CS 4059: Fundamental of Computer Vision

- Office hours:
  Thu & Thu at 11:00 am to 12:00 noon
  or by appointment, NB-010 First floor

- Textbook/Reference:
  - Computer Vision and *Image Analysis – Umbaugh*
  - *Fundamentals of Computer vision-Mubarak Shah*
  - Computer Vision: *Algorithms and Applications – Richard Szeliski*

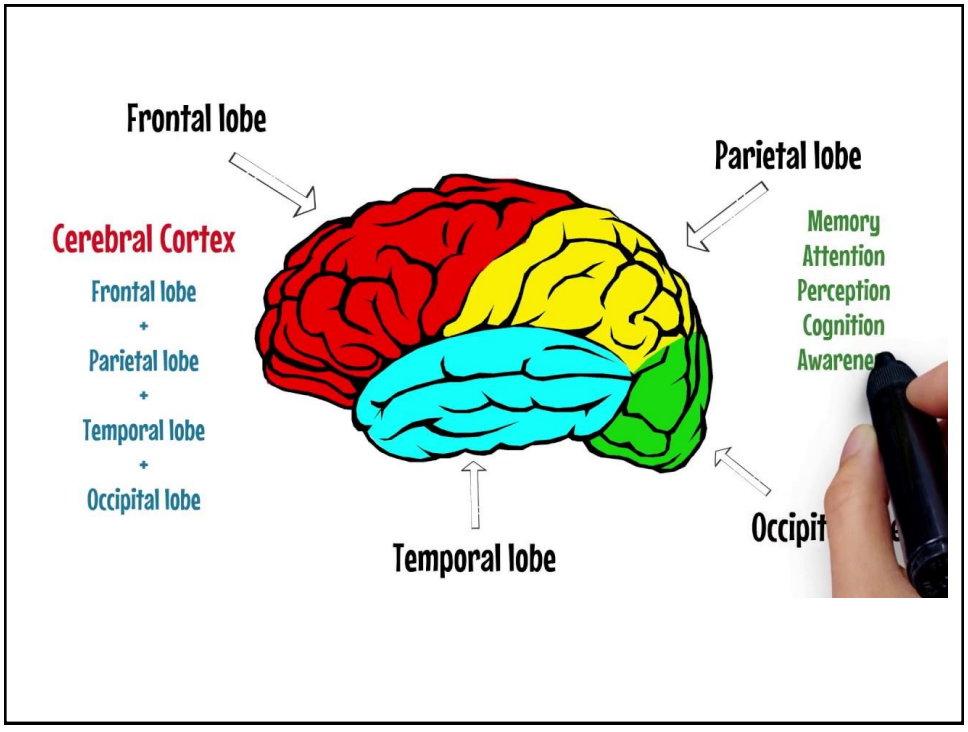
- Class material:
  Google classroom

This week

- Introduction to computer vision
- Course requirements
- Course overview (Broad)

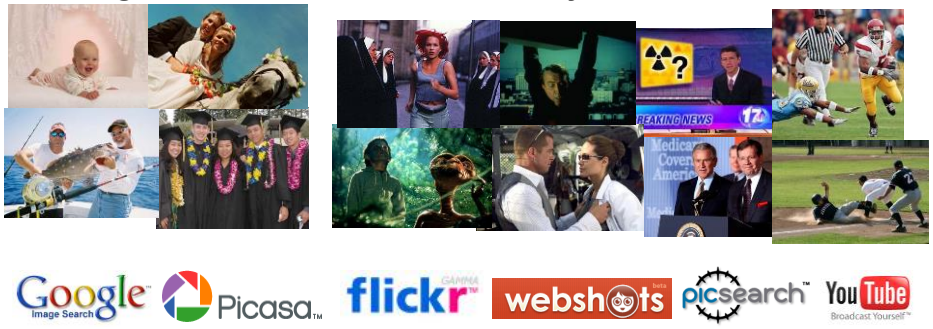
Why study computer vision?

