Computer Vision and Scene Analysis

CS-GY 6643 I, Spring 2018

6:00 PM - 8:30 PM, Tu, Location: JABS 474

Guido Gerig, CSE

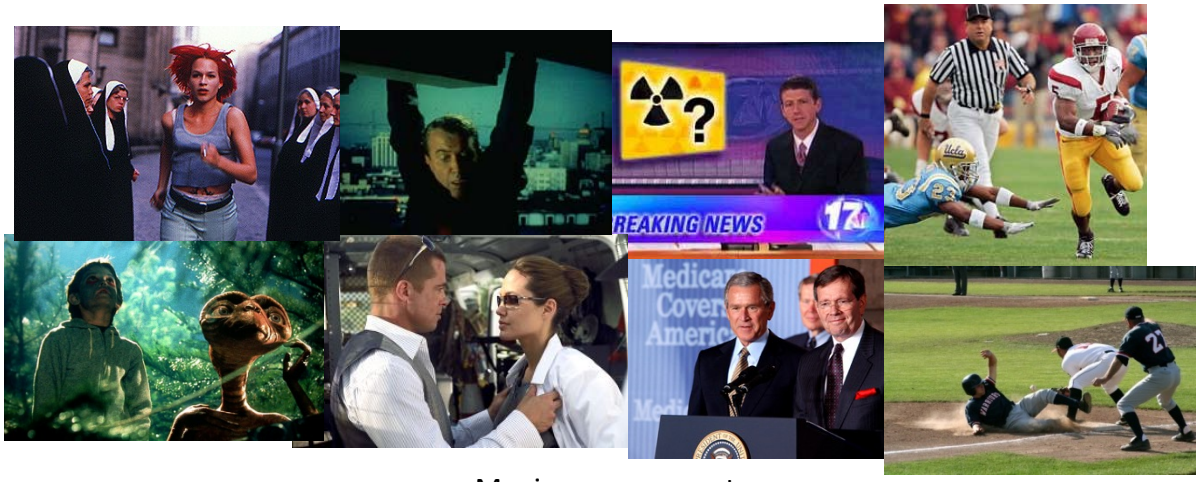
Main Goals

- to tell you what you can do with digital images
- to show you that developments in image analysis and computer vision can be fun and exciting
- to demonstrate that image processing is based on strong mathematical basic principles, applied to digital images via numerical schemes
- to demonstrate that you that you can solve typical image processing tasks on your own

Why vision?

- Images and video are everywhere!

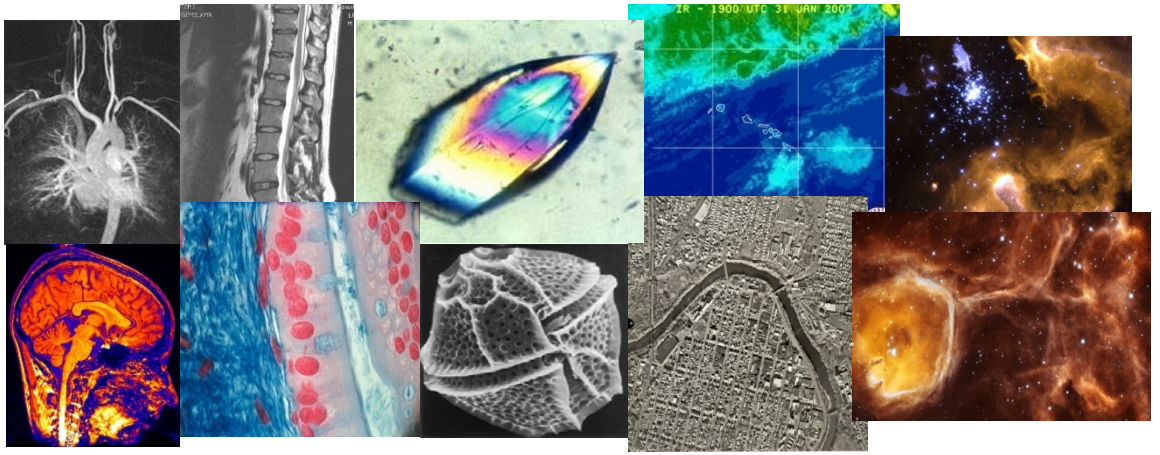
Personal photo albums



Movies, news, sports

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

Surveillance and security



Medical and scientific images

Slide credit; L. Lazebnik

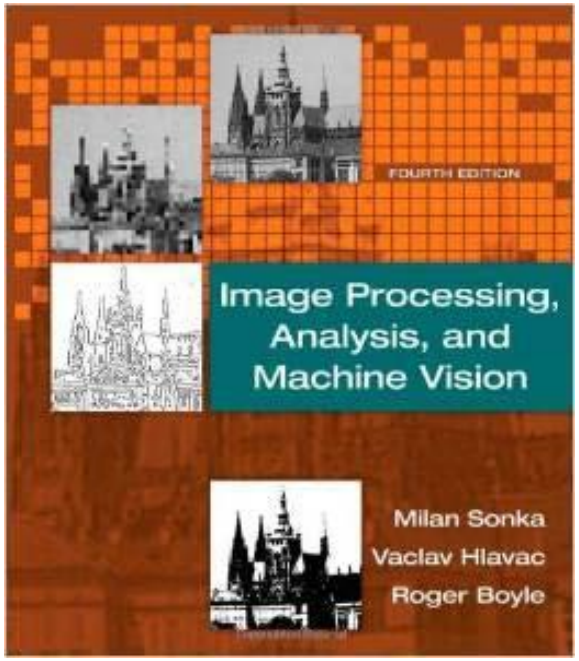
Administrivia

- Classes: Tue, 6 – 8.30pm
  Room JABS 474
- Instructor: Guido Gerig
  gerig@nyu.edu
- TA's: Zebin Xu, Monil Shah, Yida Zhou,
  Andrew Dempsey
- Prerequisites: Graduate standing, CS 5403 (Data Structures) or equivalent, proficiency in programming and familiarity with matrix arithmetic.
- Organization:

  Admin/Grading/Uploads: **NYU**Classes

Textbook

- **Required textbook:** Milan Sonka, Vaclav Hlavac and Roger Boyle, *Image Processing, Analysis, and Machine Vision*, *4th Edition*, Cengage Learning, Stamford, CT, 2015. (ebook, paperback, hardcover)



Briteclass Lecture Recording

briteCLASS

- Lectures will be recorded automatically using a system called Briteclass
- Recorded lectures will be available for viewing online, generally within 12 hours.

