

CSC 589 Introduction to Computer Vision

Instructor: Bei Xiao
American University
Monday, January 12th

Today's class

- Introduction
- What is computer vision?
- Why vision is hard?
- Applications of computer vision
- What we will learn in this course
- OpenCV challenges

What is computer vision?

- The major goal of computer vision is the **understanding** of images and videos.



What kind of scene?

What time of the day
it was taken?

How many cars?

How far is the store?

Are there any
person in the
scene?

What is computer vision?

- The major goal of computer vision is the **understanding** of images and videos.



What kind of scene?

What time of the day
it was taken?

How many cars?

How far is the store?

Are there any
person in the
scene?



Slide courtesy from Noah Snavelly



sky

building

flag

face

banner

wall

street lamp

bus

bus

cars

slide credit: Fei-Fei, Fergus & Torralba

Slide courtesy from Noah Snively

What is computer vision?



What kind of food it is?

What kind of meat?

Are there any vegetables in the image?

What kind of cuisine is it?

What are the most useful utensils to use to eat the food?
Knife and fork?

Goal of computer vision

To describe the world that we see in one or more images and to reconstruct its properties, such as shape, illumination, and color distributions, and most of all, its content.

Why vision is important?

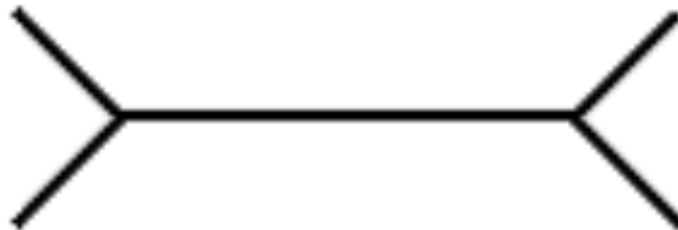
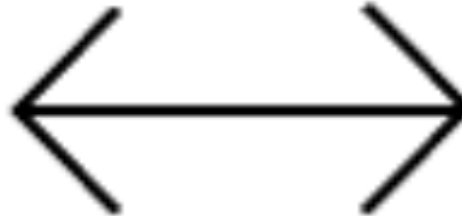
Visual cortex occupies about 50% of Macaque brain.

More human brain devoted to vision than anything else.

Visual information is essential in robotic system.

Large percentage of effort in neurobiology is dedicated to the visual system.

Vision is hard



Muller-Lyer illusion

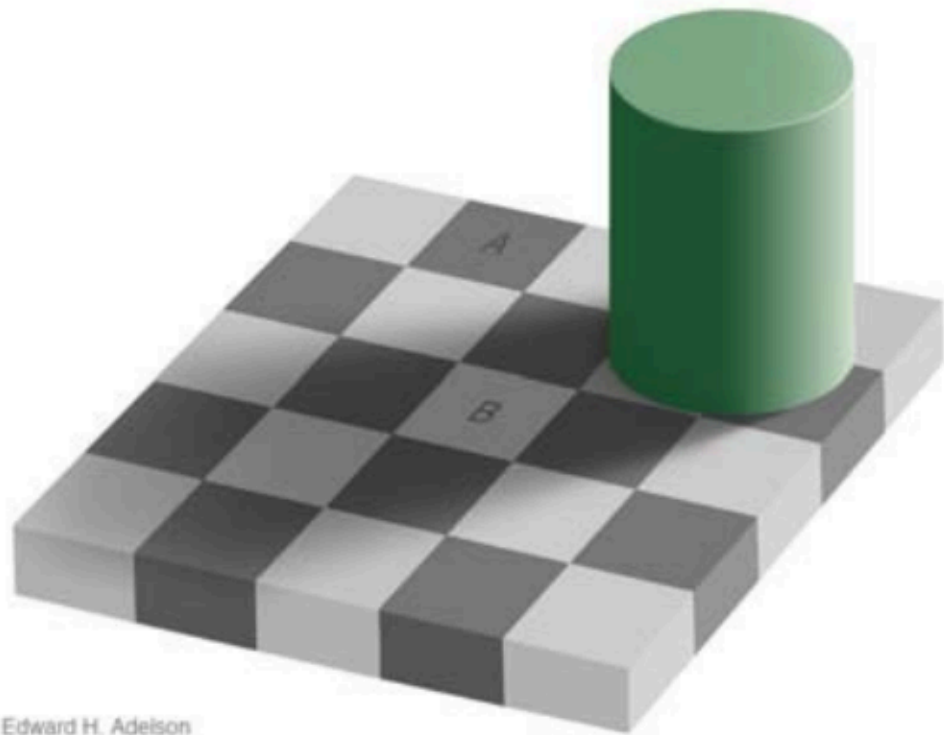
Vision is hard

What is the most likely scenario?



