ~ 50% of cerebral cortex is for vision.
  - Vision is how we experience the world.
- Engineering
  - Want machines to interact with world.
  - Digital images are everywhere.

Vision is inferential: Light



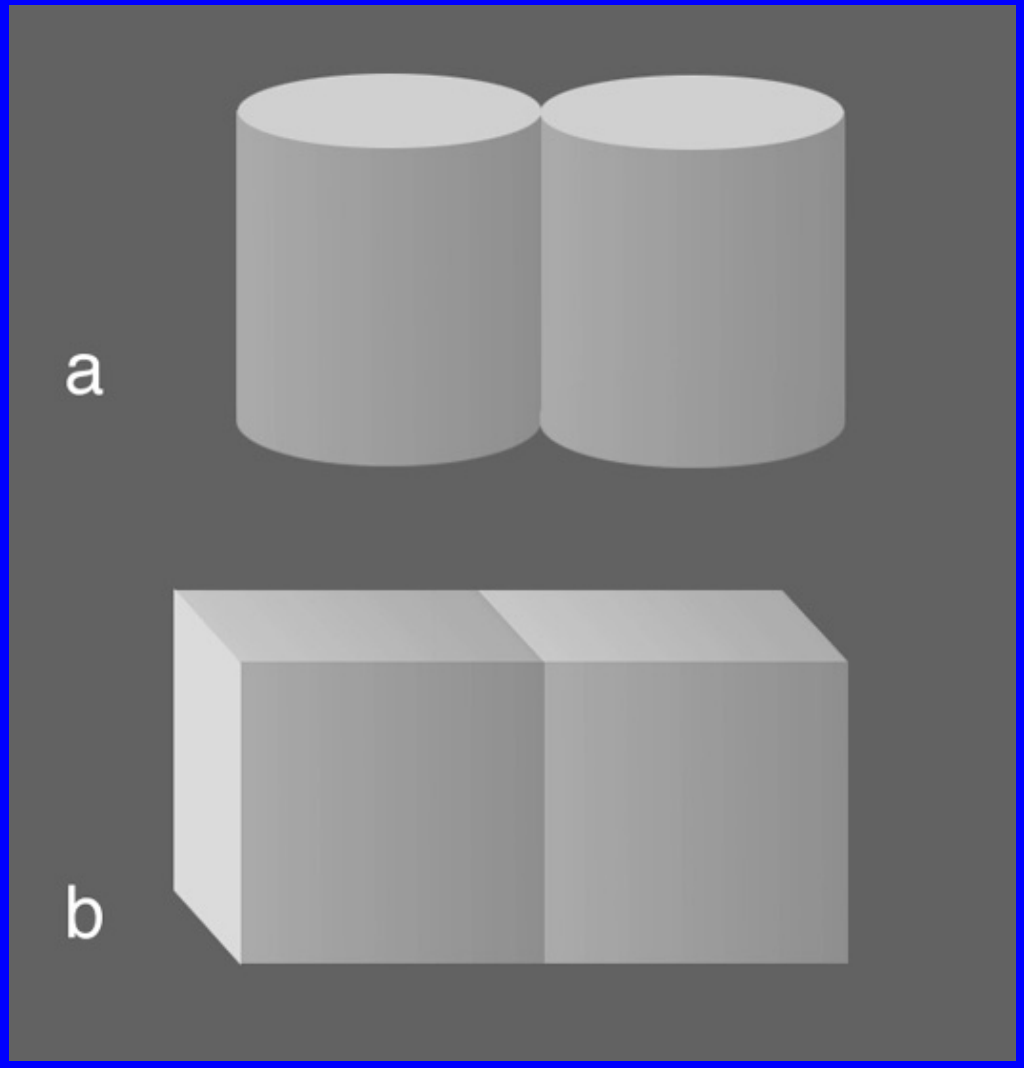
(http://www-bcs.mit.edu/people/adelson/checkershadow_illusion.html)