You may access the recordings directly from NYU Classes. For best result, use a Chrome or Firefox Browser. Other browsers won't work.

- Login to NYU Classes
- Click on your course tab
- You can access the "Briteclass Lecture Capture" recordings in two ways:
  - Click on the "BriteClass Lecture Capture" link on the Home page information section.
  - Click on the "Resources" link on the left menu bar and find the "BriteClass Lecture Capture" link to access.
- This will take you to the Briteclass dashboard where you can select the desired classroom recording based on date and time. 

TA / SW Tools / Office Hours

- **TA's:** Zebin Xu, Monil Shah, Yida Zhou, Andrew Dempsey
- **Software:** Matlab (free), Octave (free), Python
- Office Hours TA Hours/location: tbd
- Office Hours instructor: tbd

NYU Honor Code

Students are expected to work on their own, as instructed by the Professor. Students may discuss projects with other individuals either in the class or outside the class, but they may not receive code or results electronically from any source that is not documented in their report. Students must write their own code, conduct their own experiments, write their own reports, and take their own tests. Any use of sources (for projects or tests) that are not specifically given to the student by the Professor or TA, must be discussed with the Professor or TA or documented&cited in the report. Any student who is found to be violating this policy will be given a failing grade for the course and will be reported to the authorities as described in the University's Student Code of Conduct:

- NYU Tandon Student Code of Conduct
- Student Code of Conduct (PDF)

Grading - Weights

Course load:

- There will be about 5 to 6 handwritten homework assignments, plus two computer projects that require programming. You can use any high-level programming language to do the projects, but Python, C/C++, Java, or MatLab are the recommended languages.
- There will be two exams. The second exam will be held during the final exam week and will only cover materials after the first exam.

Grading - Weights

Grade Distribution:

- **Homework** **10%**
- **Programming Projects** **30%**
- **Exam I** **30%**
- **Exam II** **30%**

Some Basics

- It will be your responsibility to regularly read the Announcements on NYUClasses.
- We don't need a laptop for the class, please keep them closed !!!!!
- Please interact, ask questions, clarifications, input to instructor and TA.
- No need for cell phones …., you surely know.

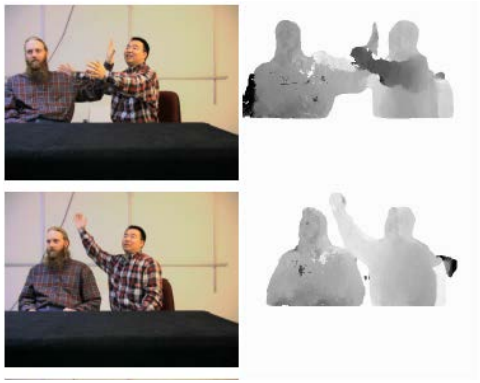
What is computer vision?

- Automatic understanding of images and video
  - Computing properties of the 3D world from visual data *(measurement)*
  - Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. *(perception and interpretation)*

Courtesy Prof. Trevor Darrell trevor@eecs.berkeley.edu

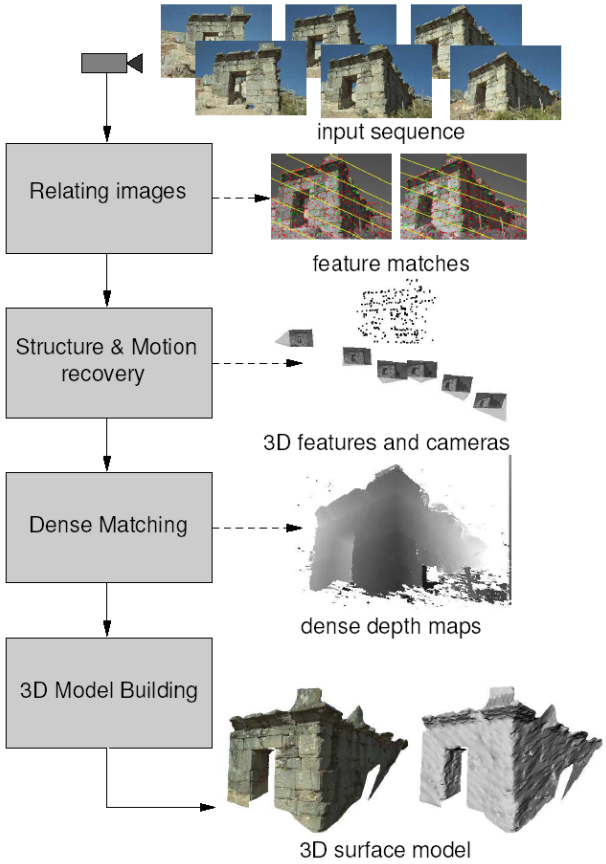
Vision for measurement

Real-time stereo

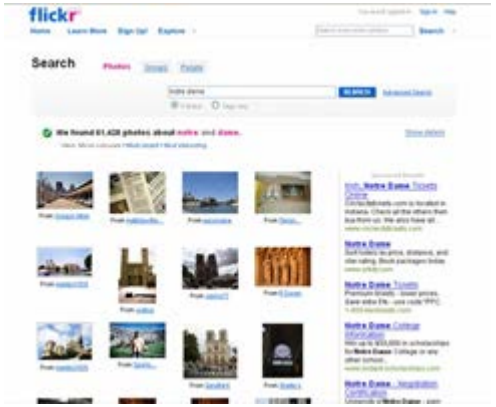


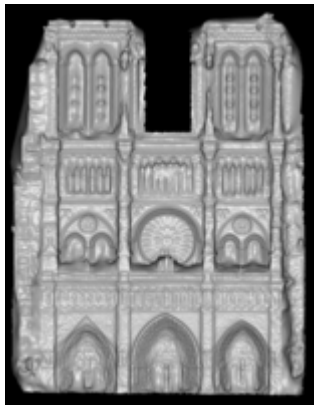
Pollefeys et al.