Prior knowledge: Occlusion is more common than L shaped object

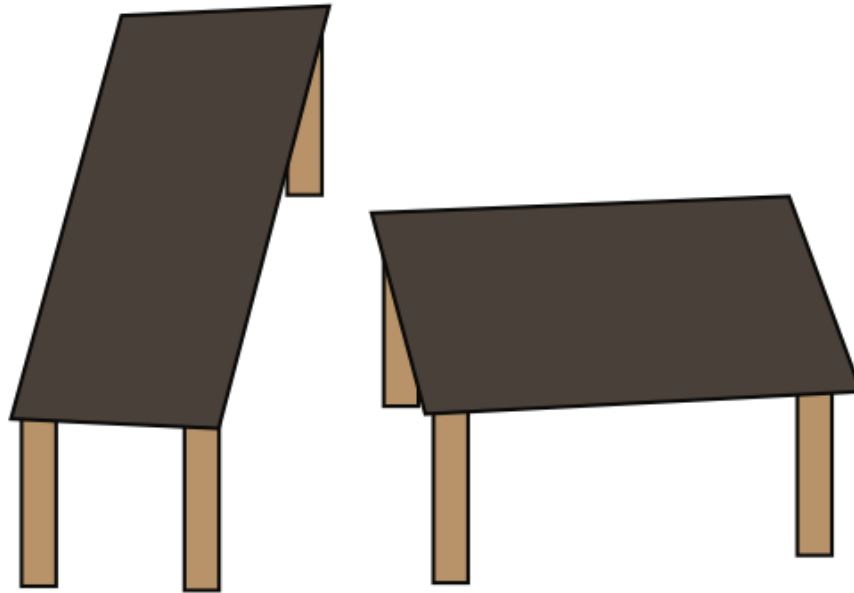
Vision is hard



Edward H. Adelson

Edward Adelson

Vision is hard



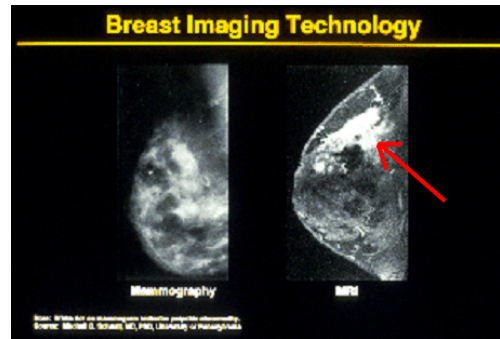
Roger Shepard

<http://www.opticalillusion.net/optical-illusions/shepards-tables-whats-up/>

Computer Vision Applications



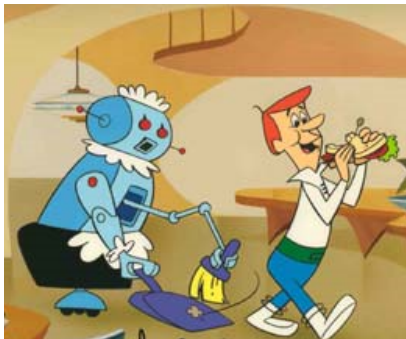
Safety



Health



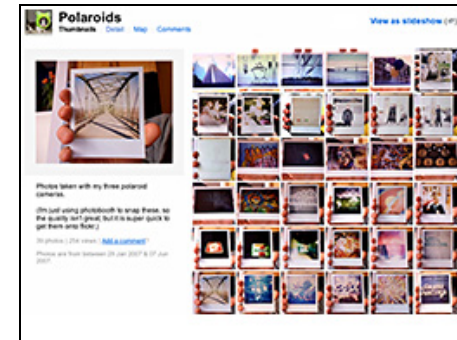
Security



Comfort



Fun



Access

Image-based search

Query image



Visually similar images



Report images

Label: Residenzplatz, Salzburg, Austria

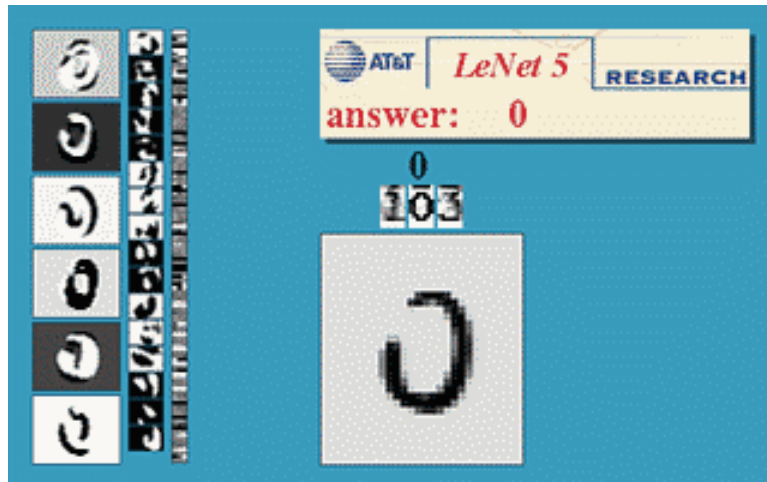
Usually compares image similarity distance such as similarity in color, texture, shape, etc.

Read about [content-based image retrieval](#)

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

Face Detection



Face Detection function keeps subjects' faces in **sharp focus**

Object recognition (in supermarkets)

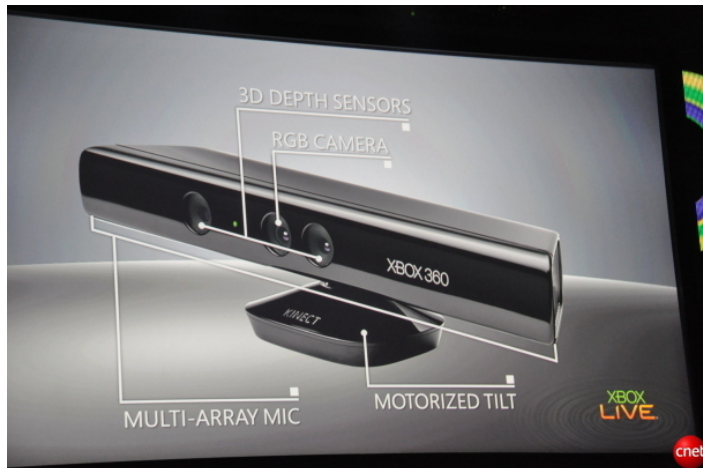


[LaneHawk by EvolutionRobotics](#)