Vision is inferential: Light



(http://www-bcs.mit.edu/people/adelson/checkershadow_illusion.html)

Vision is Inferential: Geometry



Vision is Inferential: Prior Knowledge

Vision is Inferential: Prior Knowledge

Computer Vision

- Inference → Computation
- Building machines that see
- Applying computation to images
- Modeling biological perception

Goals of this course

- Prepare students for research in CV
- Breadth and depth
  - Breadth: familiarity with (almost) all areas of computer vision.
    - We will cover a lot of ground.
  - Depth: Intuitive understanding of fundamental principles of vision.

Breadth

- Fourier transforms, convolution
- Image denoising – linear
- Image denoising – nonlinear

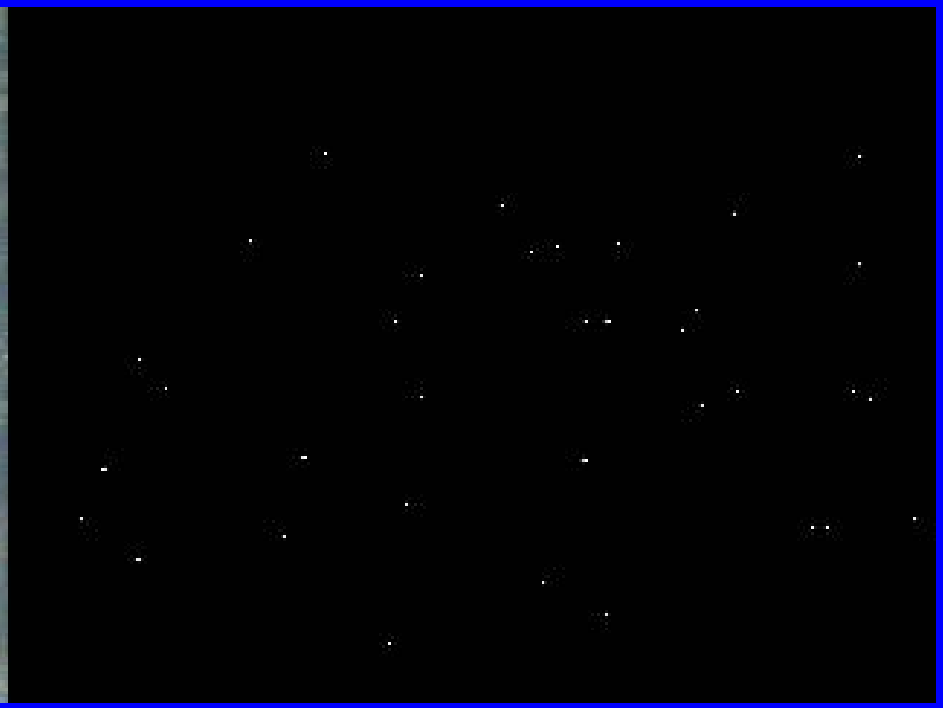
Topics (cont'd)
Edge detection

Canny edge detector (from Wikipedia)