Structure from motion



Multi-view stereo for community photo collections



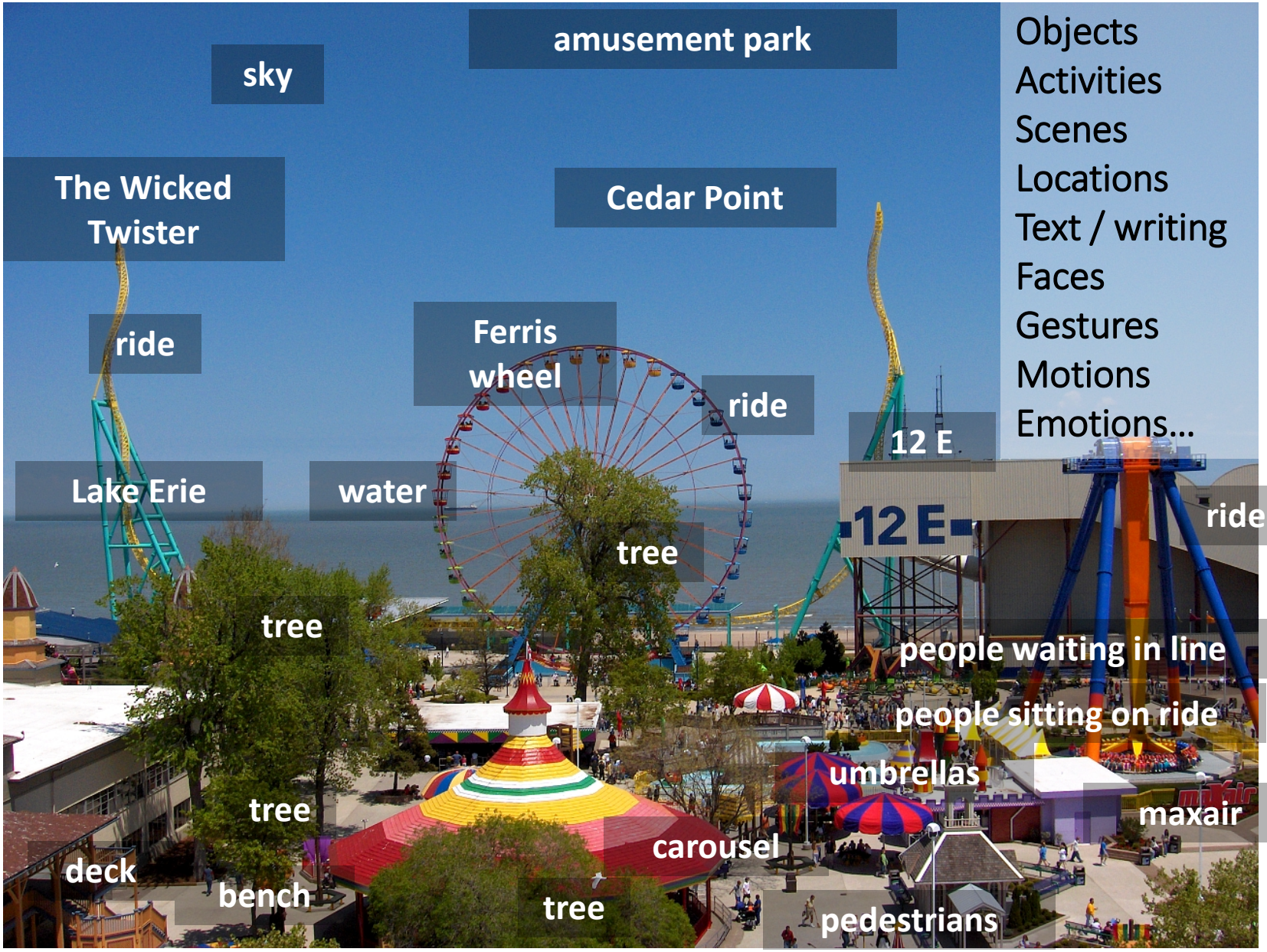


Goesele et al.

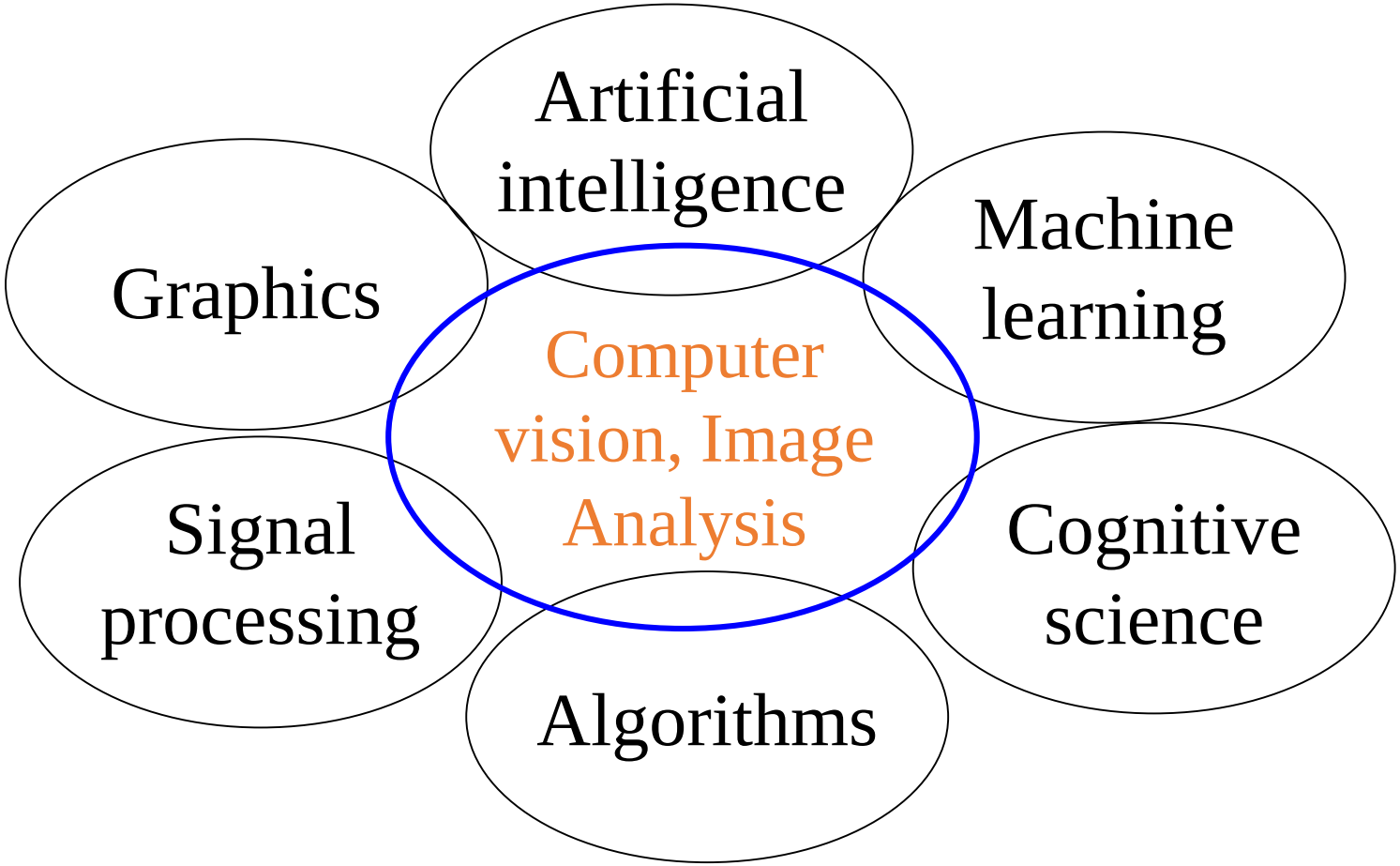
Courtesy Prof. Trevor Darrell trevor@eecs.berkeley.edu