“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... “

Slide from James Hayes

Motion Sensing: Kinect



Robot : <https://www.youtube.com/watch?v=w8BmgtMKFbY>

Object recognition:

<https://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o>

Gesture Recognition



http://www.webopedia.com/TERM/G/gesture_recognition.html

Special effects: motion capture



Dawn Of The Planet of The Apes, Director talks about the visual effects
<https://www.youtube.com/watch?v=hIWYaePmAYM>

Google Cars

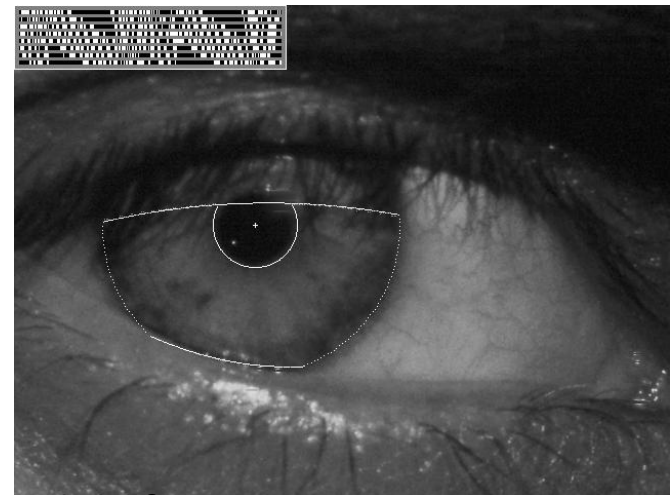
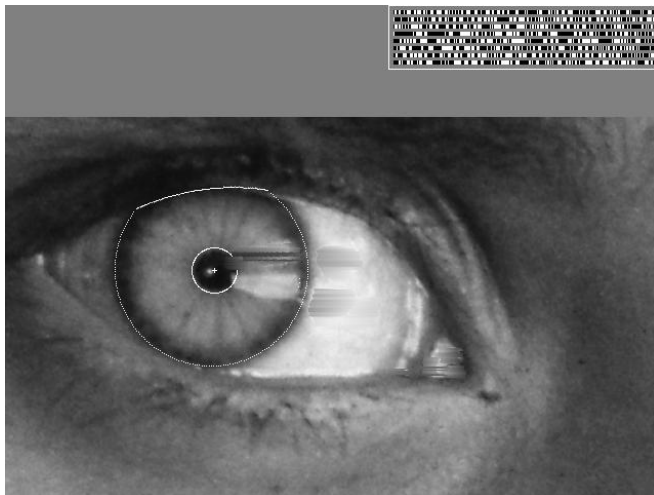


The range finder mounted on the top is a Velodyne 64-beam laser. This laser allows the vehicle to generate a detailed 3D map of its environment.

Vision-based biometrics



“How the Afghan Girl was Identified by Her Iris Patterns” Read the [story](#)
[wikipedia](#)



Slide from James Hayes

Mobile Eye for driving safety

Pedestrian Collision Warning



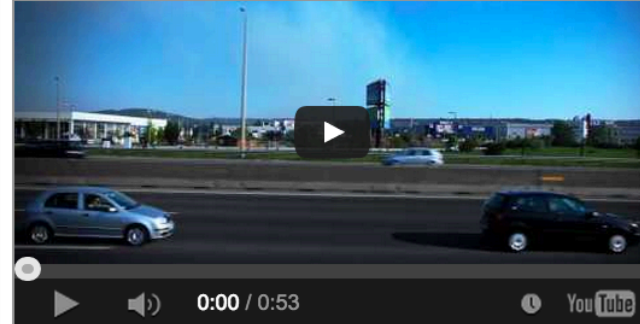
Lane Departure



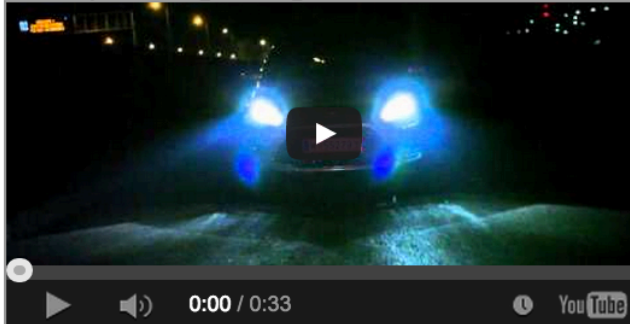
Forward Collision Warning



Headway Monitoring



Mobileye Intelligent High-Beam



Speed Limit Indication



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.

