CPSC 425: Computer Vision

Instructor: Ali Madooei
madooei@cs.ubc.ca

Department of Computer Science
University of British Columbia

Lecture Notes 2016/2017 Term 2
January 3, 2016

Topics
• What is Computer Vision?
• Why you might want to study it?
• Course Overview
What is Computer Vision?

• Computer vision, broadly speaking, is a research field aimed to enable computers to process and interpret visual data, as sighted humans can.
What do you see in this image?
What is Computer Vision?

• Computer vision, broadly speaking, is a research field aimed to enable computers to process and interpret visual data, as sighted humans can.

Figure credit: Fei-Fei Li (Stanford)
What we would like to infer...

Will person B put some money into Person C’s tip bag?
Current state of the art

Can computers match (or beat) human vision?
• Yes and no (but mostly no!)

• Let’s see some examples of state-of-the-art!
Optical character recognition (OCR)

- Technology to convert scanned docs to text
  - If you have a scanner, it probably came with OCR software

Automatic check processing

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

Digit recognition, AT&T labs
Face detection

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

https://www.youtube.com/watch?v=PcoJjqV5E6Y
Face recognition

Facebook!

Apple iPhoto Software
http://www.apple.com/ilife/iphoto/
Vision for biometrics

Fingerprint scanners on many new devices

Face Recognition Login
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… “
Object recognition (in supermarkets)

How Amazon Go Works?
https://www.youtube.com/watch?v=NrmMk1Myrxc
Object recognition (in mobile devices)

Google Goggles

Nokia’s Point & Find
3D Urban Modeling

Microsoft Virtual Earth
Bing Maps

Google Street View

Figure credit: Fei-Fei Li (Stanford)
3D Urban Modeling

Microsoft’s Photosynth
https://photosynth.net/

Figure credit: Fei-Fei Li (Stanford)
Special effects: shape and motion capture
Sports

Sportvision first down line

Nice explanation on www.howstuffworks.com

http://www.sportvision.com/video.html
Automotive Safety and Smart cars

Tesla’s Autopilot

Google Self-driving Cars

Mobileye
Interactive Games: Kinect
Vision for robotics, space exploration

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.

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

3D imaging
MRI, CT

Image Guided Surgery
Grimson et al., MIT

Slide Credit: Steven Seitz (University of Washington)
Computer Vision Problems

• Making computers see and understand

  1. Computing properties of the 3D world from visual data (measurement)

  2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. (perception and interpretation)

  3. Algorithms to mine, search, and interact with visual data (search and organization)
Why Computer Vision is hard?

• Making computers see and understand

  1. Computing properties of the 3D world from visual data (measurement)
     • *Ill-posed problem: real world much more complex than what we can measure in images: 3D -> 2D*

  2. Algorithms and representations to allow a machine to recognize objects, people, scenes, and activities. (*perception and interpretation*)
     • It is computationally intensive
       • *About half of the cerebral cortex in primates is devoted to processing visual information* [Felleman and van Essen 1991]
     • *We do not (fully) understand the recognition problem*
Why Computer Vision is hard?

• Making computers see and understand

  3. Algorithms to mine, search, and interact with visual data 
     (*search and organization*)

• The Internet and the advent of Visual Social Media in particular, 
  has brought with it an enormous explosion of visual content.

  • On Facebook, users upload more than 300 million photos per day.
  • On YouTube, 100 hours of video are uploaded every minute.
  • Instagram host to more than 20 billion shared images.
  • Similar mind-numbing statistics can be noted for other internet-based 
    media such as Flickr, Picassa, Vine, etc.
Why study computer vision?

• It is one of the most exciting areas of research in computing science.
Related Disciplines

Images \[\rightarrow\] Vision \[\leftarrow\] Graphics

Images \[\rightarrow\] Image Processing

Images \[\rightarrow\] Model
Why study computer vision?

• Images (and movies) have become ubiquitous in both production and consumption.

• Therefore applications to manipulate images (movies) are becoming core.

• As are systems that extract information from imagery

Safety  Health  Security  Search

• Look at “The Computer Vision Industry” by David Lowe

Content Credit: Aaron Bobick (Georgia Tech)
Why study computer vision?

- It is among the fastest growing technologies in today’s industry
3609 ATTENDEES
Platinum Sponsors

facebook  Adobe  Microsoft  Amazon  Intel  Google

Athena eyes  MITSUBISHI ELECTRIC

Gold Sponsors

iRobot  SKYDIO  UB ER  BA i

Sensetime  Sighthound  Linkface  Cognex

IBM Research  ebay  here  spare 5

Slide credit: CVPR 2016 Opening Speech
Computer vision is booming

Number of **attendees** at CVPR
Computer vision is booming

Number of papers at CVPR
Topic Trends

Papers per Topic

- Recognition
- Video
- Low-level
- 3D
- Face and gesture
- Motion and tracking
- Statistical methods and learning
- Computational photography
- Unsupervised learning
- Optimization methods
- Biomedical
- Vision for X
- Other

