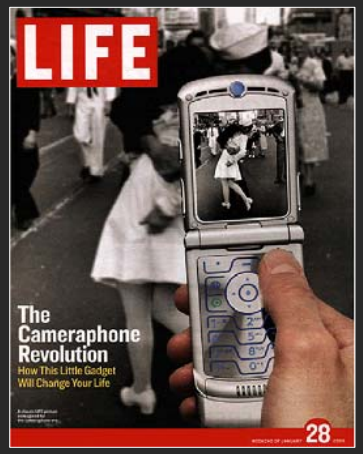
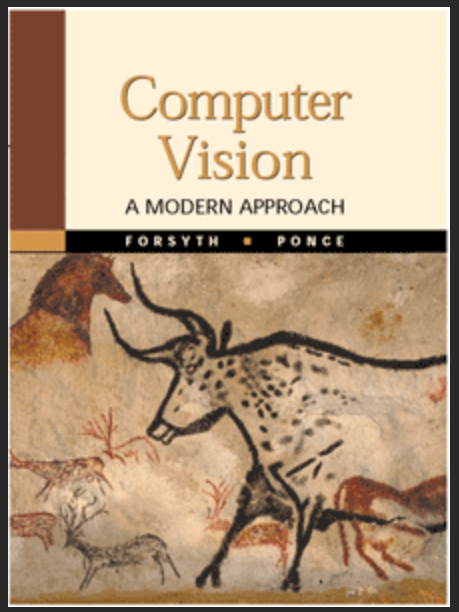
COMP 776: Computer Vision



Basic Info

- Instructor:
  Svetlana Lazebnik (lazebnik@cs.unc.edu)

- Office hours:
  By appointment, FB 244



- Textbook (recommended):
  Forsyth & Ponce, *Computer Vision: A Modern Approach*

- Class webpage:
  http://www.cs.unc.edu/~lazebnik/spring09

Today

- Introduction to computer vision
- Course overview
- Course requirements

The goal of computer vision

- To perceive the story behind the picture

What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

The goal of computer vision

- To perceive the story behind the picture

- What exactly does this mean?
  - Vision as a source of metric 3D information
  - Vision as a source of semantic information

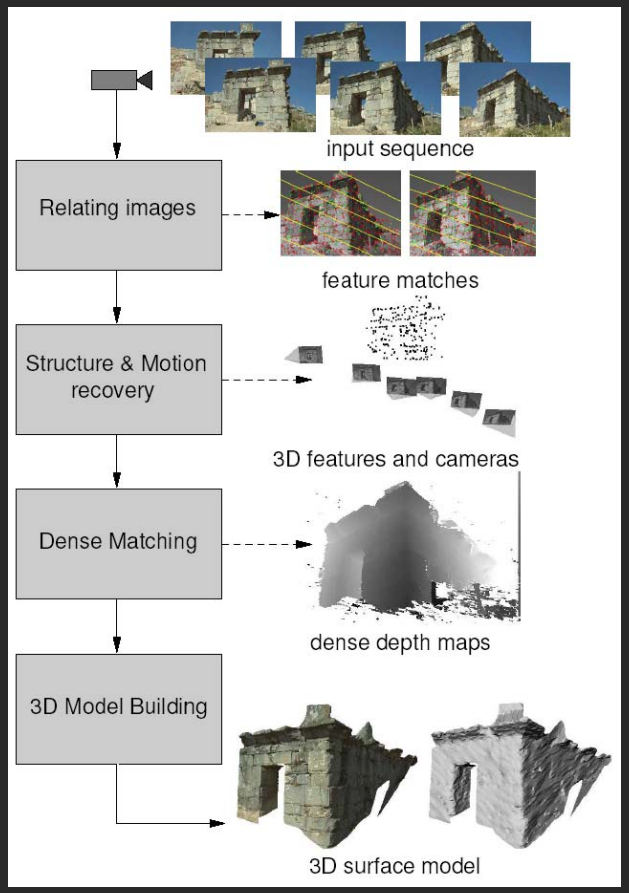
Vision as measurement device

Real-time stereo

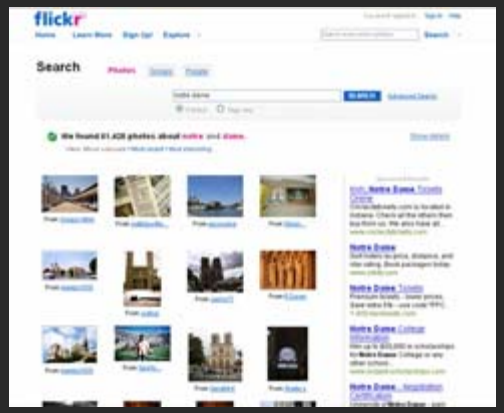
NASA Mars Rover

Pollefeys et al.

Structure from motion



Multi-view stereo for community photo collections



Goesele et al.