Slide from James Hayes

Automatic Image and Video Editing

Data driven approach of image editing

Given a single
image at day



Input image at "blue hour" (just after sunset)

Database of
time-lapse
videos



A database of time-lapse videos

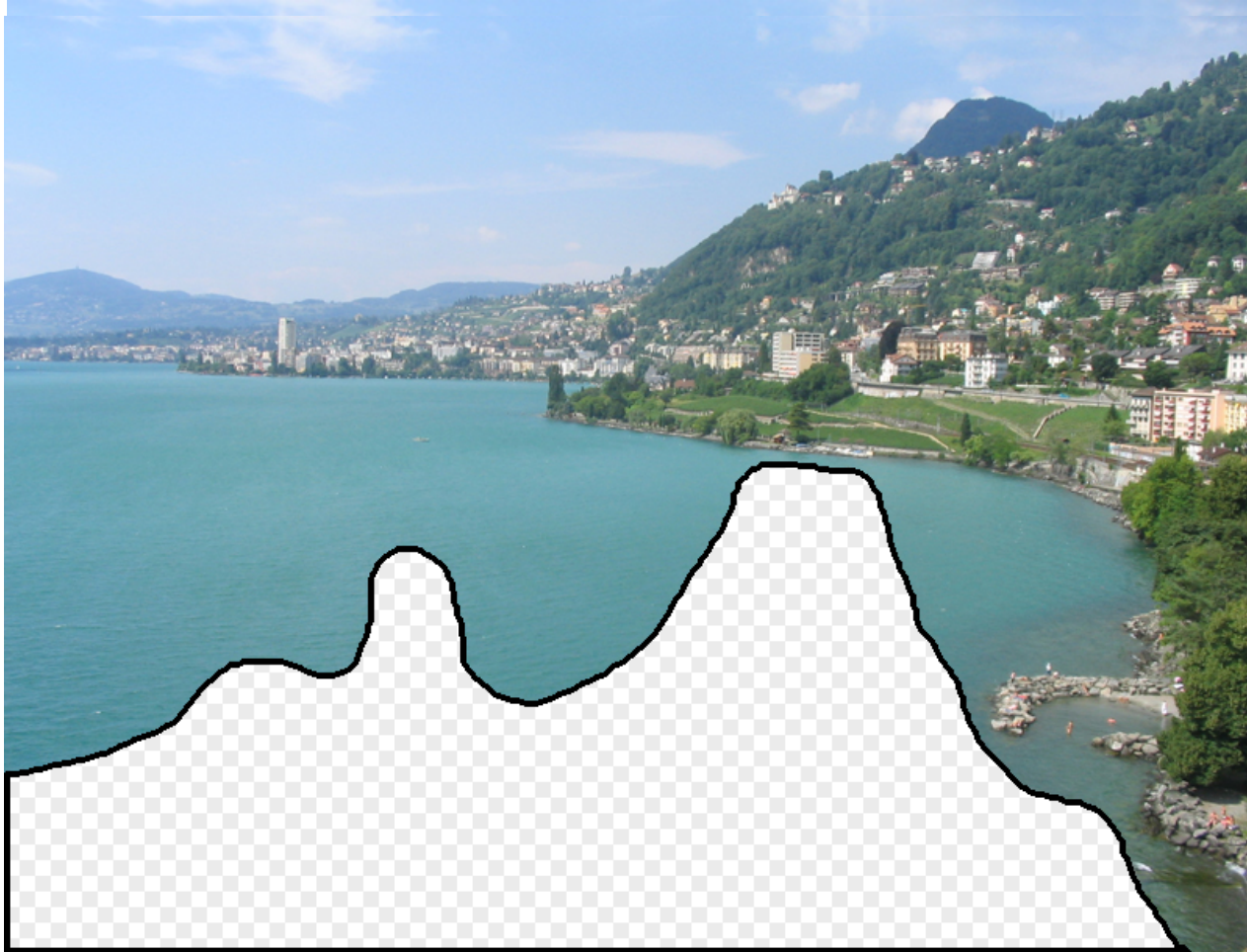
A hallucination of the same
scene at night