- 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”

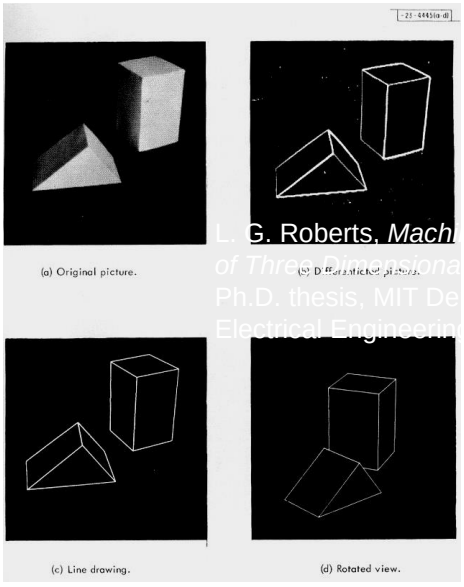


Why study computer vision…

- Images and videos are everywhere!

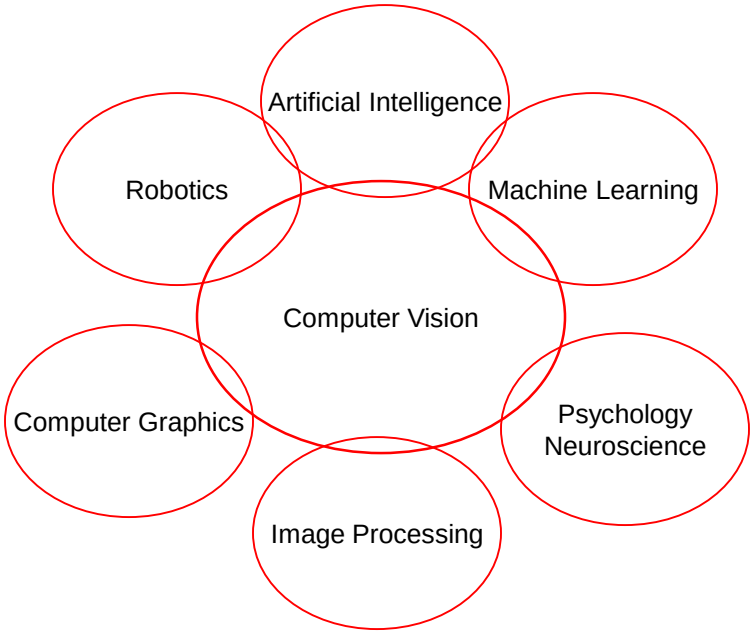


Origins of computer vision

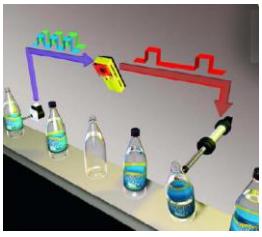


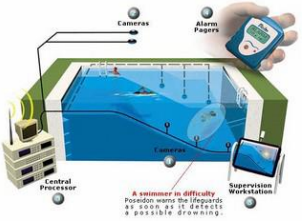
L. G. Roberts, *Machi*
of Three-Dimensiona
Ph.D. thesis, MIT De
Electrical Engineerin

Connections to other disciplines



Applications of computer vision



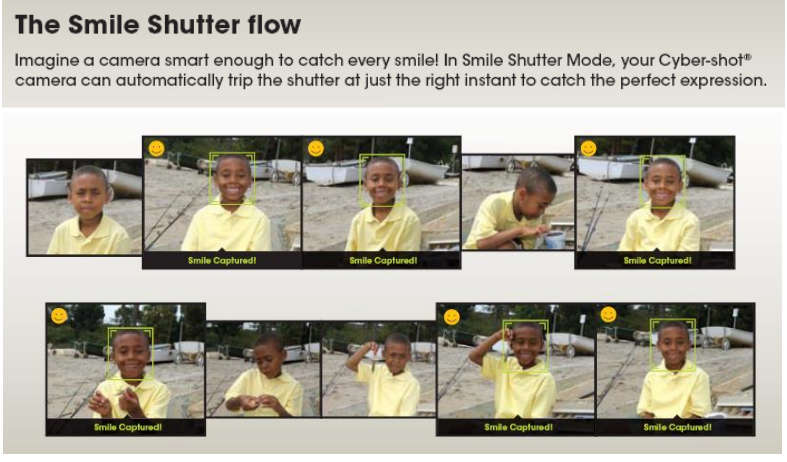


Applications of computer vision

Commercial products integrates tidbits of computer vision

- Most digital cameras now detect faces
  - Canon, Sony, Fuji, …

Smile detection?