Vision as a source of semantic information

slide credit: Fei-Fei, Fergus & Torralba

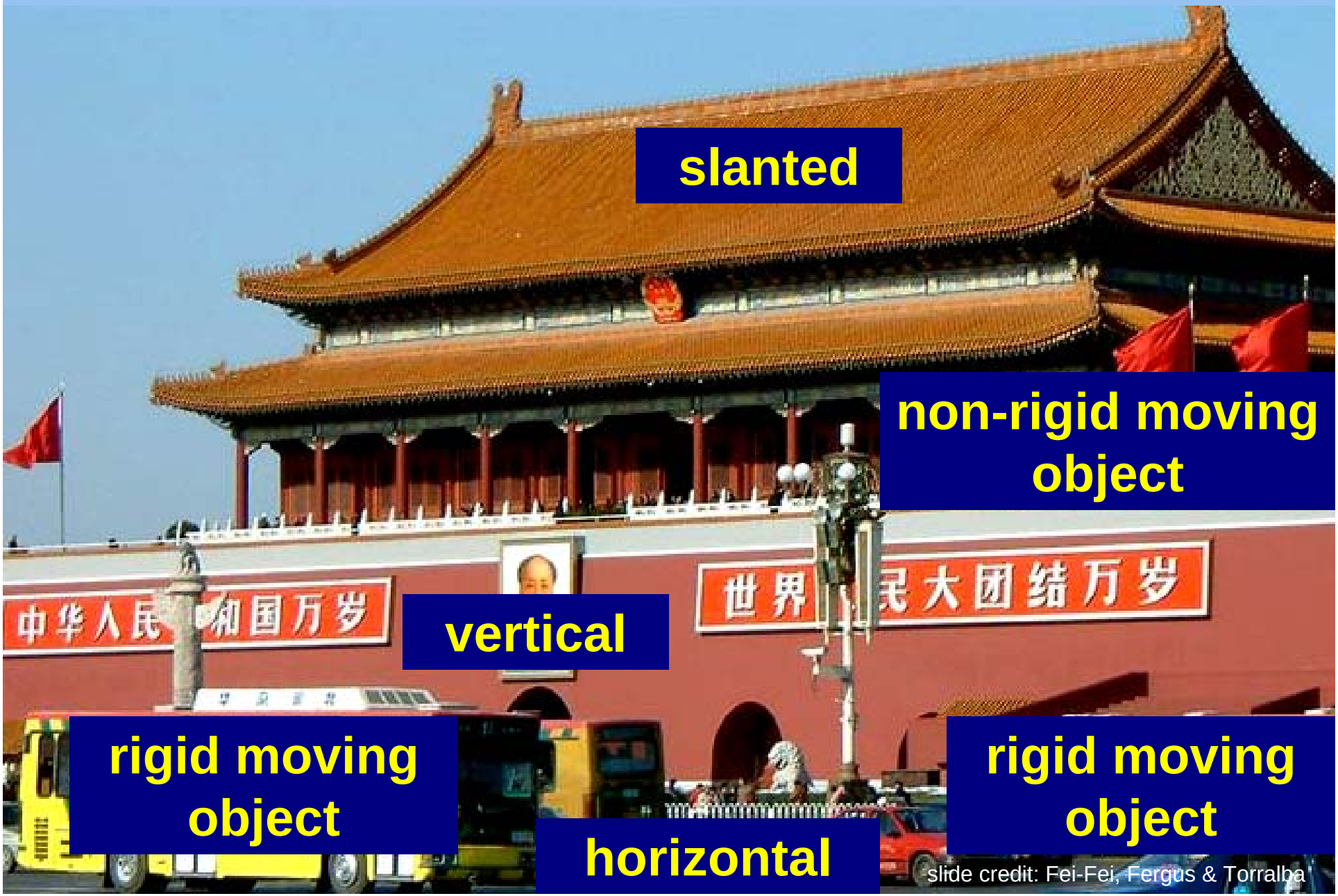
Object categorization

Scene and context categorization

- **outdoor**
- **city**
- **traffic**
- **…**

slide credit: Fei-Fei, Fergus & Torralba

Qualitative spatial information

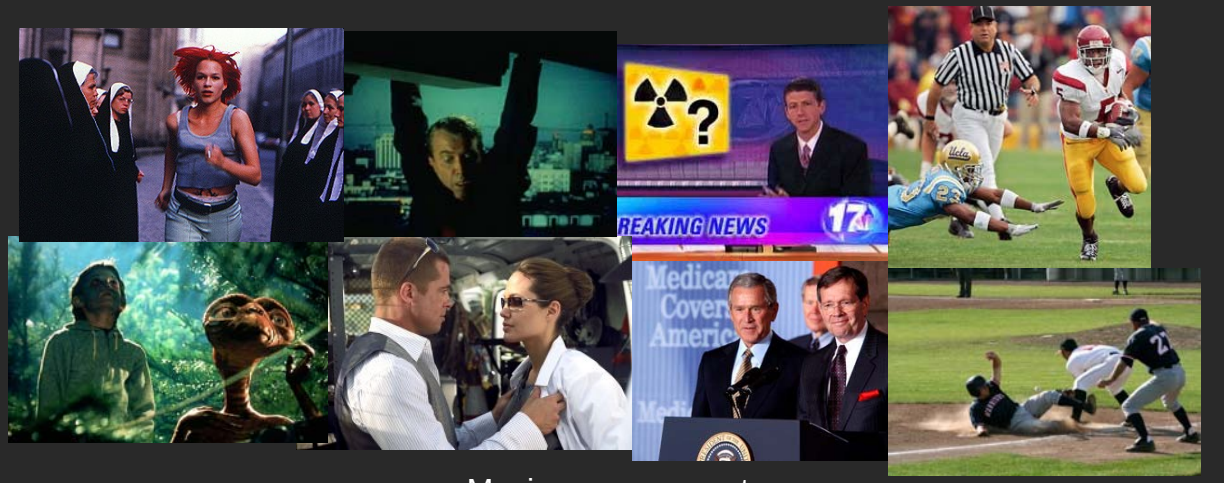


slide credit: Fei-Fei, Fergus & Torralba

Why study computer vision?

- Vision is useful: Images and video are everywhere!

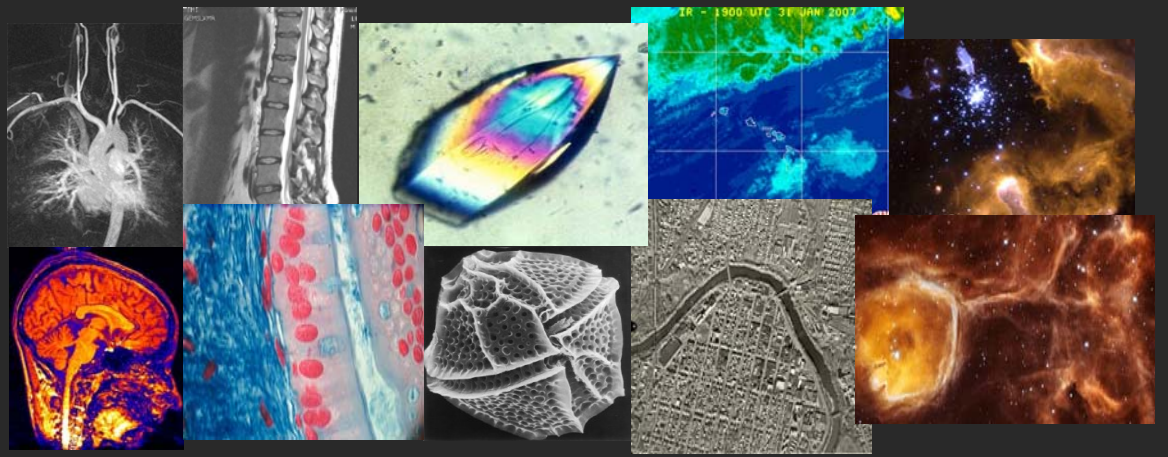
Personal photo albums



Movies, news, sports

Google Image Search™ Picasa™ flickr™ webshots picsearch™ YouTube Broadcast Yourself™

Surveillance and security



Medical and scientific images

Why study computer vision?

- Vision is useful
- Vision is interesting
- Vision is difficult
  - Half of primate cerebral cortex is devoted to visual processing
  - Achieving human-level visual perception is probably “AI-complete”

Why is computer vision difficult?

Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

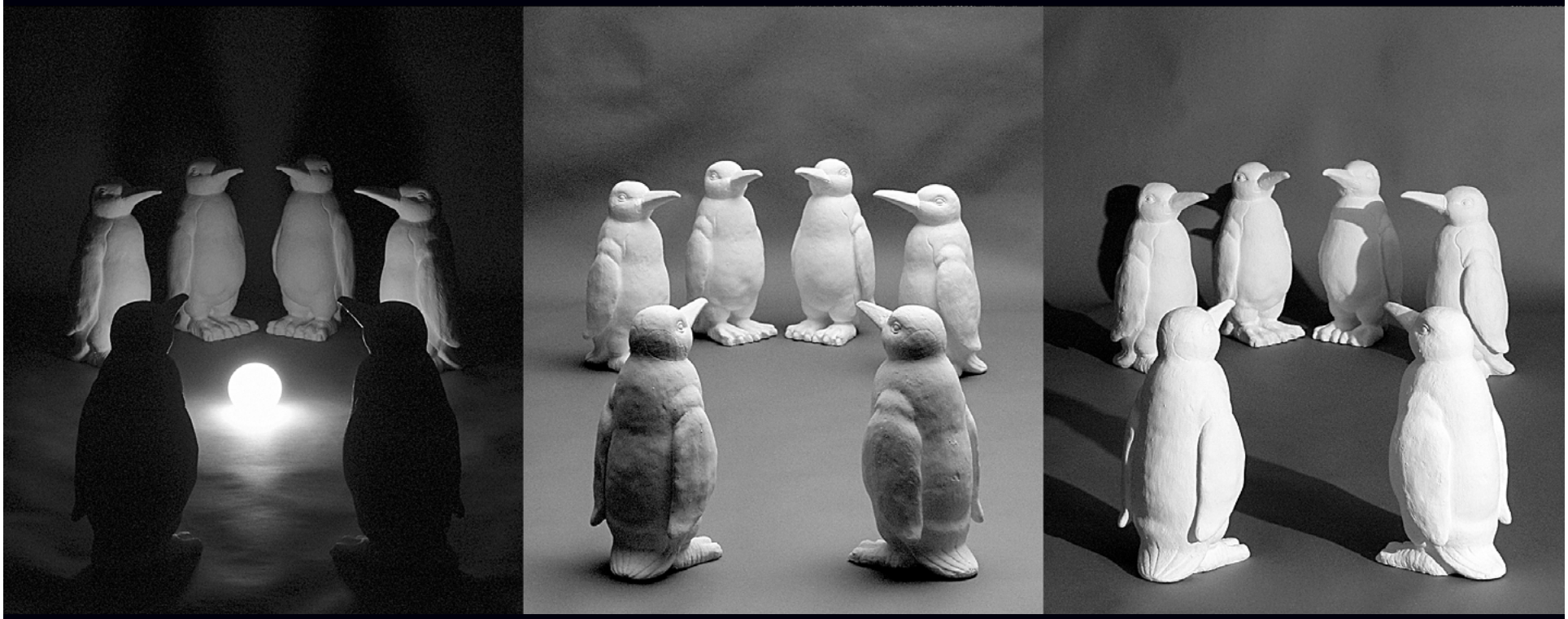
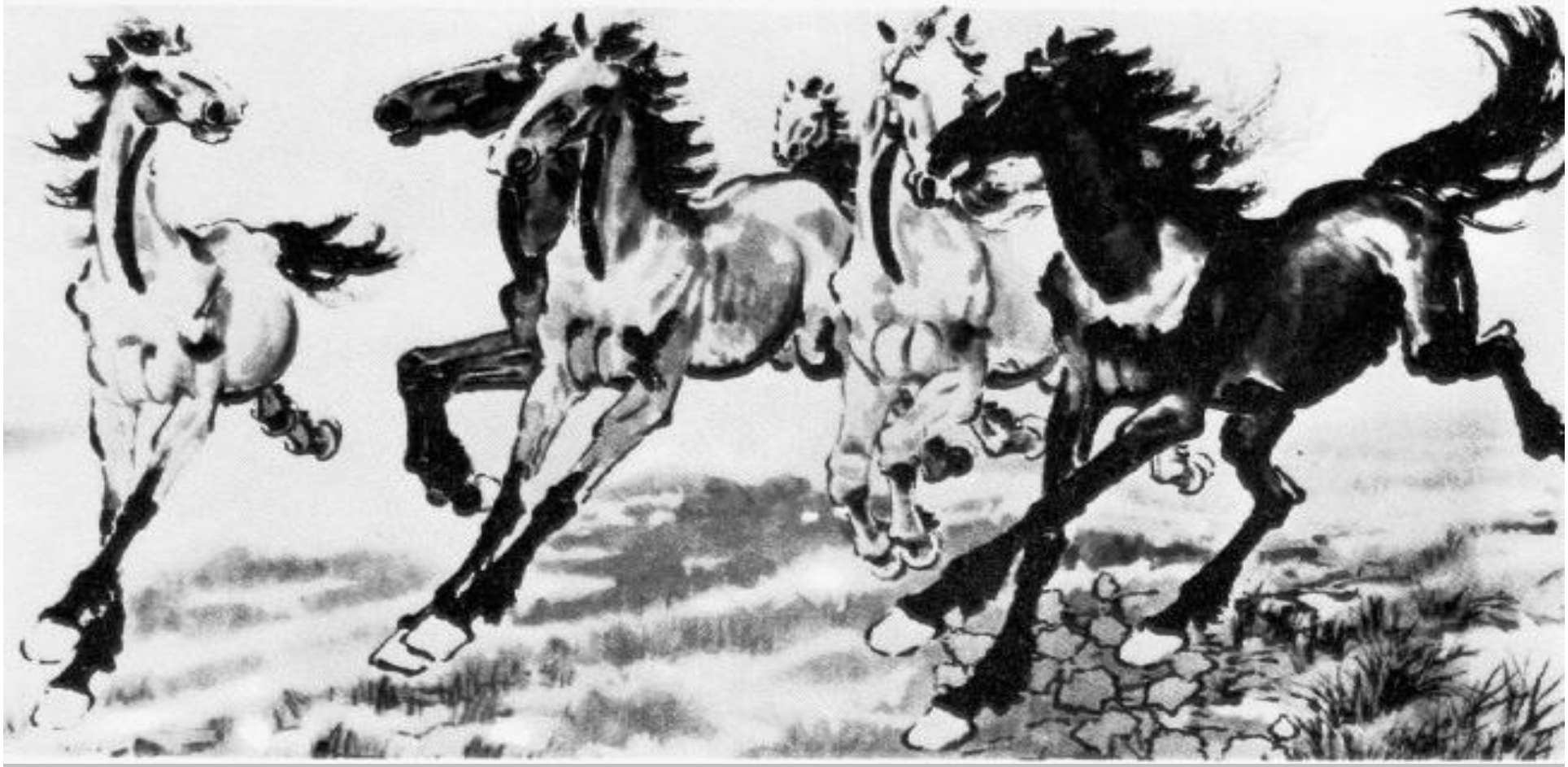