Image Segmentation

GrabCut: Interactive Foreground Extraction using Iterated Graph Cuts

Topics (cont'd) Background Subtraction

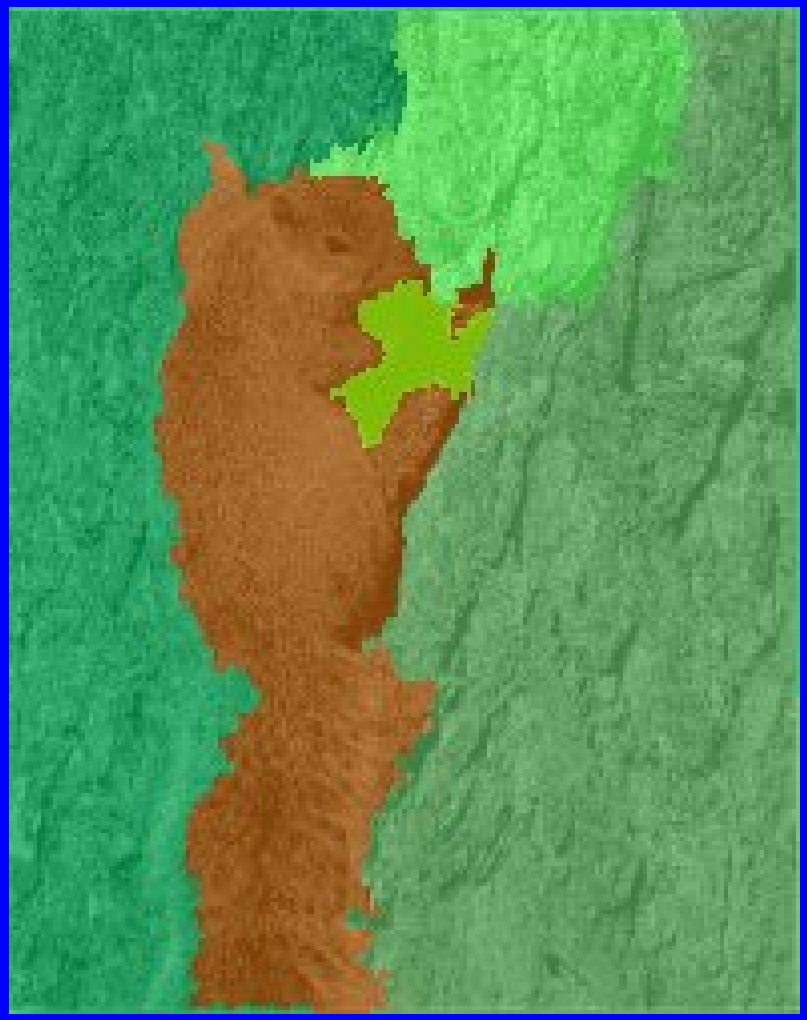


Topics (cont'd)

- Statistical modeling of images
  - Mixture of Gaussians, E-M
  - Markov processes, Markov random fields