Sony Cyber-shot® T70 Digital Still Camera

Object recognition (in supermarkets)

LaneHawk by Evoluti-onRobotics

"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… "

Login without a password…

Fingerprint scanners on many new laptops, other devices

Face recognition systems now beginning to appear more widely

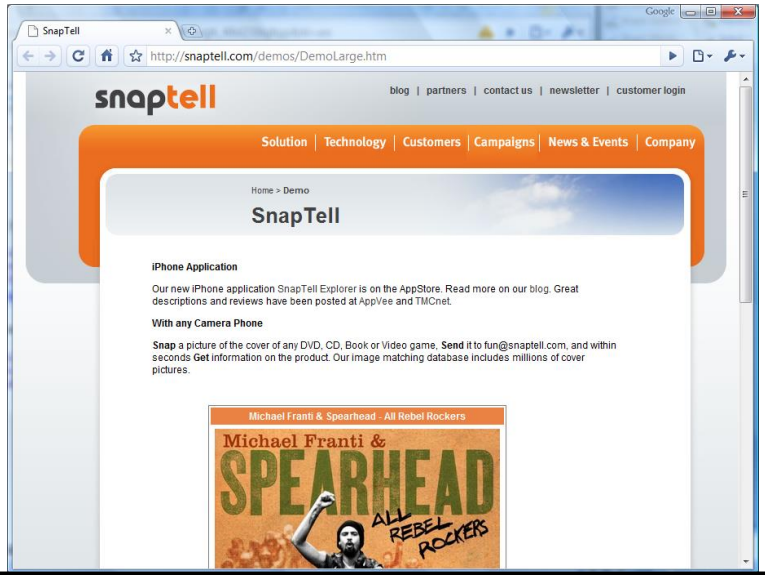
http://www.sensiblevision.com/en-us/products/forhome/overview.aspx/

Object recognition (in mobile phones)

- This is becoming real:
  - Lincoln soft Research
  - Point & Find, Nokia
  - SnapTell.com (now amazon)

Snaptell

http://download.cnet.com/ios/snaptell/3260-20_4-6312649-1.html



Special effects: shape capture

***The Matrix* movies, ESC Entertainment, XYZRGB, NRC**

Special effects: motion capture

***Pirates of the Carribean*, Industrial Light and Magic**

Click here for interactive demo

Sports

***Sportvision* first down line**

Nice explanation on www.howstuffworks.com

Smart cars

- Mobileye
  - Vision systems currently in high-end BM
  - By 2010: 70% of car manufacturers.
  - Video demo

Slide content courtesy of Amnon Shashu

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!

Digimask: put your face on a 3D avatar.

"Game turns moviegoers into Human Joysticks", CNE

Camera tracking a crowd, based on this work.

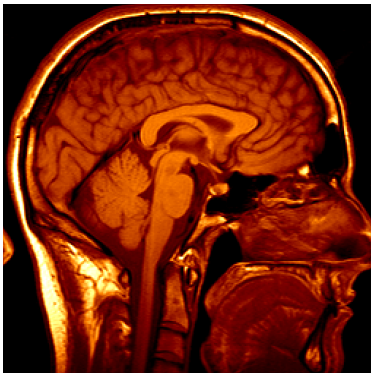
Vision in space

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.