Slide credit: L. Lazebnik

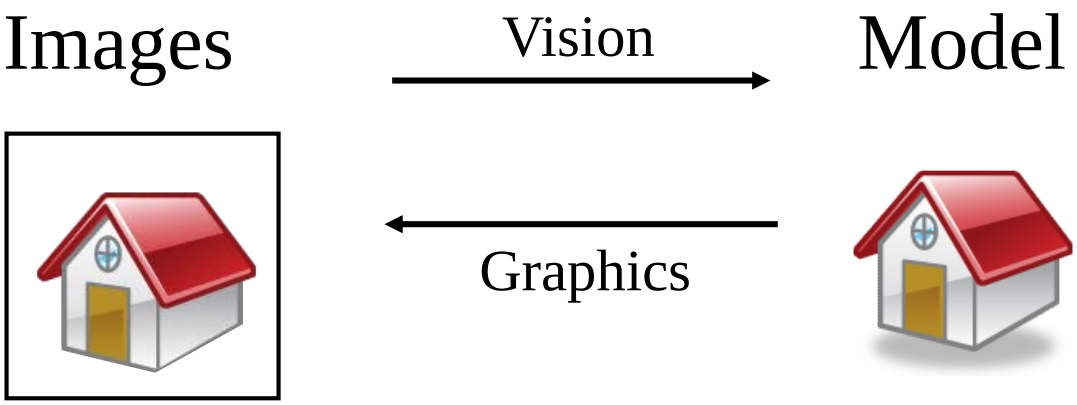
Vision for perception, interpretation



Related disciplines



Computer Vision and Graphics



Inverse problems: analysis and synthesis.

Why vision?

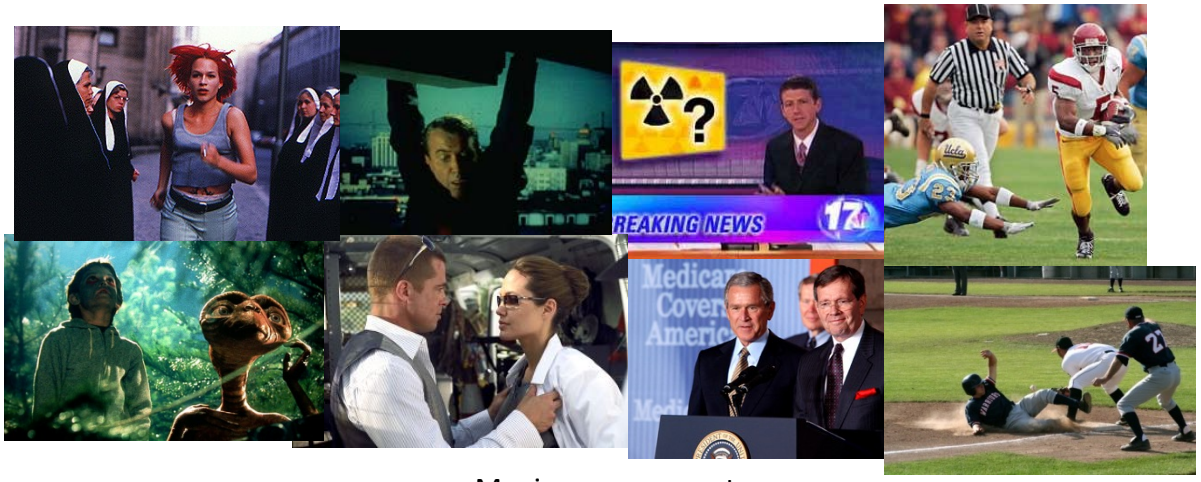
- As image sources multiply, so do applications
  - Relieve humans of boring, easy tasks
  - Enhance human abilities: human-computer interaction, visualization
  - Perception for robotics / autonomous agents
  - Organize and give access to visual content

Courtesy Prof. Trevor Darrell trevor@eecs.berkeley.edu

Why vision?

- Images and video are everywhere!