image credit: J. Koenderink

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

Challenges: deformation



Xu, Beihong 1943

slide credit: Fei-Fei, Fergus & Torralba

Challenges: occlusion

Magritte, 1957

slide credit: Fei-Fei, Fergus & Torralba

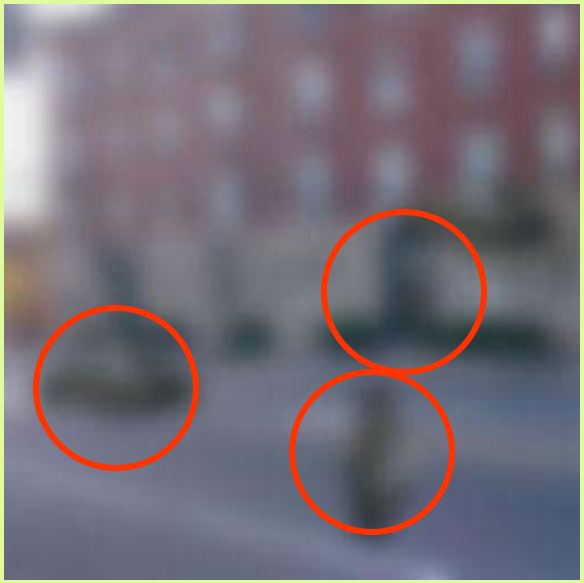
Challenges: background clutter

Challenges: Motion

Challenges: object intra-class variation

slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity



slide credit: Fei-Fei, Fergus & Torralba

Challenges or opportunities?

- Images are confusing, but they also reveal the structure of the world through numerous cues
- Our job is to interpret the cues!

Image source: J. Koenderink

Depth cues: Linear perspective

Depth cues: Aerial perspective

Depth ordering cues: Occlusion

Shape cues: Texture gradient

Position and lighting cues: Cast shadows

Source: J. Koenderink

Grouping cues: Similarity (color, texture, proximity)

Grouping cues: “Common fate”

Image credit: Arthus-Bertrand (via F. Durand)