Medical imaging



3D imaging

MRI, CT

Image guided surgery

Grimson et al., MIT

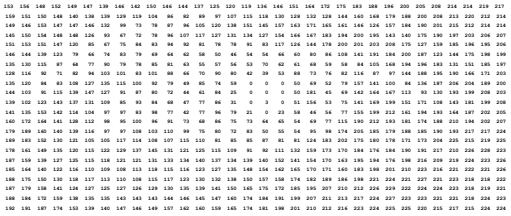
Applications of computer vision

- For more information on the computer vision industry:

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

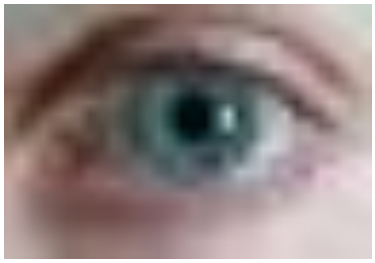
The goal of computer vision

- To perceive the “world behind the picture”



The goal of computer vision

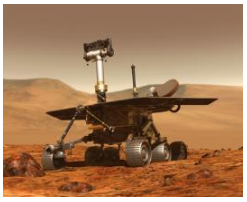
- To perceive the “world behind the picture”



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

Vision as measurement device

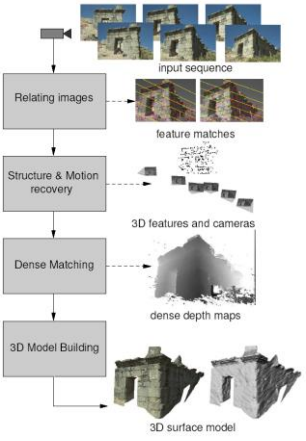
Real-time stereo

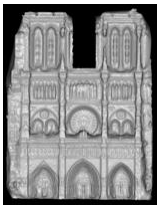


NASA Mars Rover

Pollefeys et al.

Structure from motion





Goesele et al.

- Amazing success story!

…but why do Learning for Vision?

- *"What if I don't care about this wishy-washy recognition stuff? I just want to make my robot go!"*

- Small Reason:
  - For measurement, other sensors are often better (in DARPA Grand Challenge, vision was barely used!)
  - For navigation, you still need to learn!
    Big Reason:
  - The goals of computer vision (**what + where**) are in terms of what humans care about.