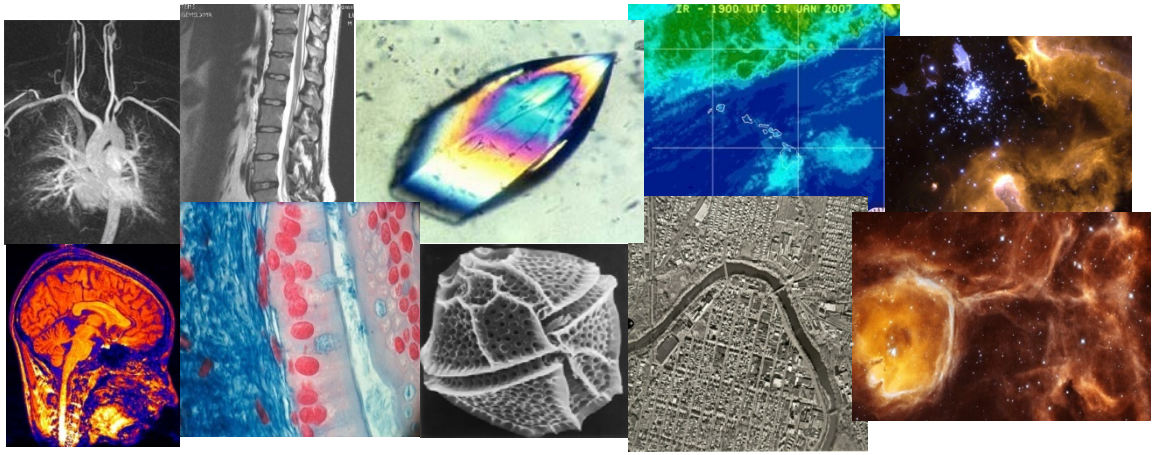
Personal photo albums



Movies, news, sports

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

Surveillance and security



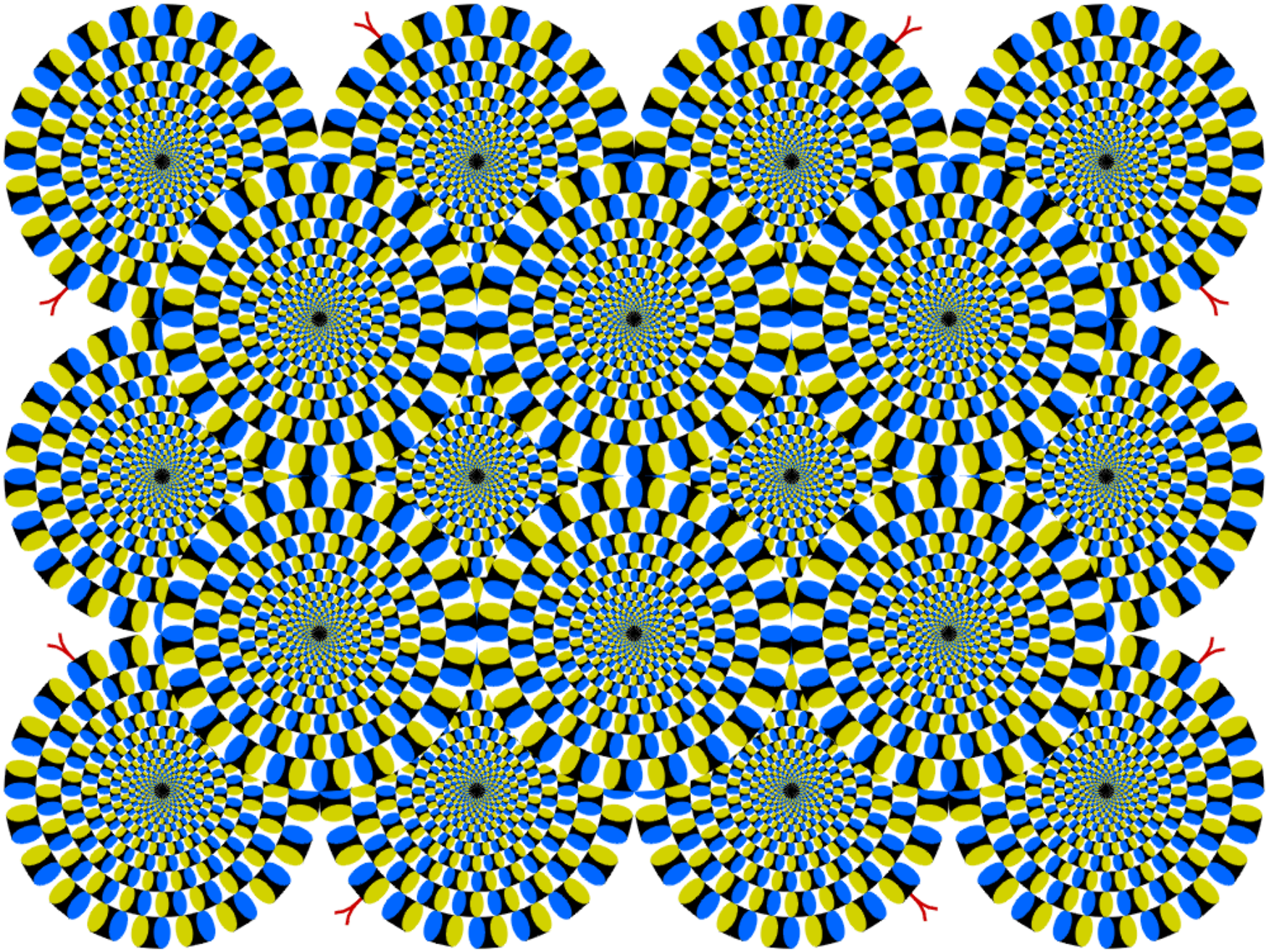
Medical and scientific images

Slide credit; L. Lazebnik

Again, what is computer vision?

- Mathematics of geometry of image formation?
- Statistics of the natural world?
- Models for neuroscience?
- Engineering methods for matching images?
- Science Fiction?