Bottom line

- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a particular 2D picture

Image source: F. Durand

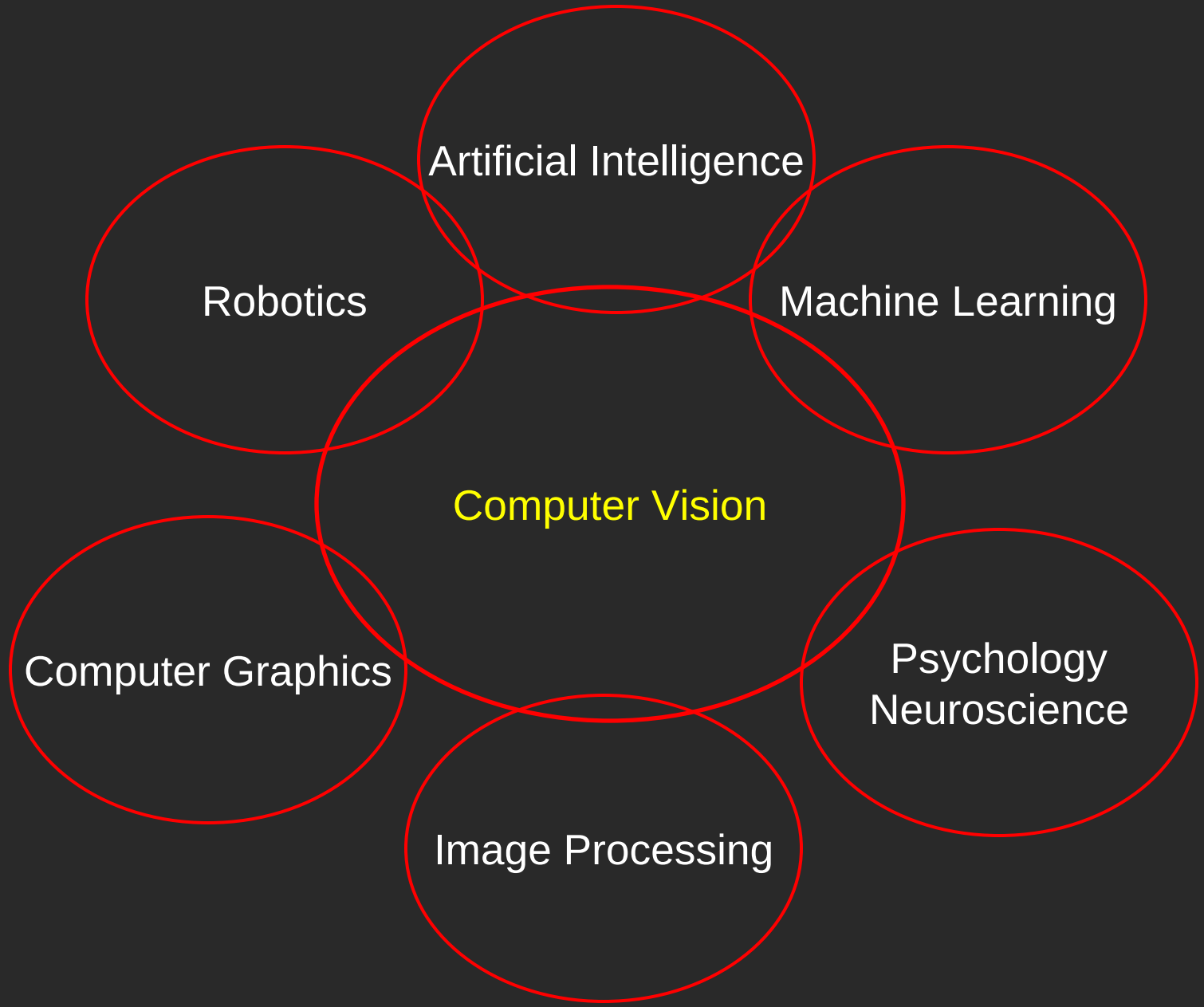
Bottom line

- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a particular 2D picture

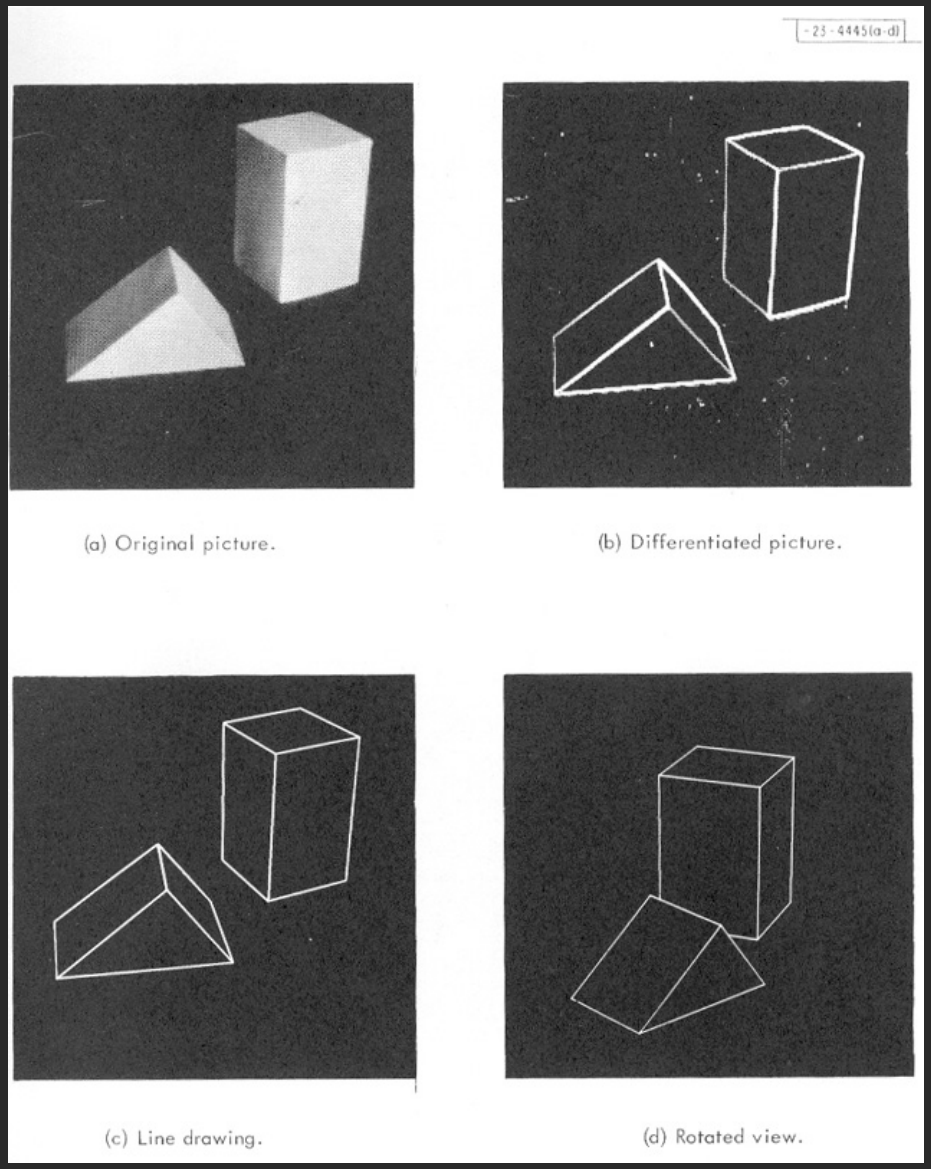
- Possible solutions
  - Bring in more constraints (more images)
  - Use prior knowledge about the structure of the world
- Need a combination of different methods

Image source: F. Durand

Connections to other disciplines



Origins of computer vision



L. G. Roberts, *Machine Perception of Three Dimensional Solids,* Ph.D. thesis, MIT Department of Electrical Engineering, 1963.

Progress to date

The next slides show some examples of what current vision systems can do

Earth viewers (3D modeling)

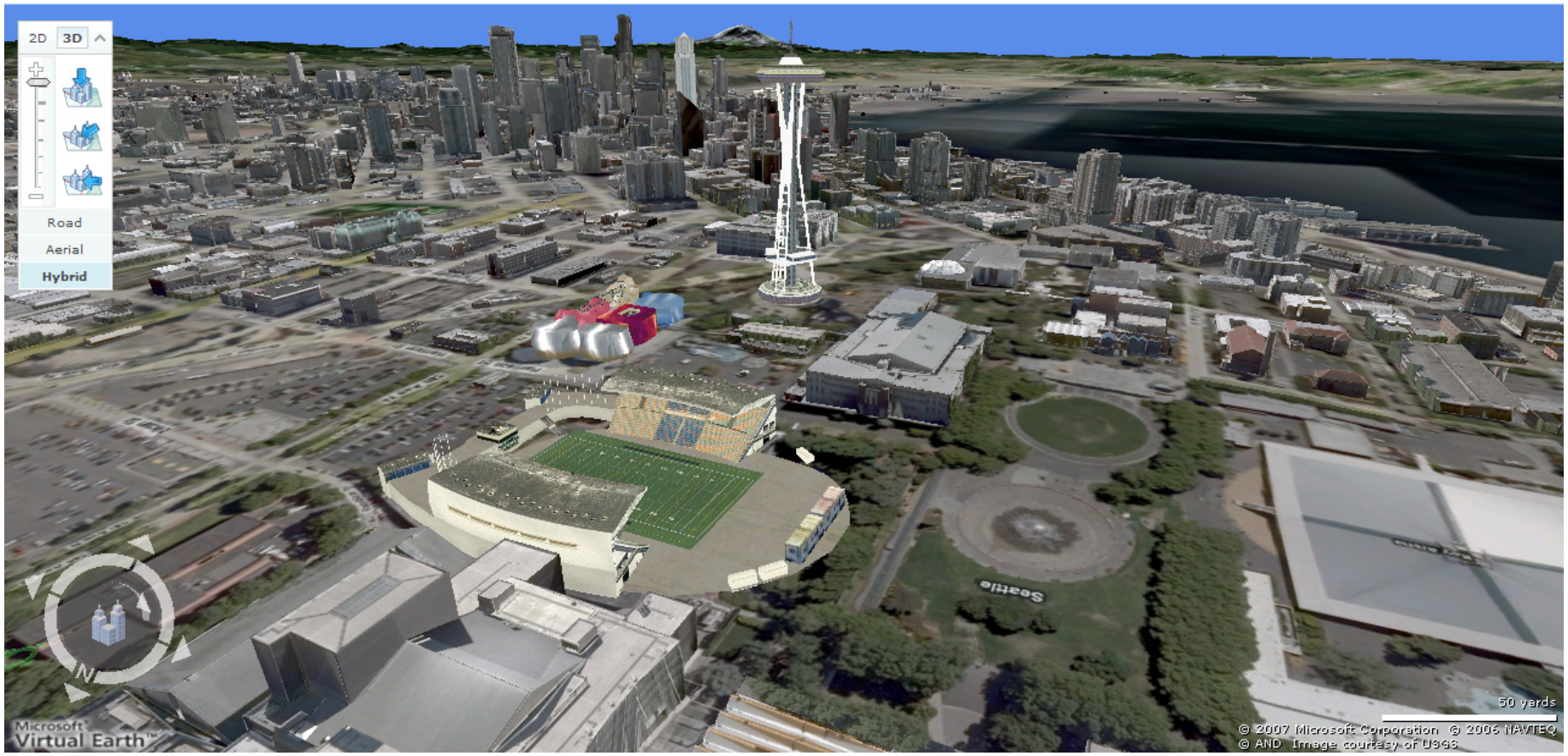
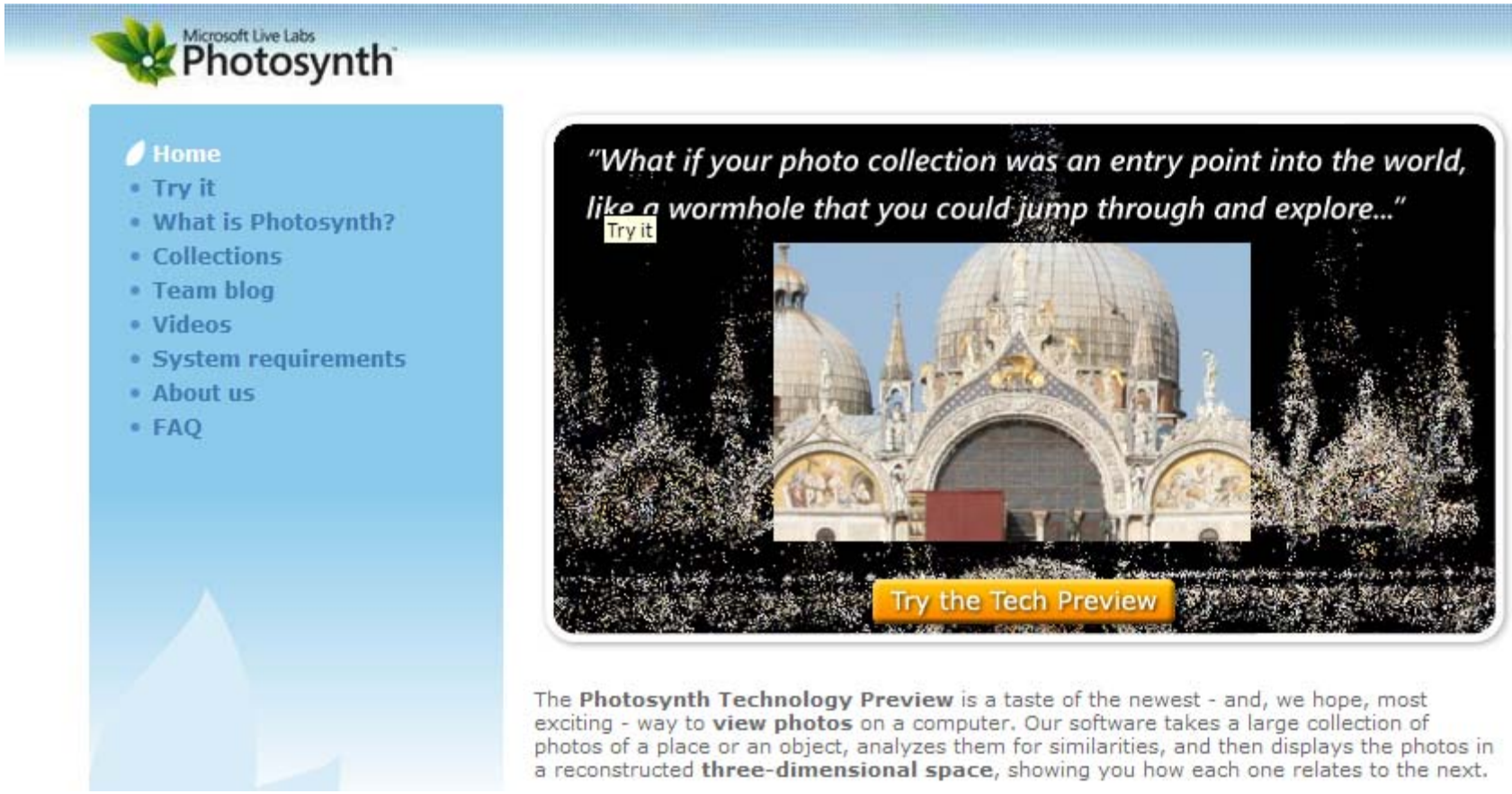