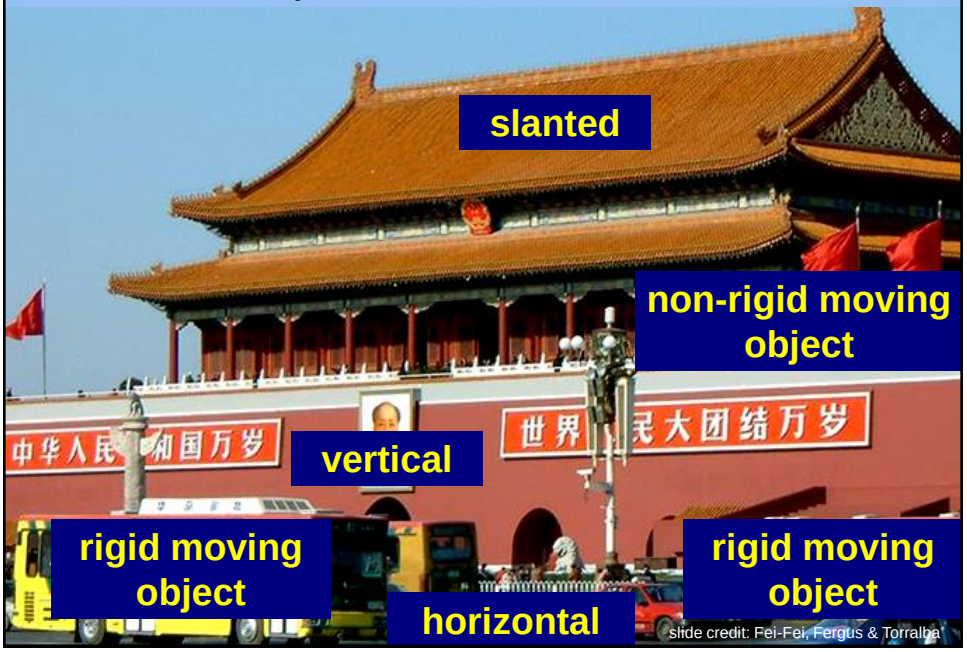
Slide credit: A. Efros

Vision as a source of semantic information

Object categorization

Scene and context categorization

Qualitative spatial information



Challenges: viewpoint variation

Michelangelo 1475-1564

slide credit: Fei-Fei, Fergus & Torralba

Challenges: illumination

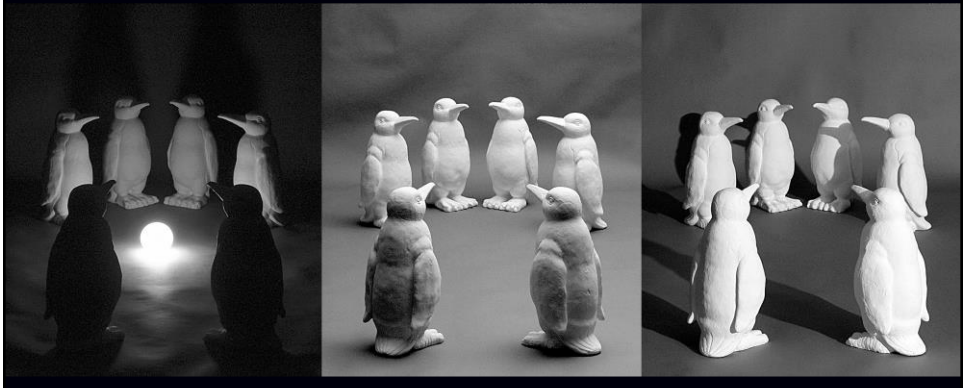
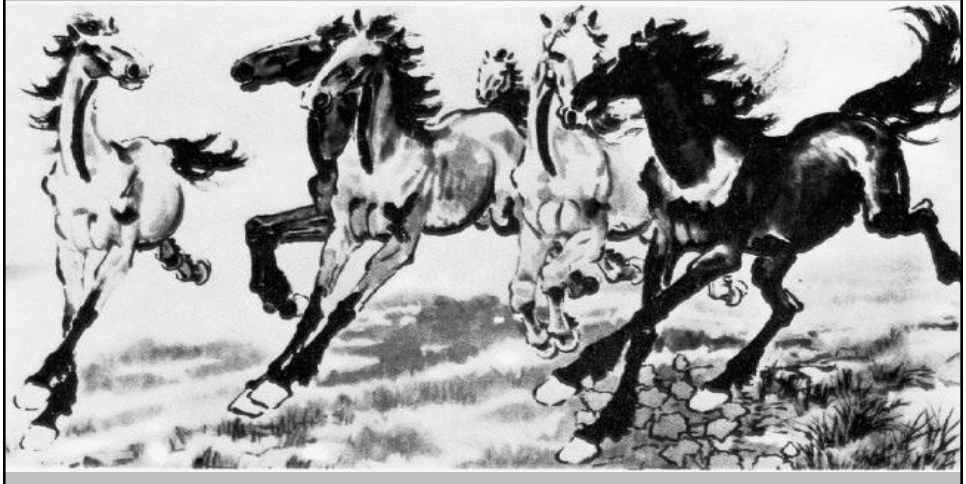