Courtesy Prof. Trevor Darrell trevor@eecs.berkeley.edu

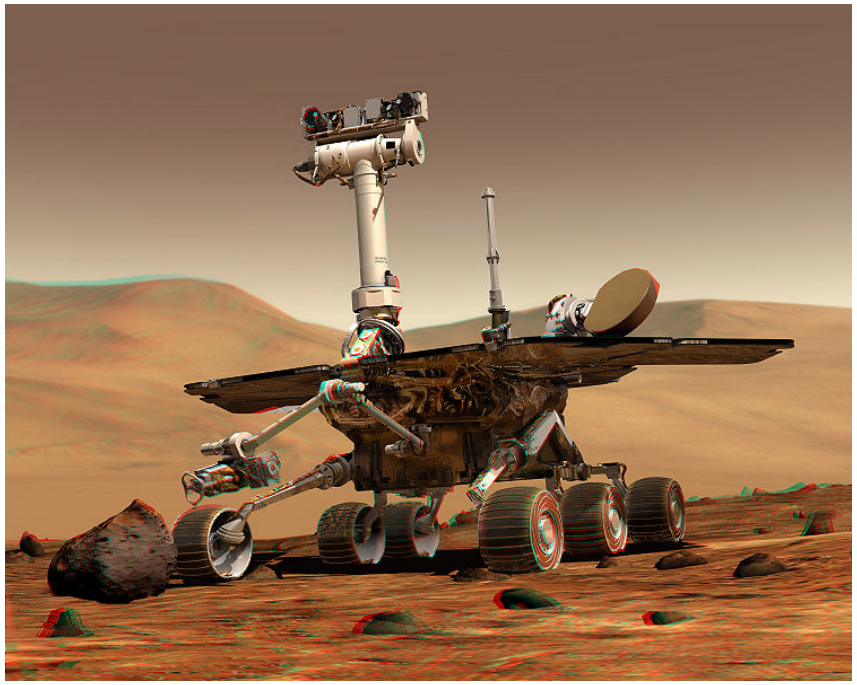


Current state of the art

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

Courtesy Prof. Trevor Darrell trevor@eecs.berkeley.edu

Robotics

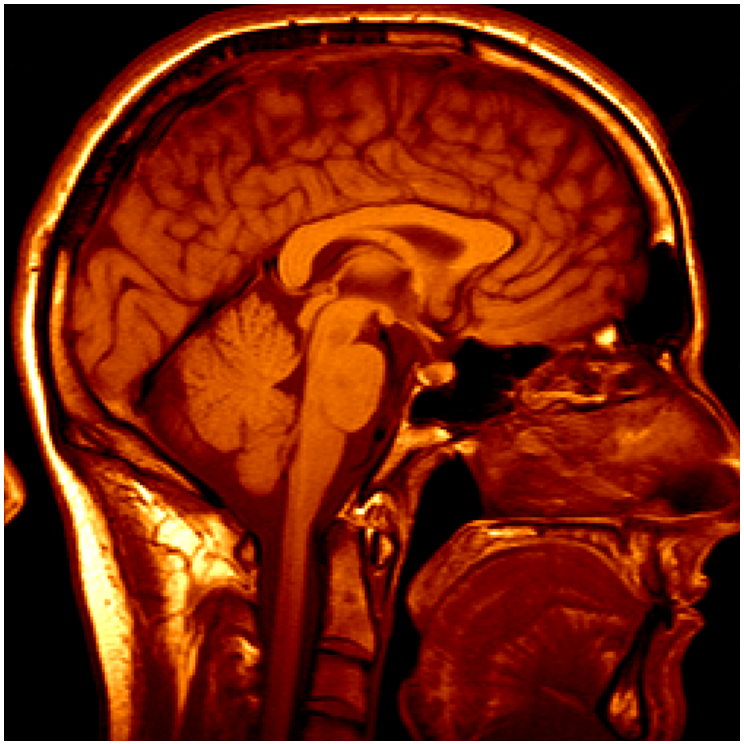


NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

http://www.robocup.org/

Courtesy Prof. Trevor Darrell trevor@eecs.berkeley.edu

Medical imaging



3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT

Courtesy Prof. Trevor Darrell trevor@eecs.berkeley.edu

Smart cars

- Mobileye
  - Vision systems currently in high-end BMW,
  - By 2010: 70% of car manufacturers.
  - Video demo
  - YouTube, TestMovie

Slide content courtesy of Amnon Shashua

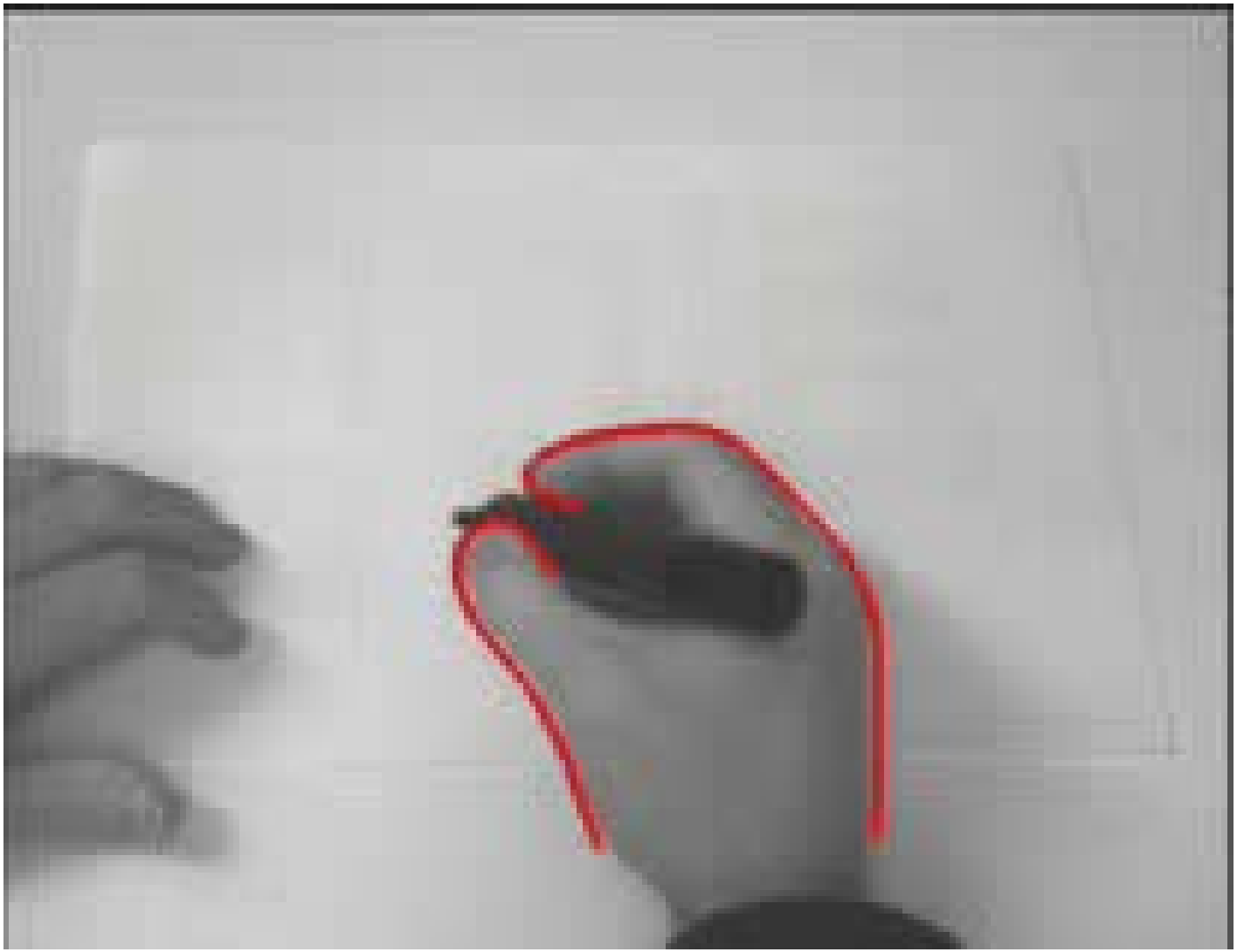
Archaeology

accuracy ~1/500 from DV video
(i.e. 140kb jpegs 576x720)