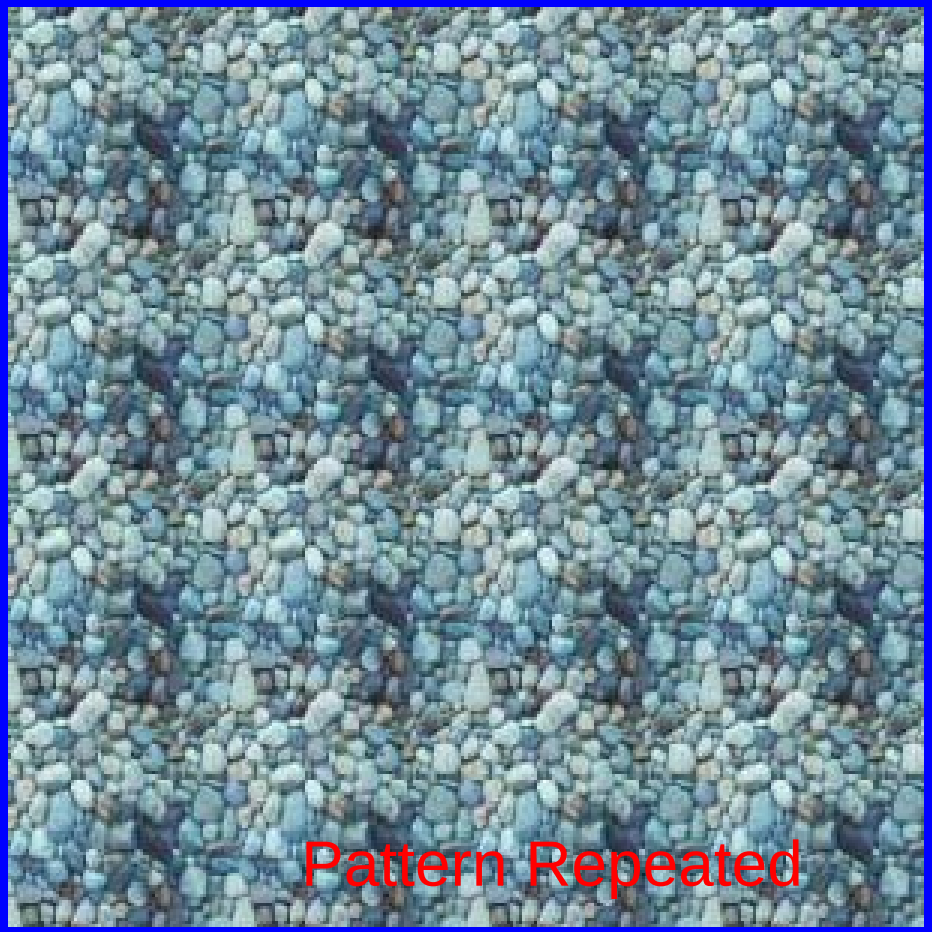
Texture in Boundary Detection

(Sharon, Balun, Brandt, Basri)



Texture Synthesis

Photo



Texture

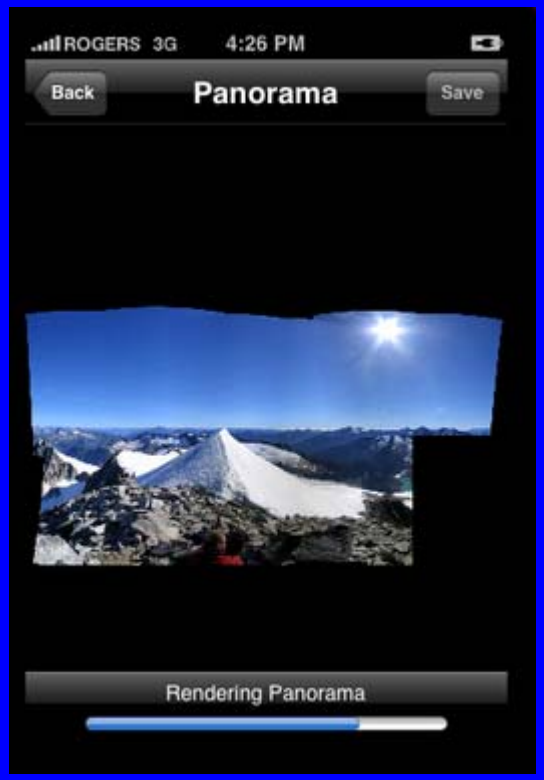
Photo

Topics (cont'd)

- Image features – corners
- Image matching – RANSAC, Hough transform
- Geometric transformations