image credit: J. Koenderink

Challenges: scale

slide credit: Fei-Fei, Fergus & Torralba

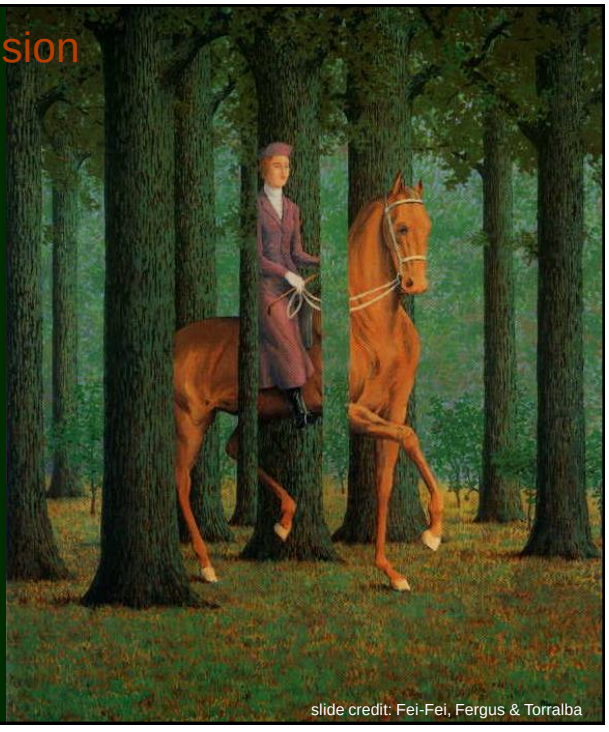
Challenges: deformation



Xu, Beihong 1943

slide credit: Fei-Fei, Fergus & Torralba

Challenges: occlusion



Magritte, 1957

slide credit: Fei-Fei, Fergus & Torralba

Challenges: background clutter

Challenges: object intra-class variation

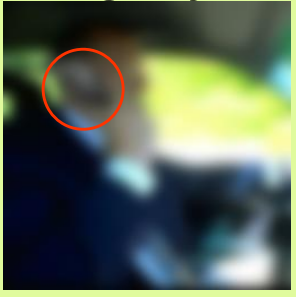
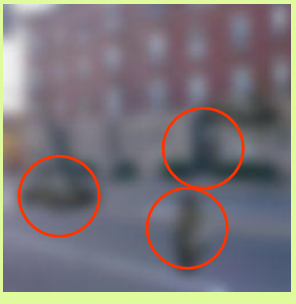





slide credit: Fei-Fei, Fergus & Torralba

Challenges: local ambiguity



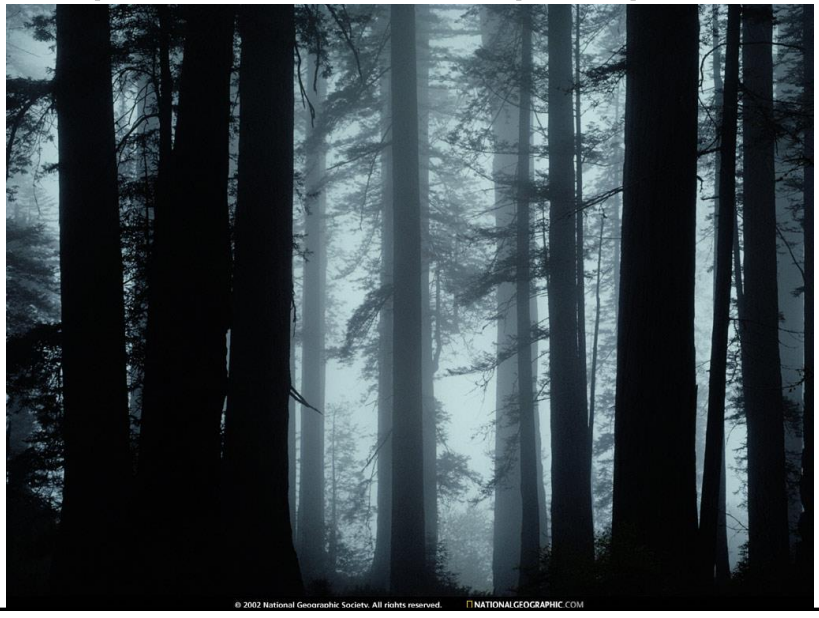

slide credit: Fei-Fei, Fergus & Torralba

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! (e.g. Texture for ICR query)