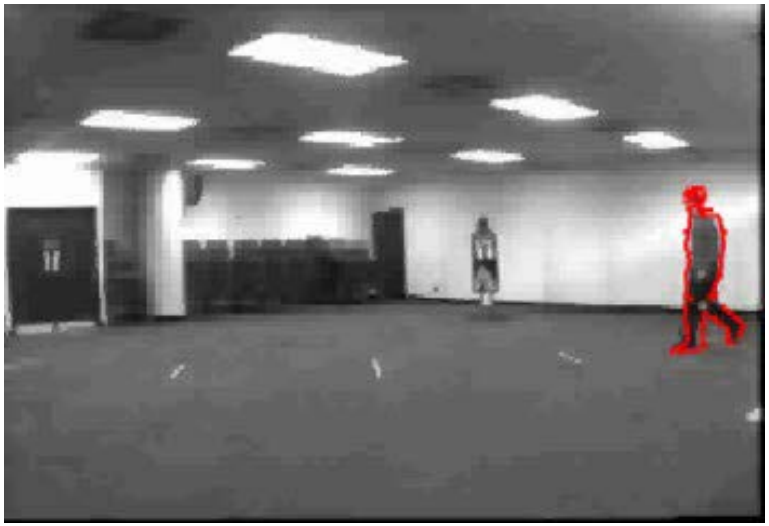
link

Object Tracking

Object Tracking: Using Deformable Models in Vision



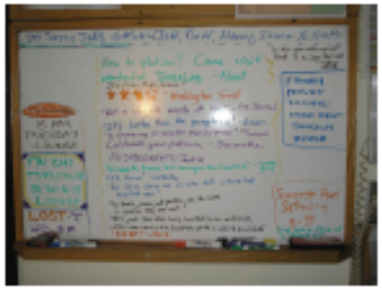
Object Tracking

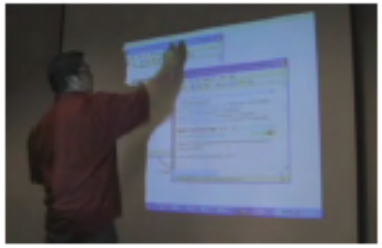


Computer Vision Systems

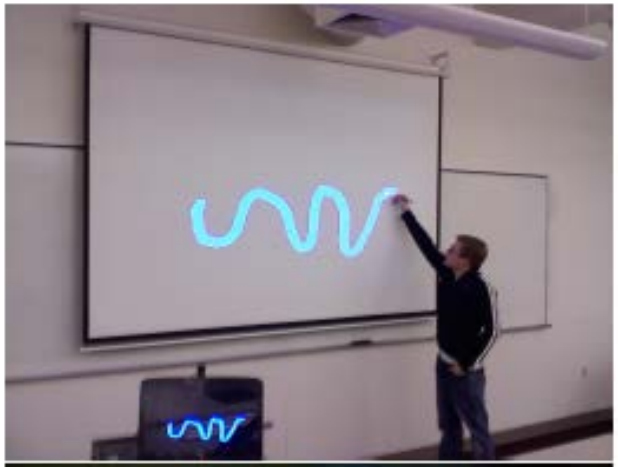
Webcam Based Virtual Whiteboard
Student Project, Jon Bronson James Fishbaugh

- Blackboards came first
- Whiteboards eventually followed
- Virtual Whiteboards are coming
- Basic Idea:
  - Write on any surface
  - Use no ink/chalk
  - Store all information to disk

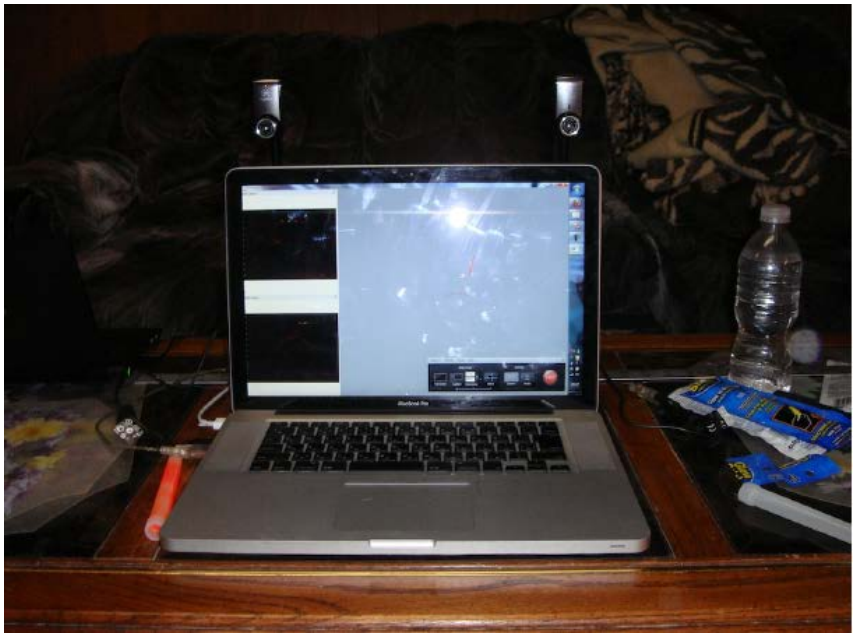




Webcam Based Virtual Whiteboard
Jon Bronson James Fishbaugh



Real-Time 3D Glowstick Detection
Computer Vision Project 2009
Andrei Ostanin



Detecting the 3D position of glowsticks in real-time using two cameras.