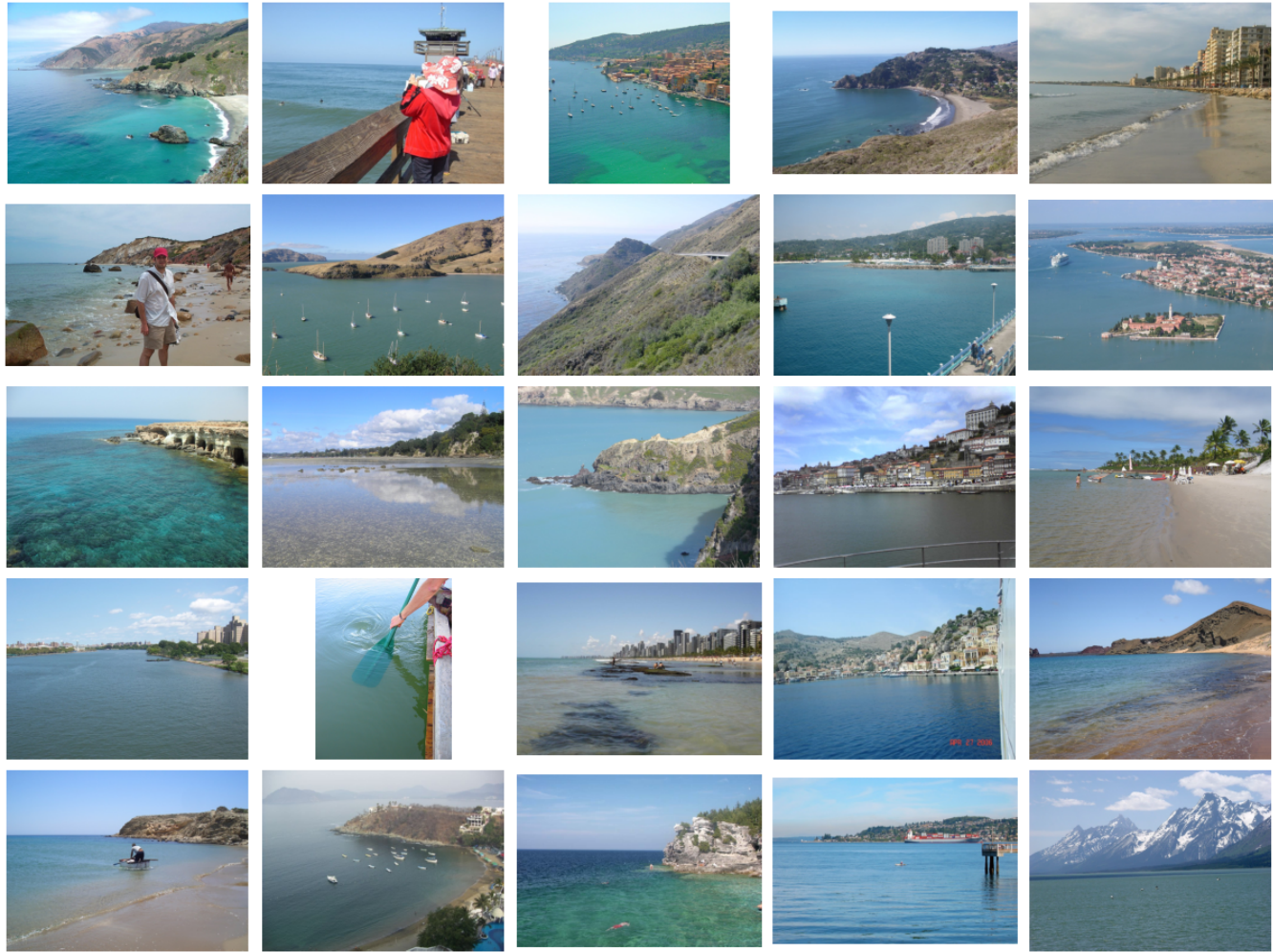
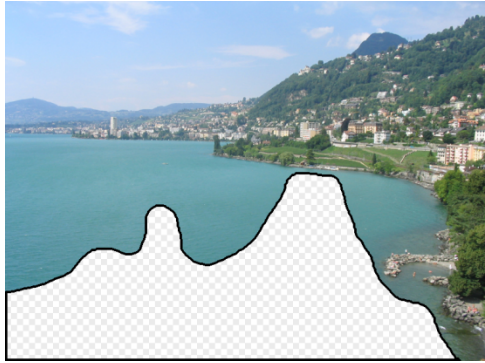
Hallucinate at night

YiChang Shi, Sylvain Pairs, Fredo Durand, and William
T Freeman, SIGGRAPH ASIA 2013

Scene Completion



[Hays and Efros. Scene Completion Using Millions of Photographs.
SIGGRAPH 2007 and CACM October 2008.]



Nearest neighbor scenes from database of 2.3 million photos



Graph cut + Poisson blending

Image Forensics



From Hany Farid, Digital Image Forensics
<http://www.cs.dartmouth.edu/farid/downloads/publications/sciam08.pdf>

Image Forensics

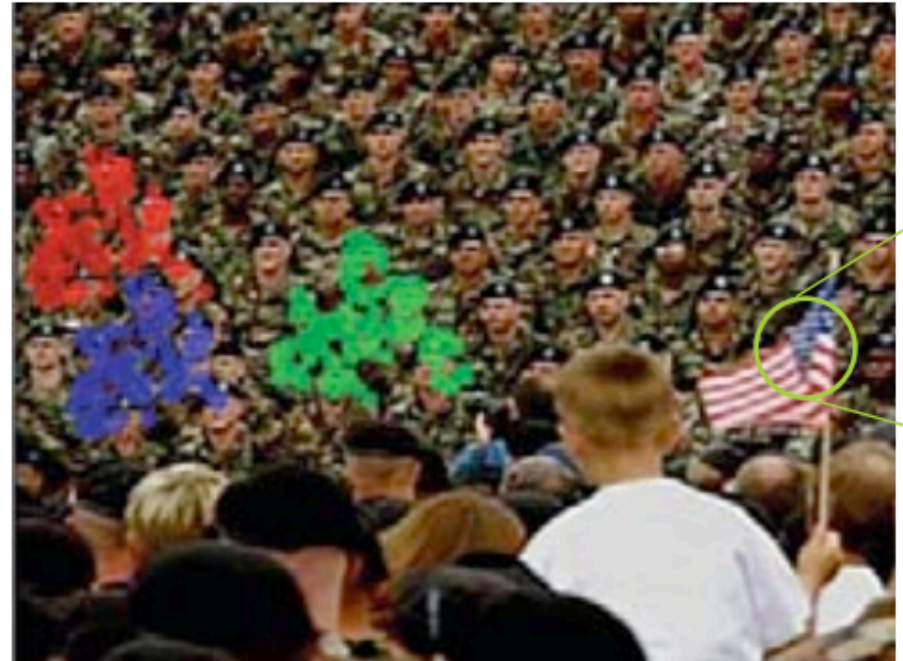


Arrows
indicates
light source
direction

1. They are not shot Together.
2. The girl's helmet is the same as the man's but recolored.

From Hany Farid, Digital Image Forensics
<http://www.cs.dartmouth.edu/farid/downloads/publications/sciam08.pdf>

Applications of Computer Vision: Detecting cloning



From Hany Farid, Digital Image Forensics

<http://www.cs.dartmouth.edu/farid/downloads/publications/sciam08.pdf>

Computer Vision Companies

- CV is a very active field. Most of new applications we just saw are within 5 years old.
- Here is a overview of CV companies:
- <http://www.cs.ubc.ca/~lowe/vision.html>

Can the computer match human perception?