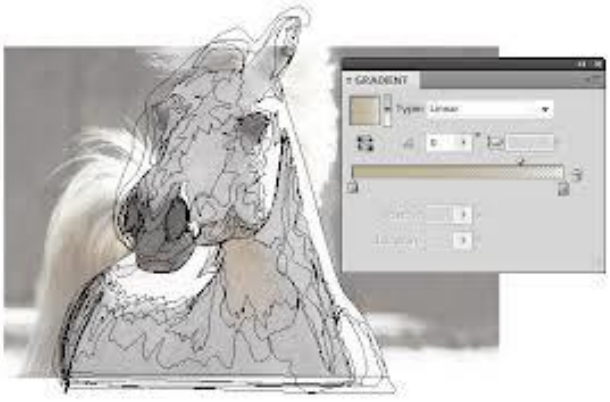
Depth cues: Linear perspective

Depth cues: Aerial perspective



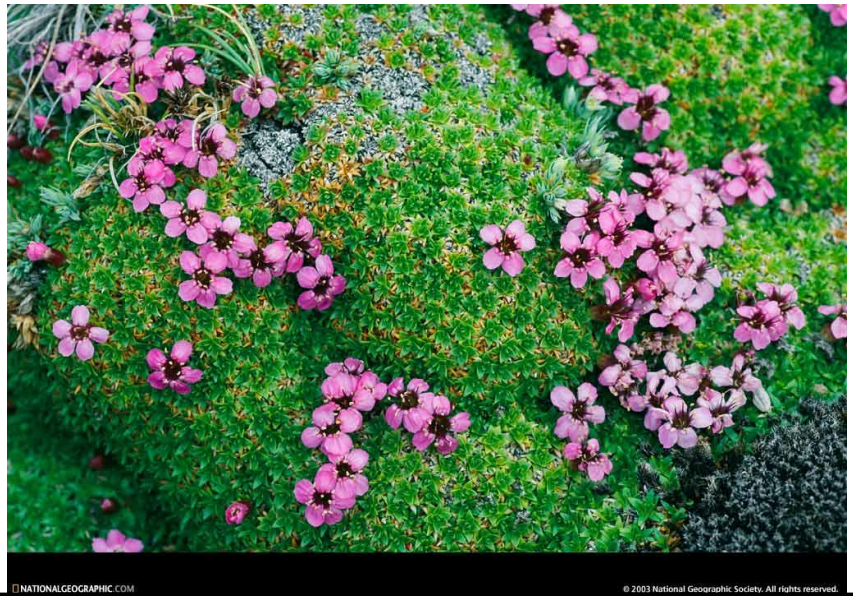
Shape cues: Texture gradient

Shape and lighting cues: Shading



Position and lighting cues: Cast

Grouping cues: Similarity (color, texture, proximity)



Grouping cues: “Common fate”

Bottom line

- Perception is an inherently ambiguous problem
  - Many
    
    given rise to a par

Bottom line

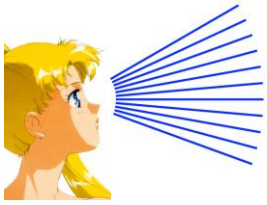
- Perception is an inherently ambiguous problem
  - Many different 3D scenes could have given rise to a particular 2D picture

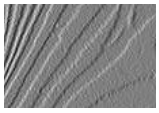
Possible solutions

  - Bring in more constraints (more images)
  - Use prior knowledge about the structure of the world
- Need both exact measurements and statistical inference!

I. Early vision

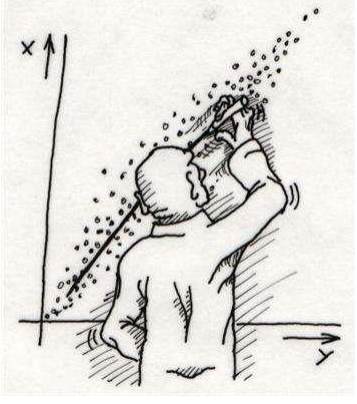
- Basic image formation and processing





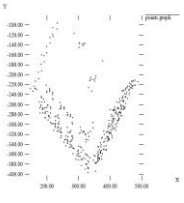
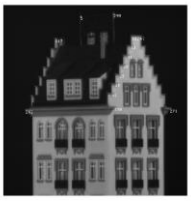
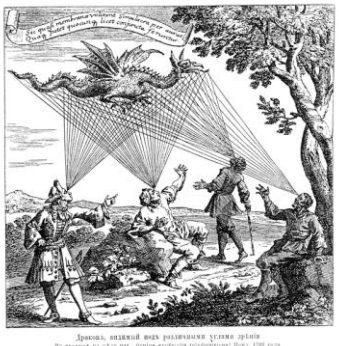
II. “Mid-level vision”

