Computer Vision

CSE/ECE 576

Linda Shapiro
Professor of Computer Science & Engineering
Professor of Electrical & Computer Engineering
Course Information

• Time:
  – MW: 1:30-2:50

• Location:
  – ECE 037

• Contact:
  – shapiro@cs.uw.edu

• TAs:
  – Kechun Liu
  – kechun@cs.washington.edu
  – Nishat Khan
  – nkhan51@uw.edu
  – Mehmet Saygin Seyfioğlu
  – msaygin@uw.edu

• Website:
  – https://courses.cs.washington.edu/courses/cse576/24sp/
Topics

• Introduction
• Color and Texture
• Image Coordinates, Transforms, and Resizing
• Filters and Convolutions
• Edges and Lines
• Interest Operators, Image Matching, Image Stitching
• Face Detection/Recognition
• Machine Learning Overview including Neural Nets
• Object Detection and Recognition with ML
• Convolutional Neural Networks
• Self-Supervised Learning
• Motion/Optical Flow
• Stereo and 3D Depth Perception
Grading (tentative)

• Six regular assignments (75%)
• One Course Project (25%)
• NO EXAMS (Yay!)
Assignments

• Build a vision library from the ground up
• Mostly in C
• Play with advanced tools, neural networks

• Beginning: lots of skeleton code, explanations
• End: less guidance, more experimentation
Assignment 1: Fun with Color
Assignment 2: Image Resizing and Filtering
Assignment 3: Panorama Stitching
Assignment 4: Neural Networks

Here are the classes in the dataset, as well as 10 random images from each:
airplane
automobile
bird
cat
deer
dog
frog
horse
ship
truck
Assignment 5: PyTorch
Assignment 6: Optical Flow
Final Course Project: Machine Learning for Some Kind of Application
Books

Older, but designed for undergrads and has the basics. Chapters available from our web page.

Newest and available as a pdf online (both the 2010 and 2020 versions).
One Look Is Worth
A Thousand Words--

One look at our line of Republic, Firestone, Miller and United States tires can tell you more than a hundred personal letters or advertisements.

WE WILL PROVE THEIR VALUE BEFORE YOU INVEST ONE DOLLAR IN THEM.

Ever consider buying Supplies from a catalog?

What's the use! Call and see what you are buying. One look at our display of automobile and motorcycle accessories will convince you of the fact.

THAT WE HAVE EVERYTHING FOR
THE AUTO

Piqua Auto Supply House

133 N. Main St.—Piqua, O.
The car is in front of the pole

Sky

Person

Horse

White

Car

Wheel

1m

Road

Shadow
Computer Vision

• Low Level Vision
  – Measurements
  – Enhancements
  – Region segmentation
  – Features

• Mid Level Vision
  – Reconstruction
  – Depth
  – Motion Estimation

• High Level Vision
  – Category detection
  – Activity recognition
  – Deep understandings
Computer Vision

• Low Level Vision
  – Measurements
  – Enhancements
  – Region segmentation
  – Features

• Mid Level Vision
  – Reconstruction
  – Depth
  – Motion Estimation

• High Level Vision
  – Category detection
  – Activity recognition
  – Deep understandings
Measurement

Brightness
Measurement

Brightness

http://www.newworldencyclopedia.org/entry/Same_color_illusion
Measurement

Length

Müller-Lyer Illusion

http://www.michaelbach.de/ot/sze_muelue/index.html
Image Enhancement

Image Inpainting, M. Bertalmíó et al.
http://www.iua.upf.es/~mbertalmio//restoration.html
Image Enhancement

*Image Inpainting*, M. Bertalmío et al.
http://www.iua.upf.es/~mbertalmio//restoration.html
Image Enhancement

*Image Inpainting*, M. Bertalmío et al.
http://www.iua.upf.es/~mbertalmio//restoration.html
Seam Carving

[Shai & Avidan, SIGGRAPH 2007]
Traditional resizing uses and stretches the whole image.

Content-aware resizing uses important areas. Extends in horizontal direction and reduces in vertical.

[Shai & Avidan, SIGGRAPH 2007]
Computer Vision

• Low Level Vision
  – Measurements
  – Enhancements
  – Region segmentation
  – Features

• Mid Level Vision
  – Reconstruction
  – Depth
  – Motion Estimation

• High Level Vision
  – Category detection
  – Activity recognition
  – Deep understandings

The car is in front of the pole
Applications: 3D Scanning

Scanning Michelangelo’s “The David”

- The Digital Michelangelo Project
- UW Prof. Brian Curless, collaborator
- 2 BILLION polygons, accuracy to .29mm
The Digital Michelangelo Project, Levoy et al.
Google’s 3D Maps
Structure estimation from tourist photos
Apple’s 3D maps

https://www.youtube.com/watch?v=InlVv-LsgZE
Computer Vision

- **Low Level Vision**
  - Measurements
  - Enhancements
  - Region segmentation
  - Features

- **Mid Level Vision**
  - Reconstruction
  - Depth
  - Motion Estimation

- **High Level Vision**
  - Category detection
  - Activity recognition
  - Deep understandings
  - Pose estimation
Face detection

• Many new digital cameras now detect faces
  – Canon, Sony, Fuji, …
Vision-based interaction: Xbox Kinect
How hard is computer vision?
“In 1966, Minsky hired a first-year undergraduate student and assigned him a problem to solve over the summer: connect a television camera to a computer and get the machine to describe what it sees.”