Image from Microsoft's Virtual Earth
(see also: Google Earth)

Source: S. Seitz



http://labs.live.com/photosynth/

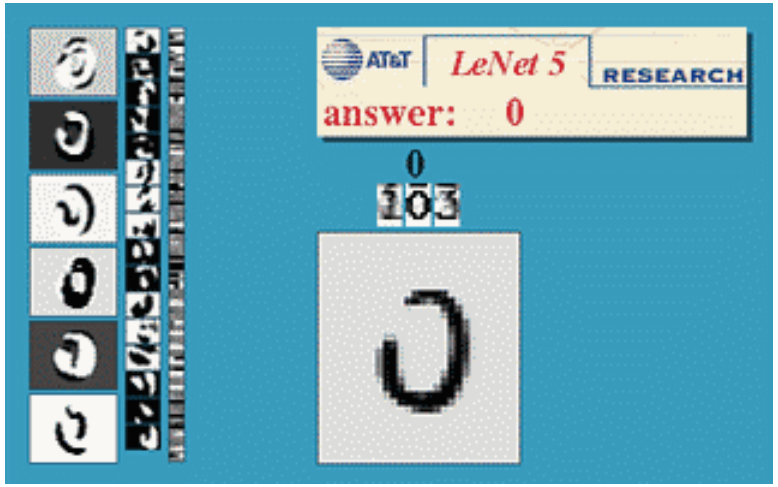
NOTE: Noah Snavely talk at GLUNCH tomorrow!

Source: S. Seitz

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Source: S. Seitz

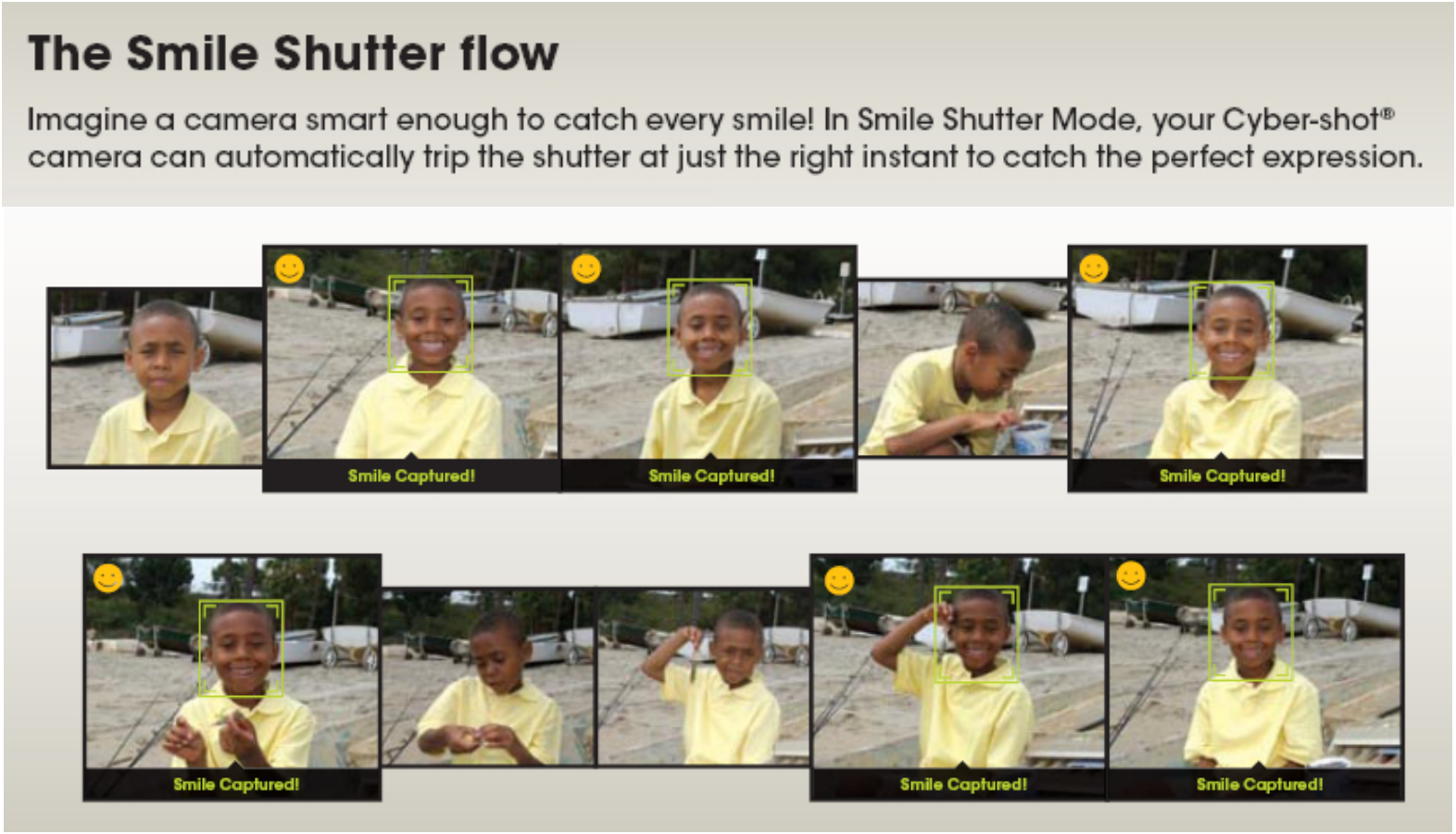
Face detection

Many new digital cameras now detect faces

- Canon, Sony, Fuji, …

Source: S. Seitz

Smile detection?



[Sony Cyber-shot® T70 Digital Still Camera]

Source: S. Seitz

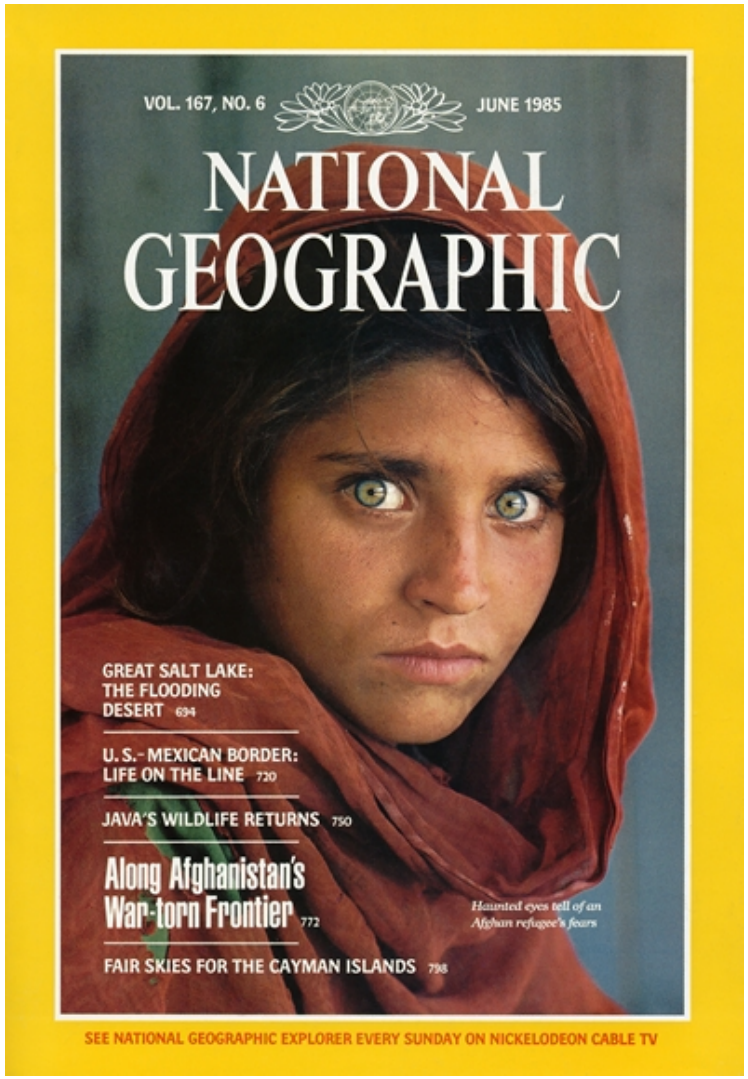
Object recognition (in supermarkets)