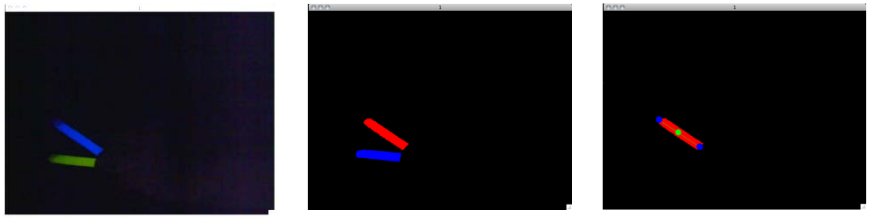
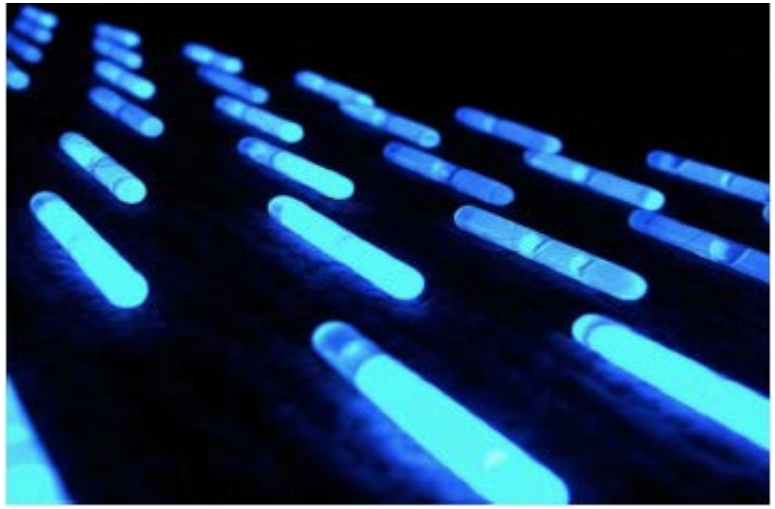
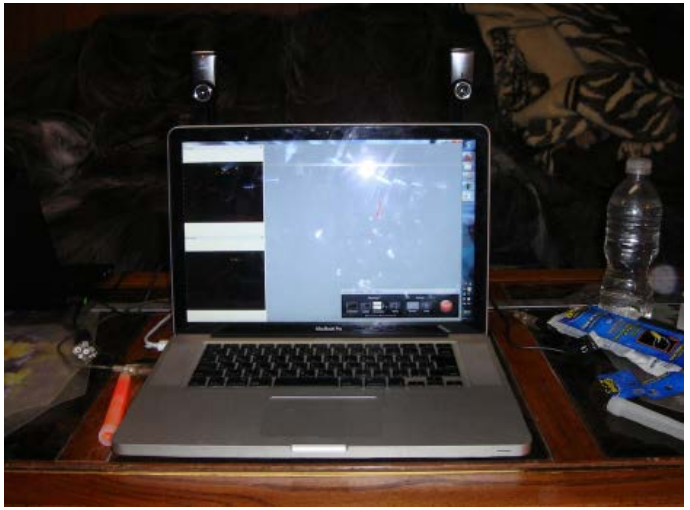


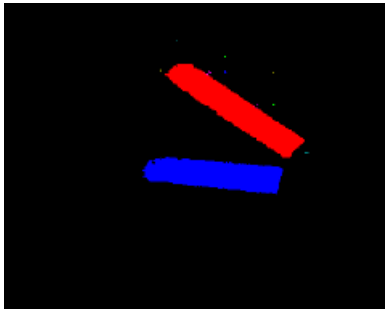
Figure 2: Camera input images.

Realtime Glowstick Detection
Andrei Ostanin





- Capture the 3D position of glowsticks in real-time using two webcams
- Environment dark enough that glowsticks are easily segmented out
- Prefer speed over correctness

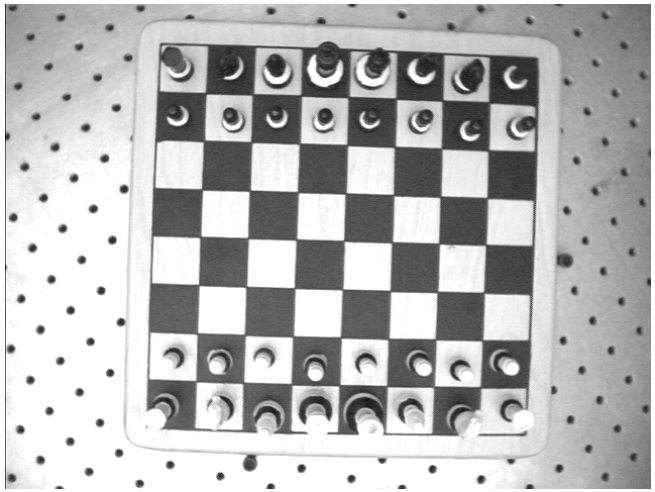


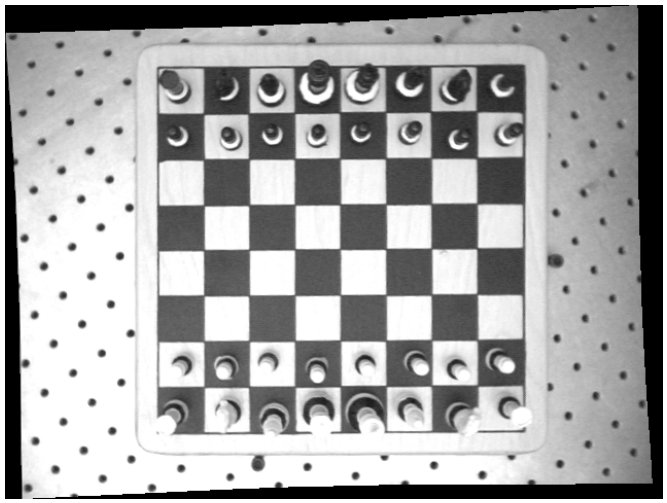
movie

Student Project: Playing Chess, Recognition and Simulation

- Track individual chess pieces
- Maintain state of board
- Graphically represent state changes and state
- D. Allen, D. McLaurin UNC
- Major ideas:
  - 3D from stereo
  - detect and describe changes
  - Use world knowledge (chess)

Calibration, Rendering & Replay





Movie