Mosaicing


ROGERS 3G
4:08 PM
Photo Albums
eldorado
Cancel
Clear
Photos to Stitch
Stitch


ROGERS 3G
4:26 PM
Back
Panorama
Save
Rendering Panorama

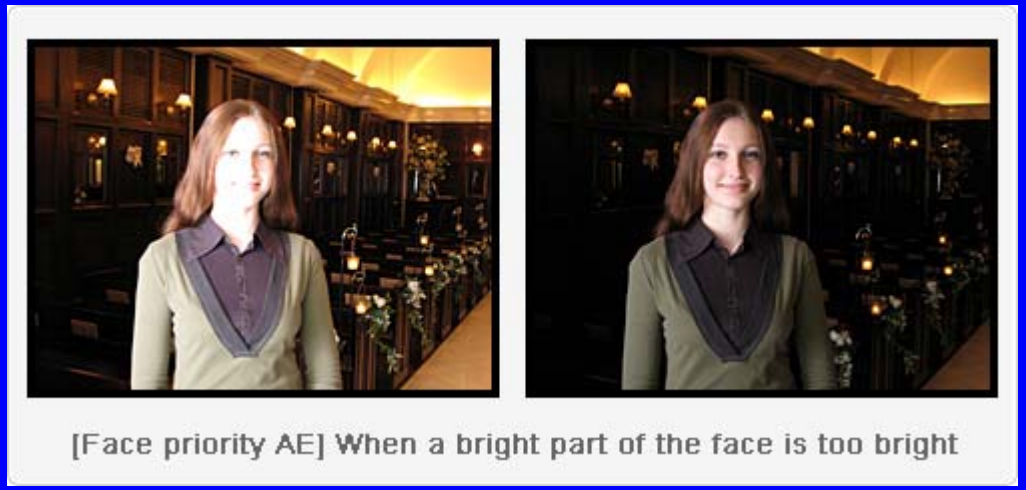
Tracking

(Comaniciu and Meer)

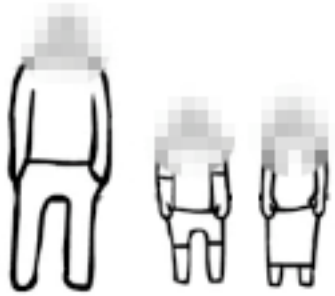
Topics (cont'd)

- Biological vision
- Cameras and perspective
- Stereo
- Optical flow
- Structure-from-motion
- Lighting, photometric stereo
- Classification – including fine-grained classification
- Shape