LaneHawk by EvolutionRobotics

“A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it… “

Source: S. Seitz

Face recognition

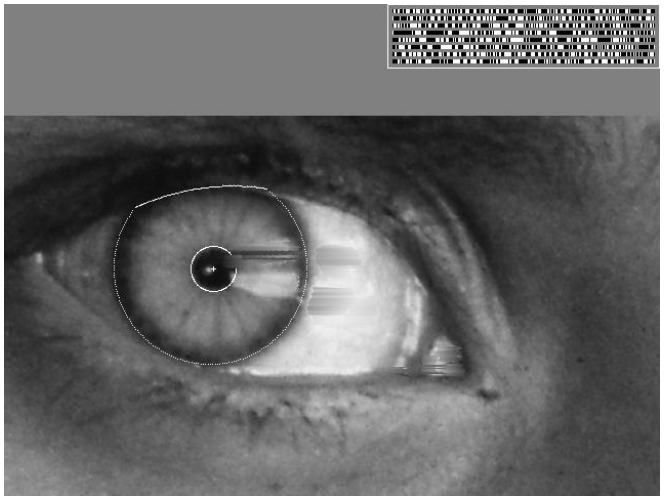


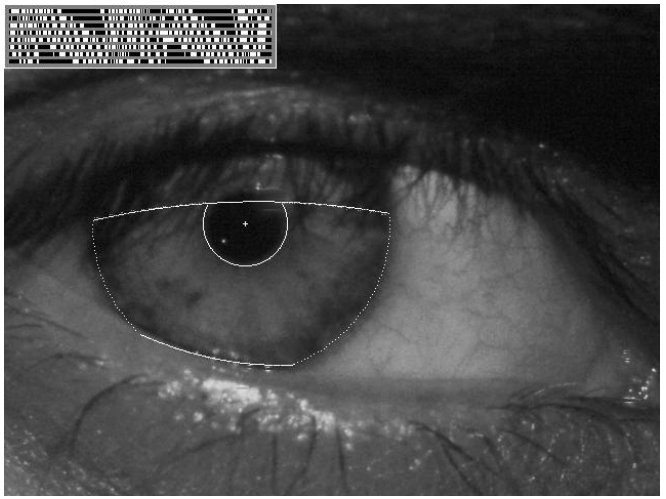
Who is she?

Source: S. Seitz

Vision-based biometrics

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

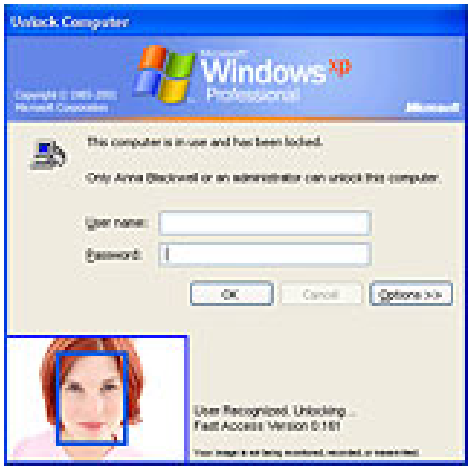




Source: S. Seitz

Login without a password…

Fingerprint scanners on many new laptops, other devices



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

Source: S. Seitz

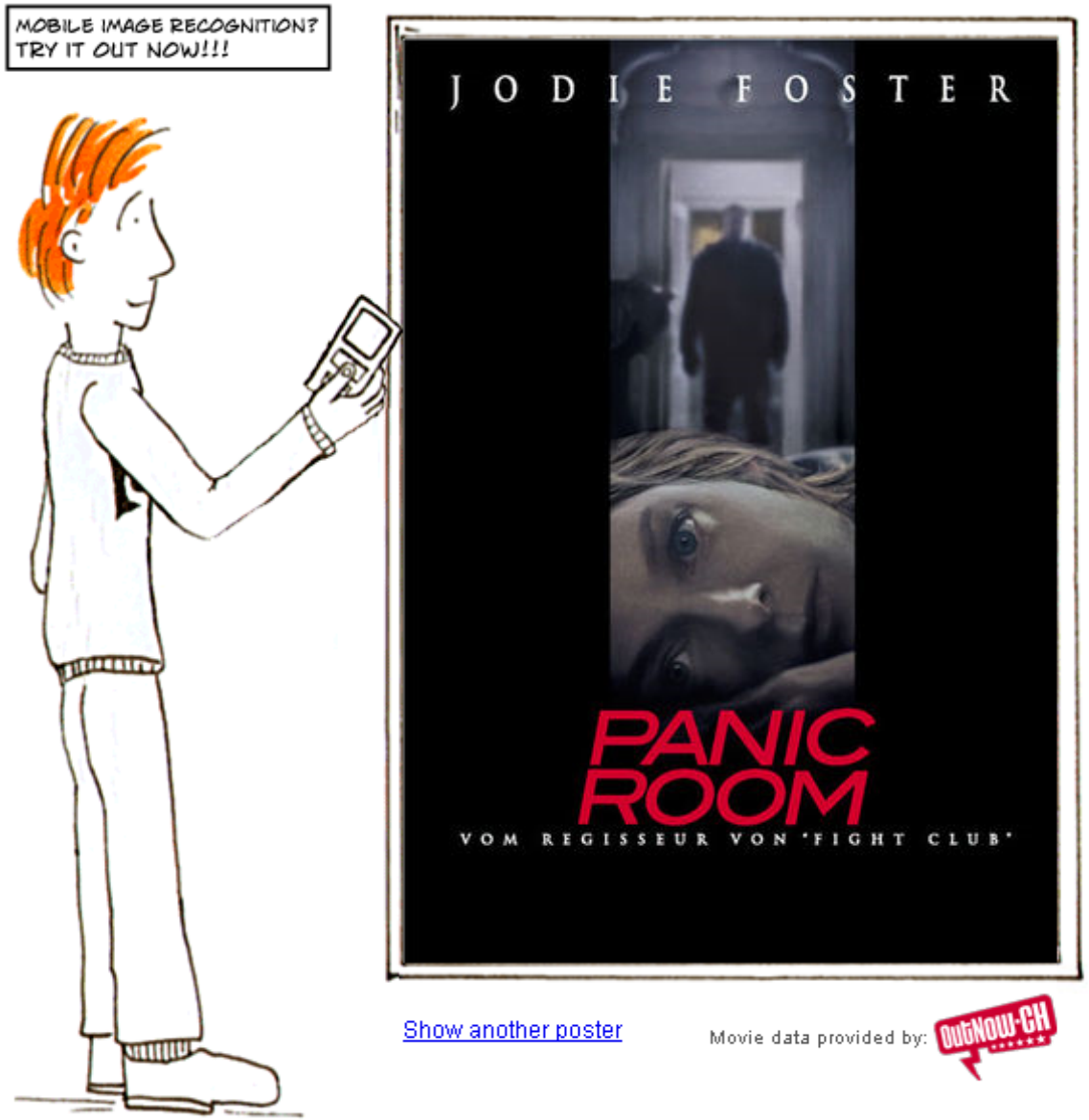
Object recognition (in mobile phones)

This is becoming real:

- Lincoln Microsoft Research
- Point & Find

Source: S. Seitz

iPhone Apps: kooaba (www.kooaba.com)



1. POINT YOUR MOBILE PHONE CAMERA TO THE MOVIE POSTER.

2. SNAP A PICTURE AND SEND IT:

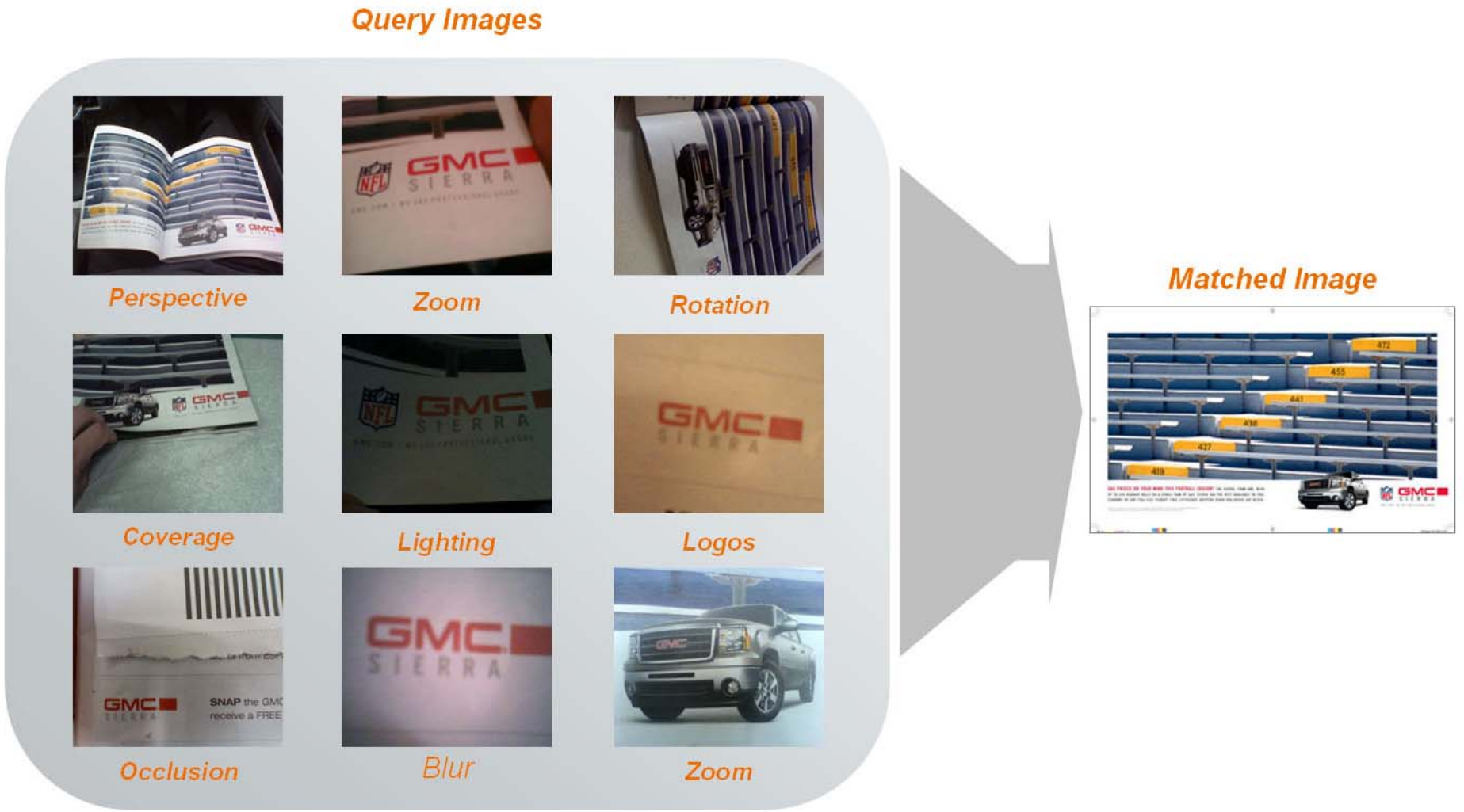
IN SWITZERLAND: MMS TO 5555 (OR 079 394 57 00 FOR ORANGE CUSTOMERS)