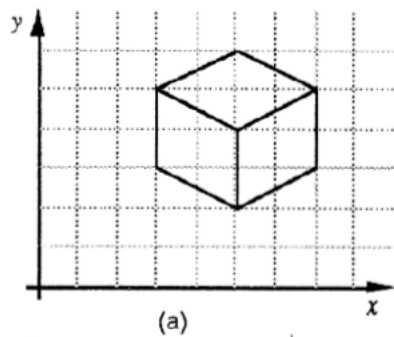
- What does human vision have:
 - Amazing eyes (sensor) and big brain
 - Experiences in the real-world (Prior knowledge)
 - Good at solving hard problems with ambiguous information
- What does computer vision have?
 - Lots of CPUs, fast speed
 - Mathematical techniques
 - Data, data, data ((could we learn from constraints from data)
 - Fast at solving simple and tedious problems
 - But tremendous progress has been made of solving hard problem such as human activity understanding

Why vision is hard?

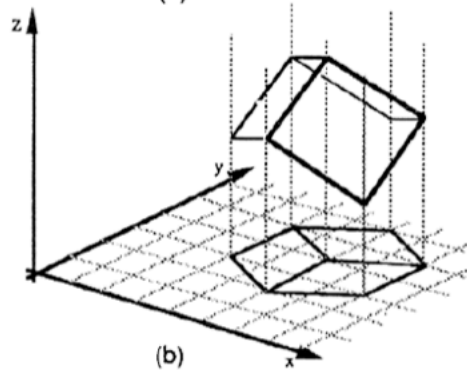
- We must estimate 3D properties (texture, color, surface properties) of a visual scene from 2D images?
- It is an inverse problem under-constrained.

Vision is an inverse problem

Why is this hard?



Construct 3D from 2D images



Vision is an inverse problem

Why is this hard?

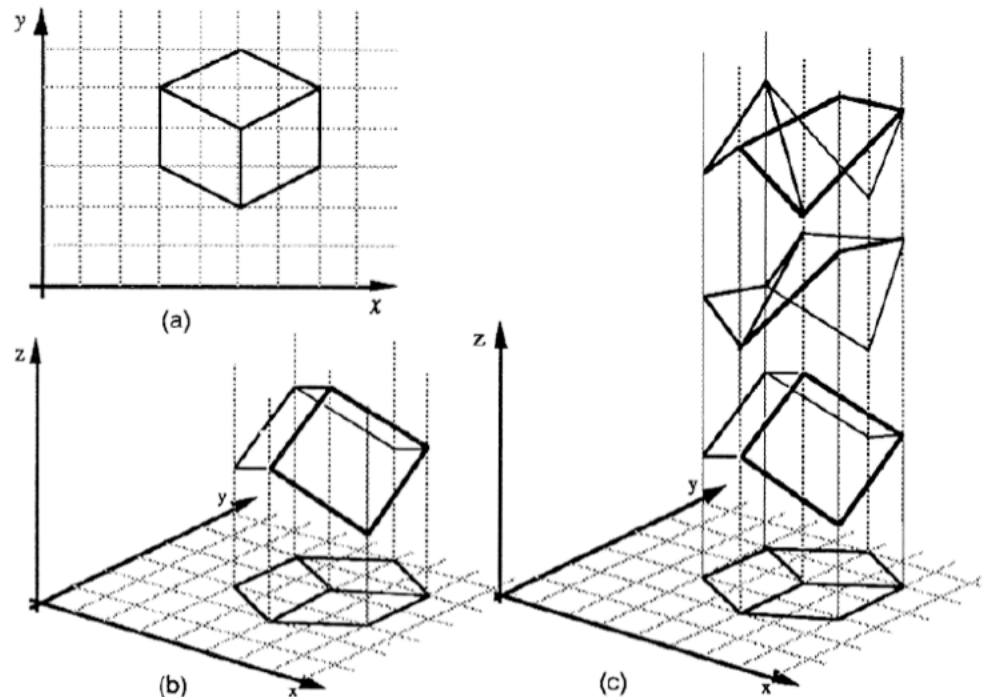


Figure 1. (a) A line drawing provides information only about the x, y coordinates of points lying along the object contours. (b) The human visual system is usually able to reconstruct an object in three dimensions given only a single 2D projection (c) Any planar line-drawing is geometrically consistent with infinitely many 3D structures.

Vision is hard

Some things have strong variations
in appearance



Why is computer vision difficult?



Viewpoint variation



Illumination



Scale

Slide courtesy from Noah Snavely

Why is computer vision difficult?