Crevier 1993, pg. 88
THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".
“You’ll notice that Sussman never worked in vision again!” – Berthold Horn
Why vision is so hard?
Why is vision so hard?

• Ill-posed problem

[Sinha and Adelson 1993]
Challenges 1: view point variation

Michelangelo 1475-1564

slide by Fei Fei, Fergus & Torralba
Challenges 2: illumination
Challenges 3: occlusion

Magritte, 1957

slide by Fei Fei, Fergus & Torralba
Challenges 4: scale
Challenges 5: deformation
Challenges 6: background clutter

Klimt, 1913

slide by Fei Fei, Fergus & Torralba
Challenges 7: object intra-class variation

slide by Fei-Fei, Fergus & Torralba
Challenges 8: local ambiguity
Challenges 9: the world behind the image
What Works Today?
• Reading license plates, zip codes, checks
Biometrics

Fingerprint scanners on many new laptops, other devices

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

Source: S. Seitz
Mobile visual search: Google Goggles

Google Goggles in Action

Click the icons below to see the different ways Google Goggles can be used.
Face detection

• Many new digital cameras now detect faces
  – Canon, Sony, Fuji, ...

Source: S. Seitz
Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.

Sony Cyber-shot® T70 Digital Still Camera

Source: S. Seitz
Face recognition: Apple iPhoto, Facebook, Google, etc
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... “
Computer alert for drowning girl

A 10-year-old girl has been saved from drowning by a computer system designed to raise the alarm when swimmers get into difficulties.

The girl, from Rochdale, was at the deep end of the pool in Bangor, north Wales, when she sank to the bottom.

The £65,000 system, called Poseidon, detected her on the pool floor and sounded the alarm. A lifeguard pulled her out and she recovered in hospital.
Cameras help confirm Scott suicide ruling

Friday, December 04, 2009

December 4, 2009 (CHICAGO) (WLS) -- Chicago police have closed the case in the death of Chicago School Board President Michael Scott.

Police Supt. Jody Weis says investigators used police cameras in the city to trace Scott's last steps in the hours before his body was found in November.

Scott's death has been ruled a suicide. The medical examiner's office concluded --not long after Scott's body was found -- that he had committed suicide. Police did not dispute the finding but wanted to pursue all the investigative leads they could. They say they have done that and have now reached the same conclusion.
Automotive safety

- **Mobileye**: Vision systems in high-end BMW, GM, Volvo models
  - Pedestrian collision warning
  - Forward collision warning
  - Lane departure warning
  - Headway monitoring and warning

Source: A. Shashua, S. Seitz
Google cars

Christine Dobby
Aug 9, 2011, "Human error blamed after Google's driverless car sparks five-vehicle crash". The Star (Toronto)
Vision-based interaction: Xbox Kinect
Augmented reality, consumer products
Special effects: shape and motion capture

Source: S. Seitz
Vision for robotics, space exploration

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking

*Source: S. Seitz*
Medical imaging

3D imaging
MRI, CT

Image guided surgery
Grimson et al., MIT
Classification of 22q11.2DS

- Treat 2D azimuth-elevation angle histogram as feature vector

<table>
<thead>
<tr>
<th></th>
<th>8x8</th>
<th>16x16</th>
<th>24x24</th>
<th>32x32</th>
<th>Experts’ median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole 2D hist</td>
<td>0.651</td>
<td>0.569</td>
<td>0.79</td>
<td>0.684</td>
<td>0.68</td>
</tr>
</tbody>
</table>
Computer vision research in healthcare

assisted living, patient monitoring
[Lan et al, PAMI 2012]

autism screening
http://www.gatech.edu/newsroom/release.html?nid=60509
Computer vision in the real-world

• Most examples are less than 7 years old
• Very active research area. Many new applications to come.
• A website of computer vision industries maintained by Prof. David Lowe (UBC)
  http://www.cs.ubc.ca/~lowe/vision.html
• Note: website is old but interesting
• Note: David Lowe retired and moved to Google 2015 to 2018
Assignments

• Assignment 1: Fun with Color
• Assignment 2: Image Resizing and Filtering
• Assignment 3: Panorama Stitching
• Assignment 4: Neural Networks
• Assignment 5: Pytorch
• Assignment 6: Optical Flow
• Course Project: Teams working on Machine Learning Projects
Assignment 1

- It’s about color, which we will cover Wednesday.
- It’s meant to be very easy, but you want to start it early.
Assignment 1 Parts

- 1. data structure for an image

```c
typedef struct {
    int h, w, c;
    float *data;
} image;
```

- So an image is a 3D array with height, width and channels (like for colors).

- data is floating point numbers between 0 and 1
Assignment 1: Parts
Read them

• TODO #1: get_pixel and set_pixel
• TODO #2: copy_image
• TODO #3: rgb_to_grayscale
• TODO #4: shift_image (shifts values)
• TODO #5: clamp_image (get values between 0 and 1)
• TODO #6: rgb_to_hsv
• TODO #7: hsv_to_rgb
Have Fun