IN GERMANY: MMS TO 84000

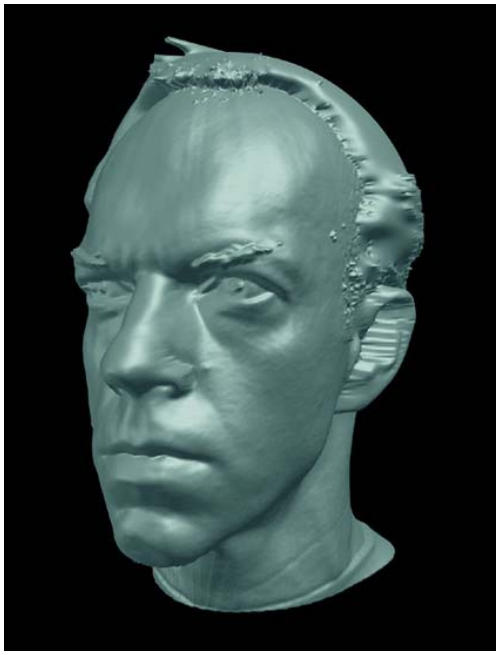
EVERYWHERE: EMAIL TO M@KOOABA.COM

3. FIND ALL RELEVANT INFORMATION ABOUT THE MOVIE ON YOUR MOBILE PHONE

iPhone Apps: snaptell (www.snaptell.com)



Special effects: shape capture



The Matrix movies, ESC Entertainment, XYZRGB, NRC

Source: S. Seitz

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic

Source: S. Seitz

Sports

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

Source: S. Seitz

Smart cars

Slide content courtesy of Amnon Shashua

Mobileye

- Vision systems currently in high-end BMW, GM, Volvo models
- By 2010: 70% of car manufacturers.

Source: S. Seitz

Vision-based interaction (and games)

Nintendo Wii has camera-based IR tracking built in. See Lee's work at CMU on clever tricks on using it to create a multi-touch display!

Sony EyeToy

Assistive technologies

Source: S. Seitz

Vision in space

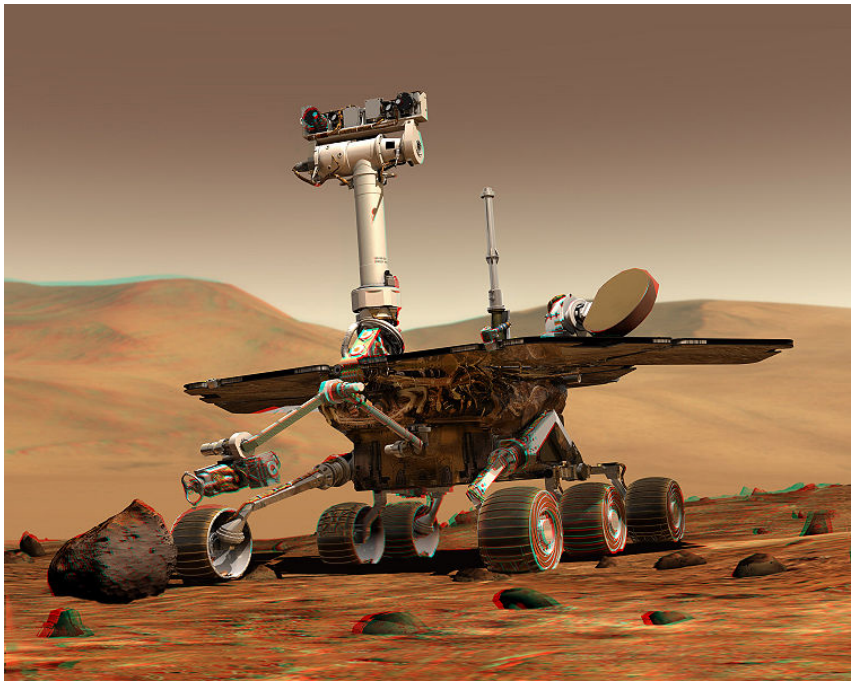
NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

Source: S. Seitz

Robotics



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

http://www.robocup.org/

Source: S. Seitz

The computer vision industry

- A list of companies here:

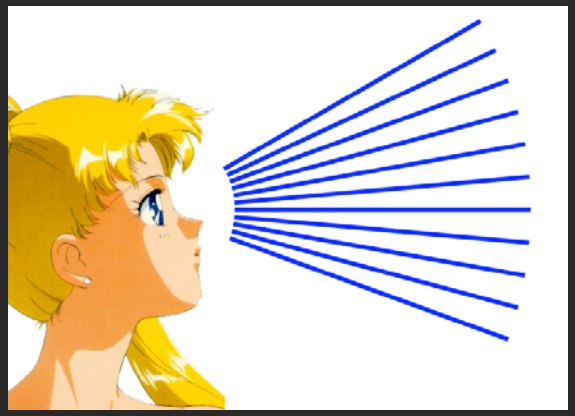
  http://www.cs.ubc.ca/spider/lowe/vision.html

Course overview

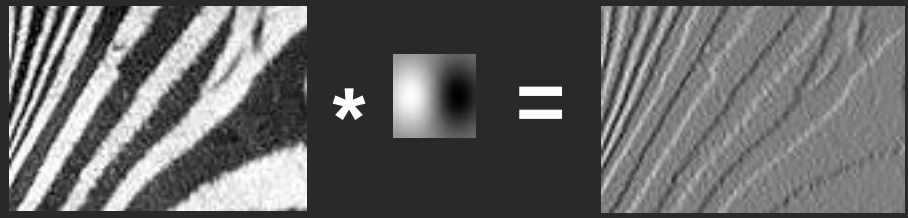
I. Early vision: Image formation and processing
II. Mid-level vision: Grouping and fitting
III. Multi-view geometry
IV. Recognition
V. Advanced topics