Intra-class variation



Motion (Source: S. Lazebnik)



Background clutter



Occlusion

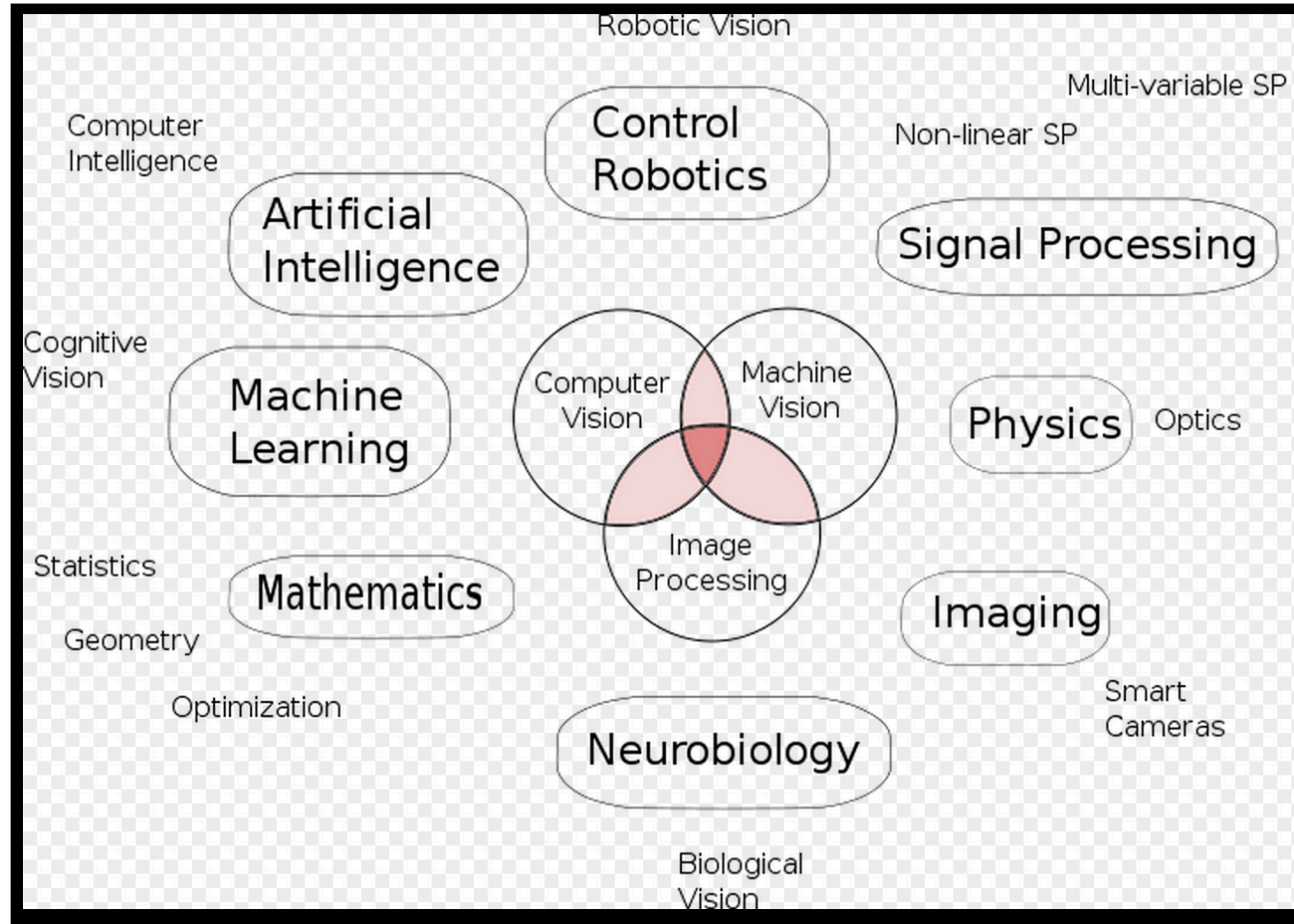
But what do we have?

- Lots and lots and lots of data.
- We can learn from humans.
- We have Prior knowledge, which can provide constraints of the problem.

Computer Vision and Nearby Fields

- Computer Graphics: Models to Images
- Computational Photography: Images to Images
- Computer Vision: Images to Models
- Human perception: understand how humans perceive the world
- Machine learning (tools that are common in all above areas): algorithm of how to learn from data

Computer Vision and Nearby Fields



What we are going to learn?

1. Low-level vision and image formation.
 - Image processing, camera models, edges.
2. Mid-level features. Geometry and algorithm.
 - motion, stereo, multiple view geometry.
3. High-level Recognition and categorization
 - face detection, object recognition,
4. Light, color, and reflectance.
5. Advanced topics

General description

The course is extremely hands-on and requires working knowledge of Python or MATLAB. CSC 280 is recommended. If you haven't taken CSC 280, please talk to me after class.

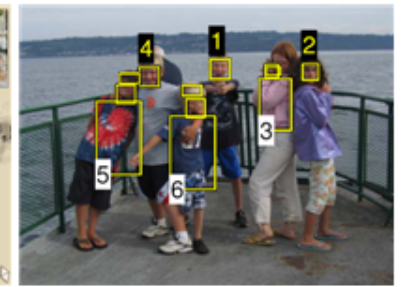
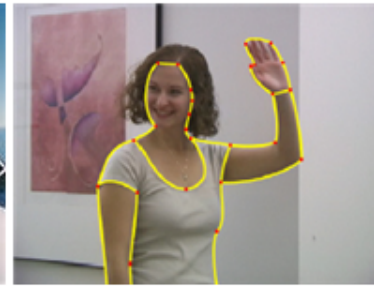
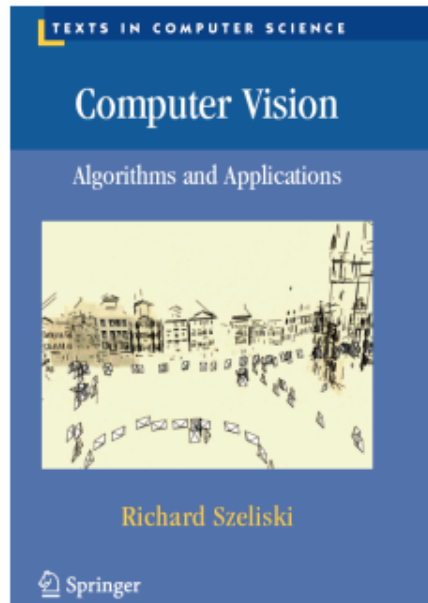
Class time is often used for group discussions and quiz. Attendance is **STRONGLY** recommended. Absence for 3 times without medical/ family illness proof will receive **0%** attendance score.