Slide credit: CVPR 2016 Opening Speech
A rough timeline of most active topics

- From (Szeliski, 2010) [See textbooks]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital image processing</td>
<td>Block world, line labeling</td>
<td>Optimal flow</td>
<td>Topological learning</td>
</tr>
<tr>
<td>Generalized cylinders</td>
<td>Pictorial structure</td>
<td>Scale-space processing</td>
<td>Feature-based recognition</td>
</tr>
<tr>
<td>Stereo correspondence</td>
<td>Intrinsic images</td>
<td>Shape from shading, texture, and focus</td>
<td>Region-based modeling</td>
</tr>
<tr>
<td>Optical flow</td>
<td>Image pyramids</td>
<td>Physical-based modeling</td>
<td>Graph cuts</td>
</tr>
<tr>
<td>Structure from motion</td>
<td>Kalman filters</td>
<td>Projective invariants</td>
<td>Particle filtering</td>
</tr>
<tr>
<td>Szeliski, 2010</td>
<td>Markov Random Fields</td>
<td>Kalman filters</td>
<td>Particle filtering</td>
</tr>
<tr>
<td>Physically-based modeling</td>
<td>Projective invariants</td>
<td>Kalman filters</td>
<td>Particle filtering</td>
</tr>
<tr>
<td>3D range data processing</td>
<td>Factorization</td>
<td>Graph cuts</td>
<td>Particle filtering</td>
</tr>
<tr>
<td>Energy-based segmentation</td>
<td>Face recognition and detection</td>
<td>Texture synthesis and inpainting</td>
<td>Category recognition</td>
</tr>
</tbody>
</table>
Course Overview

• This course provides an introduction to the fundamental principles and applications of computer vision.

• Course website
  http://www.ugrad.cs.ubc.ca/~cs425/ or
  http://www.cs.ubc.ca/~Madooei/cpsc425/

• Course Piazza
  https://piazza.com/ubc.ca/winterterm22016/cpsc425
Course Staff

- Instructor: Ali Madooei
- Office: ICCS 237
- Email: madooei@cs.ubc.ca
- Office hours:
  - Friday 3:00 – 4:00 p.m.
  - Or, via appointment.
  - I have open door policy.
- Teaching Assistance (Office hours: TBA)
  - Amon Ge (amon@cs.ubc.ca)
  - Kai Wu (kaywu@ece.ubc.ca)
  - Moumita Roy (moumita@cs.ubc.ca)
  - Mir Rayat Imtiaz Hossain (rayat137@cs.ubc.ca)
Course Topics

• Week 1 - Welcome & Introduction / Image as Function
• Week 2 & 3 - Image Processing (Linear Filtering, Convolution)
• Week 4 - Filters as Templates
• Week 5 - Image Feature Detection (Edges & Corners)
• Week 6 - Texture & Colour
• Week 7 - Midterm (Review and Exam)
• Week 8 - Mid-term break!
• Week 9 - Image Feature Description (SIFT)
• Week 10 - Model Fitting (RANSAC, The Hough Transform)
• Week 11 - Camera Models, Stereo Geometry
• Week 12 - Motion and Optical Flow
• Week 13 - Clustering and Image Segmentation
• Week 14 - Learning and Image Classification
Course Origin

• CPSC 425 was originally developed by Bob Woodham and has evolved over the years. Much of the material this year is adapted from material prepared by Bob.

Previously taught by

• 2015-2016 W2 by Fred Tung
• 2015-2015 W2 by Jim Little
Lectures

• Tue & Thu, 11:00-12:20, DMP 110.
  • Slides (and other teaching materials) will be posted to course homepage.

• Bring your i>Clickers.
  • Register your remote https://connect.ubc.ca/
  • There will be clicker questions:
    • ½ point for participation
    • ½ point for correct answer.
  • The clicker questions contribute 5% to your total grade.
Assignments

• There will be 7 assignments (6 marked)
  • Approximately one every two weeks.

• You will use the Enthought Python Distribution (EPD), specifically Python and four libraries, Python Imaging Library (PIL), NumPy, Matplotlib, and SciPy
  • Assignment-1 will introduce you to this.

• Assignments contribute 25% to your final score.
Midterm Exam

- [Tentatively] on Thursday, Feb 16
  - Here in class during the lecture time.
  - The last lecture before mid-term break.

- Multiple-choices, True/false, Short-Answer questions
- It is aimed to test your "understanding" of the content of this course.

- The midterm is closed book, no notes allowed
- The midterm contributes 25% to your final score.
Final Exam

• The final exam is held during the regular examination period, April 10–28, 2017, and is scheduled by the Registrar’s office.

• Similar to midterm but more extensive with more short/medium answer questions.

• For the final exam, you are allowed one (standard) handwritten double-sided sheet of notes.

• The final exam contributes 45% to your final score.
Grading Scheme

• In-class (clicker questions): 5%
• Assignments: 25%
• Midterm exam: 25%
• Final exam: 45%
Textbook

The course uses the following textbook, which is recommended (but not required):

Computer Vision: A Modern Approach (2nd edition)

- **By:** D. Forsyth & J. Ponce
- **Publisher:** Pearson
- **Pub. Date:** 2012
Textbook

Two useful alternatives that can be freely downloaded as a PDF from SpringerLink, through UBC Library Website (must login using CWL).
Textbook

• These textbooks are on reserve in the reading room:
Readings

• You will be assigned readings from the Forsyth & Ponce textbook.
  • And occasionally from other resources.

• Ideally, you want to read the assigned reading before coming to lecture.
  • Reading assignments will be given in class [posted on slides].
  • A [tentative] list is posted online if you want to read ahead.
Readings

Come to each class prepared to discuss that day’s material:

**Problem:**
• What is the problem addressed?

**Key Idea(s):**
• What is the key idea (or ideas) behind the approach taken?
• What assumptions are made?
• Are there alternative approaches?

**Technical Detail(s):**
• What theory underlies the approach taken?
• What are important practical aspects of implementation, experimentation and application?
Prepare for the next lecture

• Reading from textbook: none!

• [Optional] you may want to read sections 1.1 and 1.2 from (Szeliski, 2010).

• [Optional] watch TED talk by Prof. Fei-Fei Li (Stanford): https://www.youtube.com/watch?v=40riCqvRoMs

• Start working on Assignment-1