I. Early vision

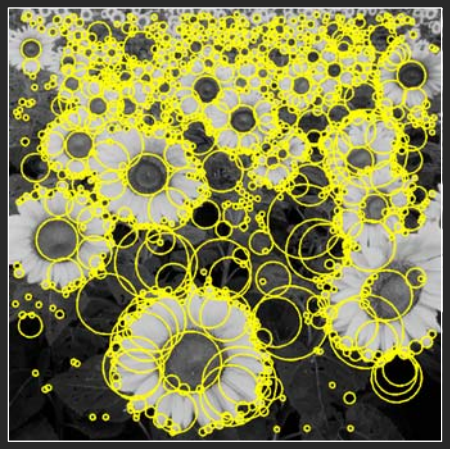
- Basic image formation and processing



Cameras and sensors
Light and color



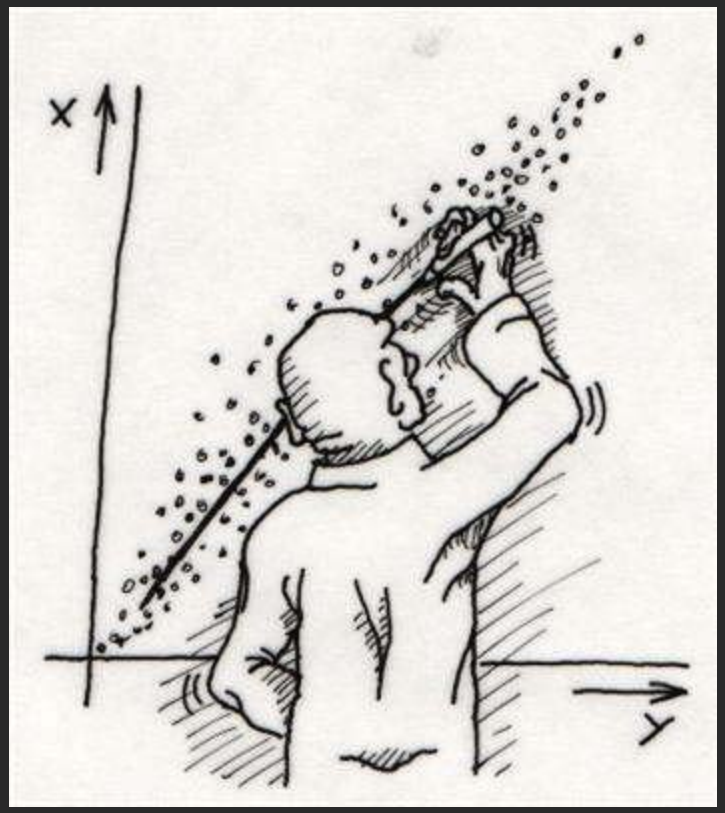
Linear filtering
Edge detection



Feature extraction: corner and blob detection

II. “Mid-level vision”

- Fitting and grouping

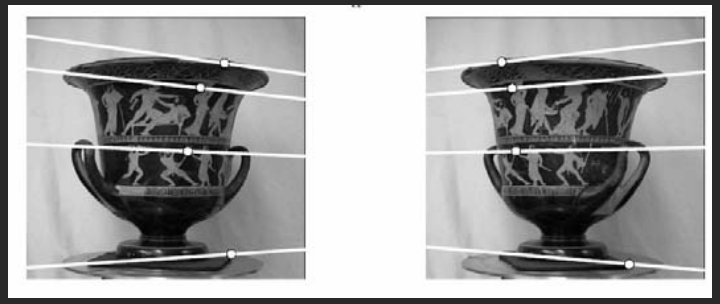


Fitting: Least squares
Hough transform
RANSAC

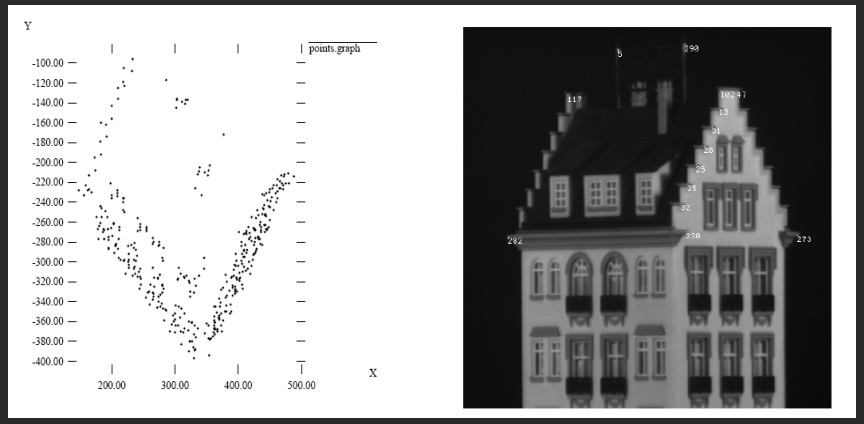
Alignment

III. Multi-view geometry

Stereo



Epipolar geometry

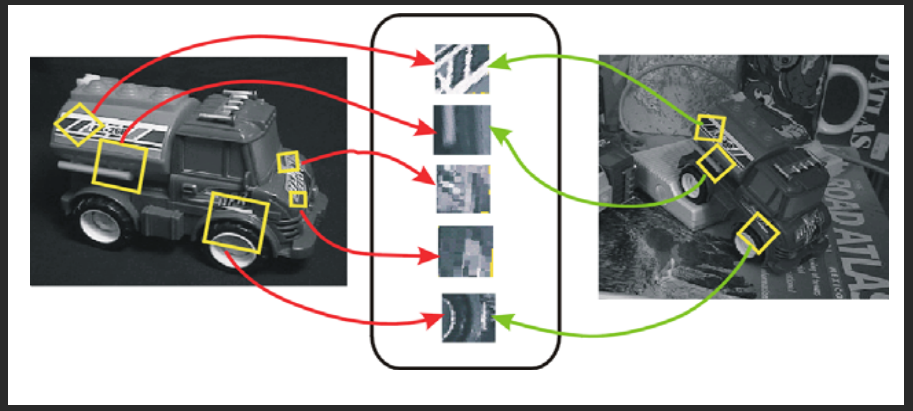


Tomasi & Kanade (1993)

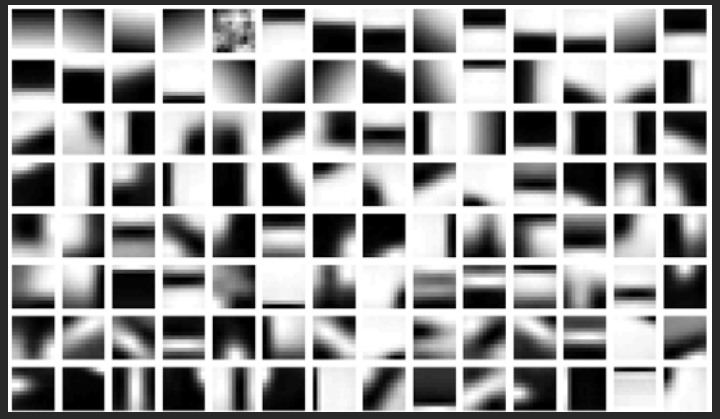
Affine structure from motion

Projective structure from motion

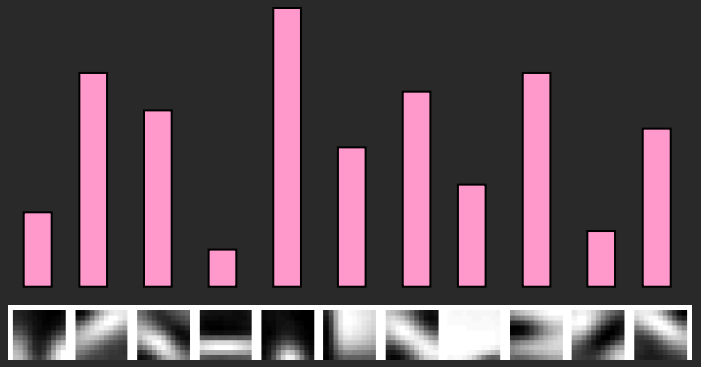
IV. Recognition



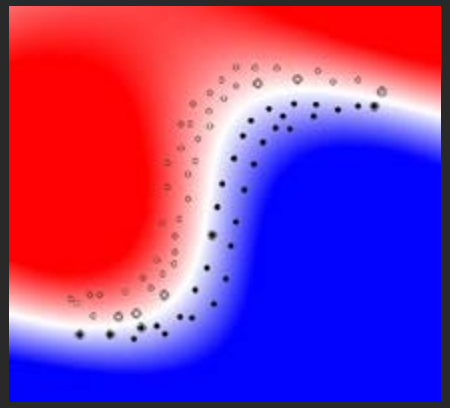
Patch description and matching



Clustering and visual vocabularies



Bag-of-features models



Classification

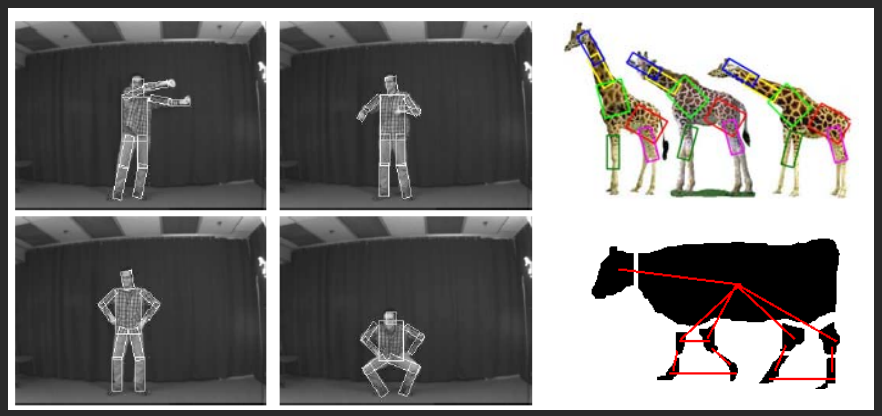
Sources: D. Lowe, L. Fei-Fei

V. Advanced Topics

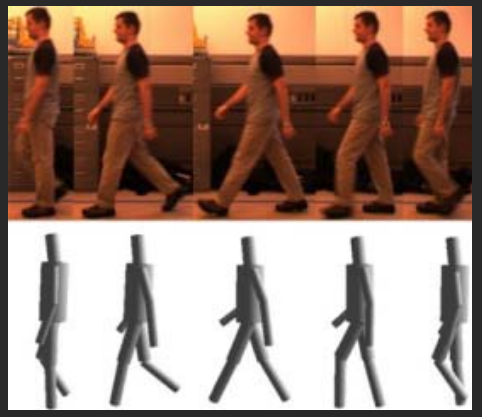
- Time permitting…

Segmentation

Face detection



Articulated models



Motion and tracking

Course requirements

- Philosophy: computer vision is best experienced hands-on
- Programming assignments: 50%
  - Three or four assignments
  - Expect the first one in the next couple of classes
  - Brush up on your MATLAB skills (see web page for tutorial)
- Final assignment: 30%
  - Recognition competition
  - Winner gets a prize!
- Participation: 20%
  - Come to class regularly
  - Ask questions
  - Answer questions

Collaboration policy

- Feel free to discuss assignments with each other, but coding must be done individually
- Feel free to incorporate code or tips you find on the Web, provided this doesn’t make the assignment trivial and you explicitly acknowledge your sources
- Remember: I can Google too!

For next time

- Self-study: MATLAB tutorial
- Reading: cameras and image formation (F&P chapter 1)