Textbooks are **ONLY** supplementary to lectures. Content can change depending on progress. Assigned reading could be papers.

Textbook

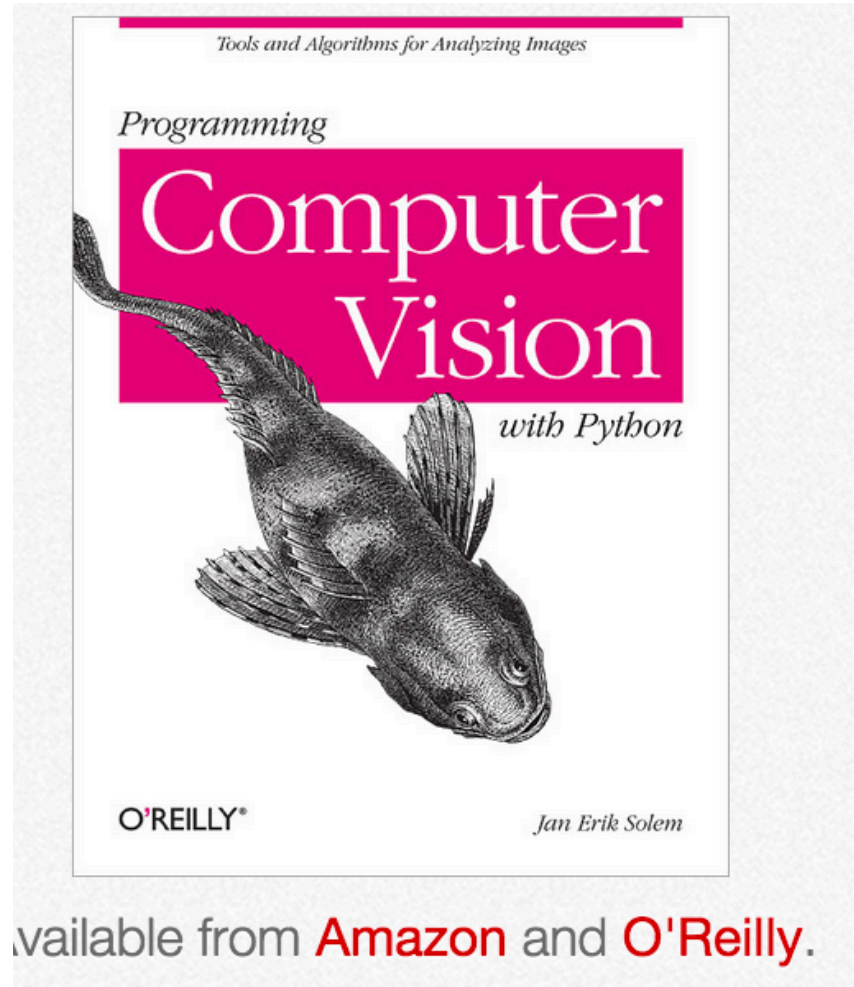
Computer Vision: Algorithms and Applications

© 2010 [Richard Szeliski](#), Microsoft Research



<http://szeliski.org/Book/>

Textbook



available from **Amazon** and **O'Reilly**.

<http://programmingcomputervision.com/>

OpenCV-Python tutorial

http://docs.opencv.org/trunk/doc/py_tutorials/py_tutorials.html

Syllabus

Grading:

60 % Homework assignments.

15 % Mid-term exam

20% Final project

5 % Attendance, including quiz.

Syllabus

Homework late policy:

- 5 free late days for homework only (**not final project**)– use them in your ways;
- After you have finished using all of your late days, there is a 25% off per day late penalty;
- No submission is accepted after 3 late days per PS (zero points).

Course Website:

<http://nw08.american.edu/~bxiao/CSC589/CSC589.html>

Prerequisites

- **Linear algebra**, basic calculus, and probability
- Experience with image processing and Python or MATLAB are recommended.

Academic Integrity

You must not copy lines of code from other people unless teamwork is allowed. Final project, for example, allows teamwork. But individual homework assignment does not. You can discuss your homework with other people but you must declare with whom you discussed with.

You must not copy lines of code from internet including online forums.

OpenCV Challenges

- OpenCV is an open source computer vision library. It has a Python wrapper.
- OpenCV is launching a community-wide challenge to update and extend the OpenCV library.
- <http://code.opencv.org/projects/opencv/wiki/VisionChallenge>

Ton

m



Image processing and animal Camouflage, Dr.
Akkanyank, PhD. MIT

Take-home Reading

Szelisky : Chapter 1, What is computer vision?

For people who hasn't taken CSC 280:

Please go through the basic Numpy tutorial:

http://wiki.scipy.org/Tentative_NumPy_Tutorial