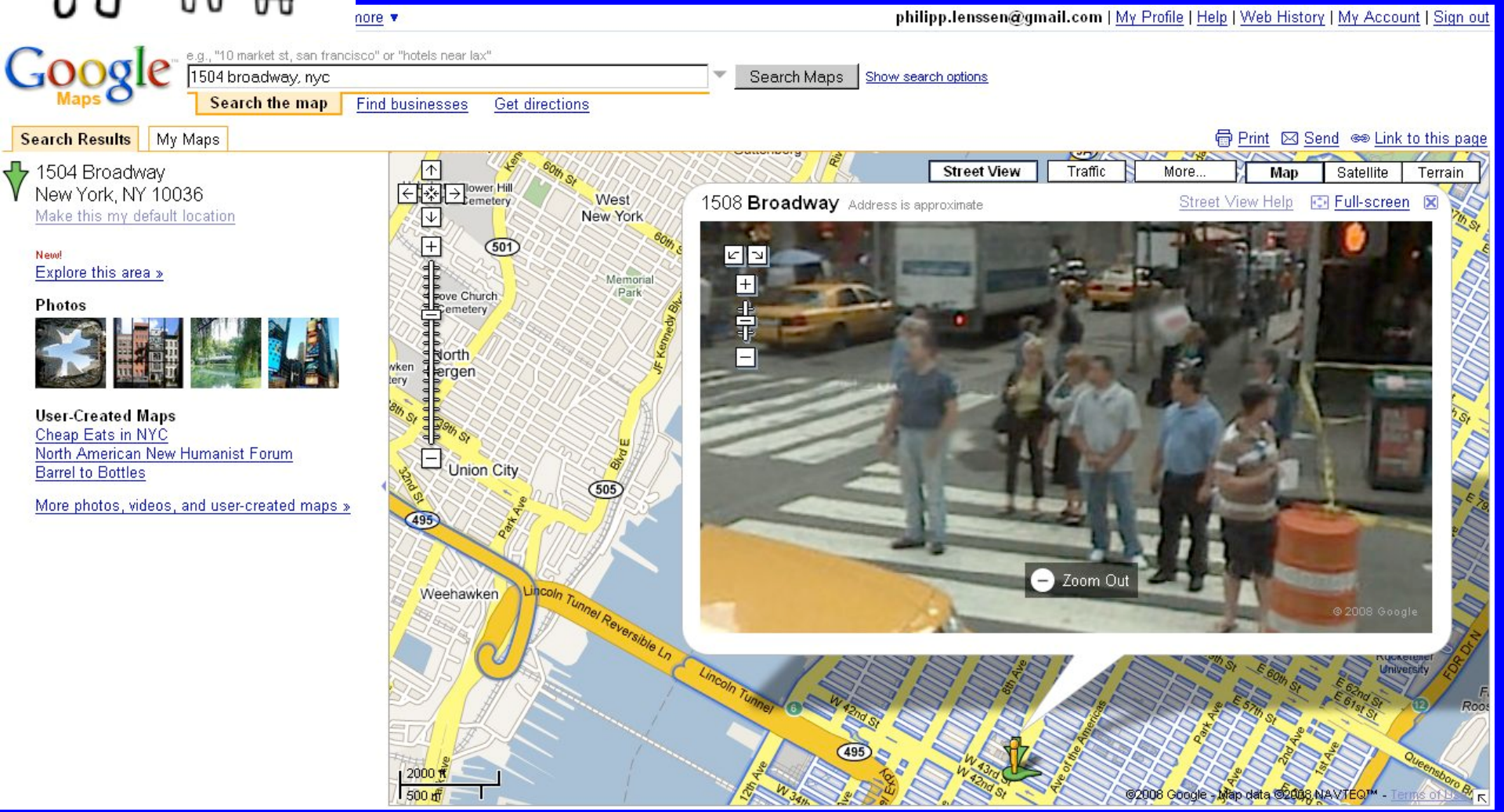
Detection



(Slide from Antonio Torralba)