- Fitting and grouping

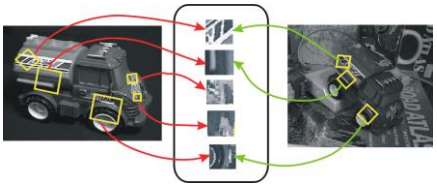
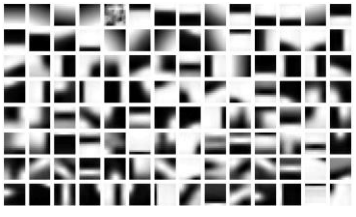
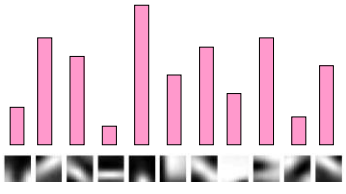
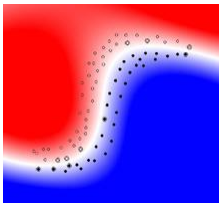


III. Multi-view geometry



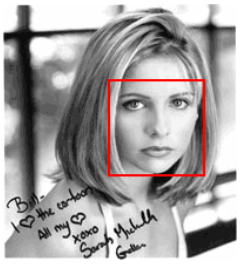




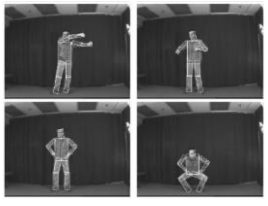
IV. Recognition

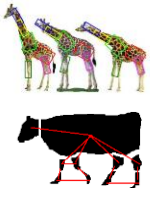





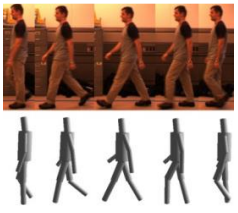
V. Advanced Topics

- Some in class other in term projects…









Course requirements

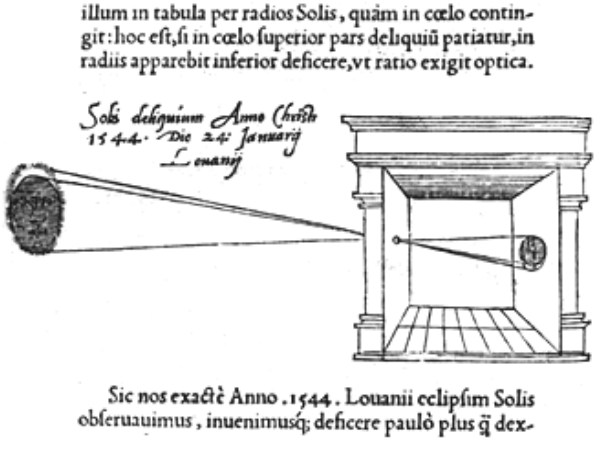
- Philosophy: computer vision is best experienced if hands-on
- Quiz & Programming assignments: 17%
  - Three or four Quizzes and assignments (Surprise Quiz)
  - Expect the first one in a couple of weeks
  - Brush up on your MATLAB/Python skills (see web/slate page for tutorial)
- Final project:
  - Putting several pieces together
  - List of options will be posted in the next few weeks (some great ideas can be find on web)
  - Expect to commit to a project idea by the end of Aug/September
- Participation: ?%
  - Ask questions
  - Answer questions
  - Give me feedback: I'm learning too!

•Mid & Final Terms: 75%

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!

- Homework: MATLAB tutorial (self-study, not collected)
- Reading: cameras and image formation (WEB)


illum in tabula per radios Solis, quàm in cœlo contingit: hoc eſt, ſi in cœlo ſuperior pars deliquiũ patiatur, in radiis apparebit inferior deficere, vt ratio exigit optica.

Solis deliquium Anno Chriſti 1544. Die 24. Januarij Louanij

Sic nos exactè Anno .1544. Louanii eclipſim Solis obſeruauimus, inuenimusq; deficere paulò plus q̃ dex-