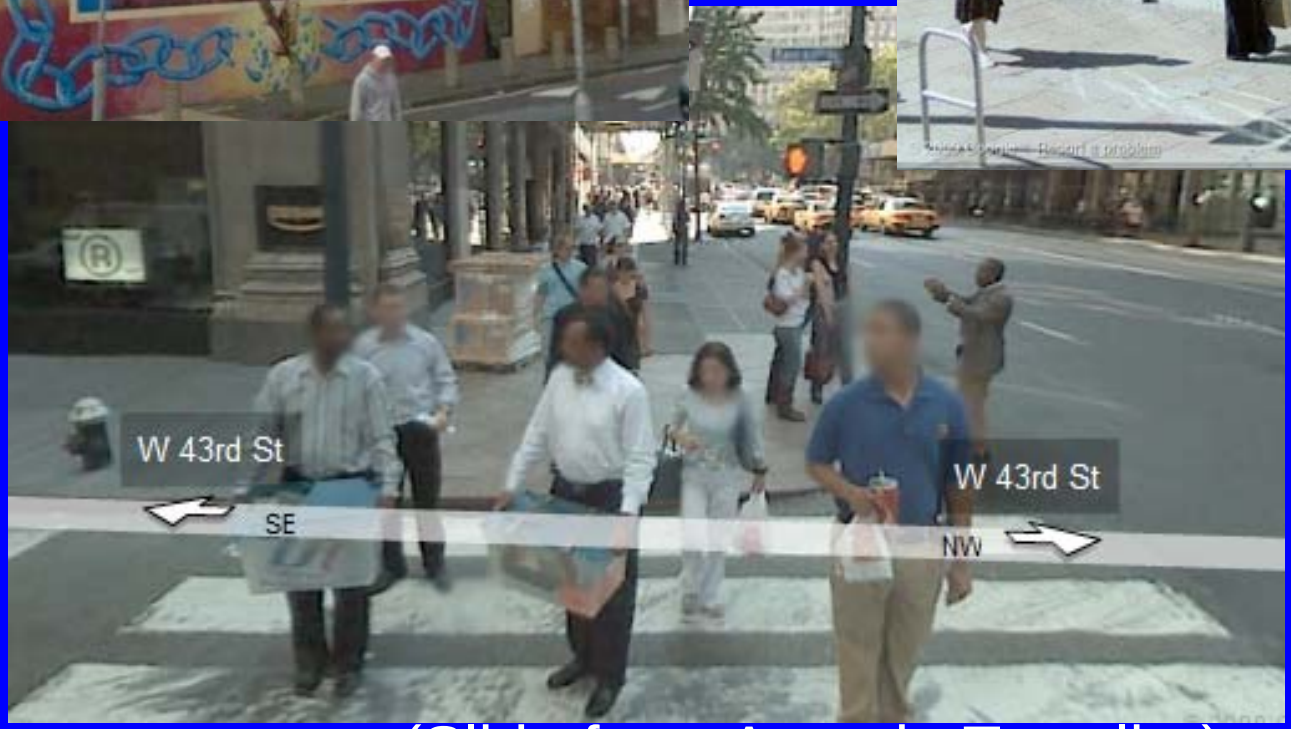


Google street view



(Slide from Antonio Torralba)

Google street view



(Slide from Antonio Torralba)

Some themes

- Correspondence between images
  - Stereo is all about correspondence
  - Optical flow is correspondence w/ linear approximation.
  - Classification requires corresponding features.
- Image gradients and image change
  - Edge detection = gradients
  - Corners: gradients in 2 directions
  - Optical flow: combines temporal and spatial gradients.
  - Biological vision sensitive to *changes*
- Statistical modeling
  - Diffusion processes, background, texture, classes

Depth: key equations and algorithms. For example…

- Convolution theorem
- Diffusion equation
- Brightness change constraint equation
- Matrix factorization for SFM
- Graph cuts for MRFs
- Dynamic programming for stereo matching
- E-M

Course Work

- Lectures
- Problem sets/projects (40%)
  - Implement 5 classic algorithms
- Midterm – Take home (20%)
- Final Exam – In class (40%)
- Readings
  - Listed on class web page
  - Optional texts: Szeliski, Forsyth and Ponce

Modeling + Algorithms

- Build a simple model of the world
  (eg., flat, uniform intensity).
- Find provably good algorithms.
- Experiment on real world.
- Update model.

Problem: Too often models are simplistic or intractable.

Bayesian inference

- Bayes law: P(A|B) = P(B|A)*P(A)/P(B).
- P(world|image) =
  P(image|world)*P(world)/P(image)
- P(image|world) is computer graphics
  - Geometry of projection.
  - Physics of light and reflection.
- P(world) means modeling objects in world.

  Leads to statistical/learning approaches.

Problem: Too often probabilities can't be known and are invented, or learned from inadequate data.

Discriminative Vision

- Develop problem with well defined i/o
- Gather lots of labeled data
- Train classifier/regressor
- *Problem:* Generalization between problems is difficult, can we get enough training data? Does this really provide understanding?

Engineering

- Focus on definite tasks with clear requirements.
- Try ideas based on theory and get experience about what works.
- Try to build reusable modules.

Problem: Solutions that work under specific conditions may not generalize.

Marr

- Theory of Computation
- Representations and algorithms
- Implementations.

- Primal Sketch
- 2½D Sketch
- 3D Representations

Problem: Are things really so modular?

Contact Info

Prof: David Jacobs Email: djacobs@cs.umd.edu

Office: Room 4421, A.V. Williams Building (Next to CSIC).

Office Hours: Tuesday, 2-3

TA: Jin Sun Email: jinanderson@gmail.com

Tools Needed for Course

- Math
  - Calculus
  - Geometry
  - Probability
- Computer Science
  - Algorithms
  - Programming, we'll use Matlab.
- Signal Processing (we'll teach a little).

Course Organization

- Reading: Class notes, no required text.
- 8 Problem sets, six with programming.
- Midterm, Final Exam.
- Grading: Problem sets 40%, midterm 20%; final 40%.
- Web page: www.cs.umd.edu/~djacobs/CMSC733/CMSC733_13.htm

Questions?