

Lecture 1: CS677

Aug 25, 2020

Introduction

- Course number: CSCI 677 (4 units)
- Course name: Advanced Computer Vision
- Instructor: Prof. Ram Nevatia
 - My background, research interests...
- Today's objective
 - Describe course content
 - Conduct of the class
 - Required work, grading
 - Pre-requisites
 - Intro to the potential and problems of vision

Background

- Course web page: <http://courses.uscden.net>
- Office hours
 - Instructor, Tu, Th; 4:00-5:30 P.M., other times by appointment
 - Online Zoom Meeting <https://usc.zoom.us/j/96312659047>
 - Meeting ID: 963 1265 9047
 - Email [nevatia AT usc DOT edu](mailto:nevatia@usc.edu),
 - TA: We have 2 TAs, Xuefeng Hu and Jiaxin Cheng; a third one may be added
 - Contact info and office hours will be posted on course web page

Waiting List

- Waiting list students have access to live classes, video recordings and all class material at DEN class page
 - Access will be turned off in 2 weeks for those unable to register
- Trying to expand capacity, by ~20 students, by adding an additional TA
- More spaces may open if registered students drop the class
- Still, will not be able to accept all on waiting list
 - CS and EE handled on equal priority
 - May prioritize students in 2nd year
- CS566 covers some of the same material with focus on deep learning methods

Online Class

- We are forced to conduct this class online
- All lectures, office hours, TA consultations and Exams will be online only
- How to maintain interactivity in lectures?
 - On-campus class is highly interactive, how to duplicate in online environment?
 - Please ask questions, participate in discussion
 - “Raise hand” if too many speaking
- Slack channel, Piazza other sharing tools can be added
 - Should not turn into sharing assignment solutions
- Learning experience for instructor, TA and students
 - What hardware/software work best?
 - We may take some time to find the answers

What is “Vision”?

- What does the term “vision” mean when applied to human vision?
 - Eyes provide images to the brain; perception happens in the brain
 - Perception consists of computing “good” descriptions of the environment
 - Agree with other measurements
 - Example: Room one is in
- Computer Vision
 - Study of techniques that attempt to create artificial vision capabilities for computers
 - Camera is like an eye: Provides images
 - “Vision” is really about perception

What is Vision useful for?

- Biological organisms
 - Navigate: avoid obstacles, don't get lost...
 - Recognize objects of interest: predators, food, friends...
 - Estimate motion, observe activities, predict...
 - Visual communication, Entertainment
 - ...
- Note: capabilities of biological vision span a broad spectrum
 - Insects, birds, mammals, primates, humans
- Artificial mechanisms
 - Capabilities depend on the goals and the environment
 - Human level generality is difficult to achieve and not required in all cases; in special cases, capabilities may exceed those of humans

Application Areas

- **Mobile robots**
 - Self-driving cars, vehicles for hazardous environments, military vehicles..
- **Visual surveillance, security; facial identification**
- **Manufacturing:**
 - Inspection, assembly...
- **Biometrics**
 - Face/Iris, fingerprint recognition
- **Visual aids**
 - Lane warnings, aid for the blind, ...
- **Multi-media**
 - Visual communication, enhanced reality, models for virtual environments, content based retrieval...
- **Remote Sensing**
 - Maps (2-D and 3-D), change detection, crop/weather assessment...
- **Biomedical**
 - Large population screening, assisted surgery....

Studying Computer Vision

- Study of computer vision is made difficult due to the fact that it is not a “solved” problem
 - Many advances made in the last 5-10 years and many practical systems exist performing important functions
 - However, many basic problems, such as object recognition, do not have solutions approaching that of human vision
 - For comparison, consider some other fields: e.g. OS, Compilers, Databases, Internet/WWW...
 - These fields are also in search of better solutions but strong capabilities exist
- In vision, even the right approach (“paradigm”) is not agreed on
 - Modern trend is to rely on learning from “big data” and neural networks but this alone does not always provide a path to fundamental problems

Course Objectives

- Understanding the key **problems** of vision
- **Alternative approaches** to solving the fundamental problems
- Specific **applications** will be covered only to illustrate the basic techniques
- Provide enough **background for further study** and for **implementation of some practical vision systems**
- We begin with no prior knowledge of computer vision but still will study several very recent techniques (hence “advanced” in the title)
- However, it is **not possible to cover “everything” about “everything”**
 - Not even all state-of-art methods can be covered; these change rapidly
 - >2000 major papers published each year

“Traditional” vs “Modern” Vision

- Field of vision is only ~50 years old, yet some methods are called “classical” or “traditional”
 - “Modern Physics” is more than 100 years old
- Neural network based Deep Learning is the modern approach and dominant in research
- CV courses in major universities continue to focus on traditional methods, e.g. see list at <http://szeliski.org/Book/>
- We will study **both** the “**traditional**” and “**DL**” approaches
 - Traditional approaches continue to be used in practice
 - CV courses in major universities continue to focus on traditional methods, e.g. see Szeliski
 - Stanford CS231n
 - 3-unit, one quarter course; we have ~2x class hours
 - Students interested in **DL only** may prefer to enroll in CS566

Class Difficulty

- CS 677 is **not** an “easy” course
 - Other options to satisfy M.S. requirements may be easier
 - Requires skills in several math topics and good programming skills
- Though only eight years old, the first **textbook is quite out of date**. In particular it lacks any coverage of DL methods.
 - Also, coverage of existing topics is uneven and some parts are very hard to read **but still it is the best available**
- Even though we assume no previous knowledge of computer vision, it is a 600-level course so we will move rapidly with some details for students to work out on their own
- At this time, there is no plan to offer a separate course oriented primarily to MS students (cs574) nor the more advanced courses (cs674 a/b)

Pre-requisites: Programming and CS

- This is a graduate course in CS!
 - Proficiency in Python and C++
 - Must have experience with actual programming, not just concepts of programming
 - Familiarity with various data structures and basic algorithms
- Be able to take informal or mathematical descriptions of algorithms and convert them into working programs
 - We will not discuss code-level details in class
- Work with libraries of code with limited documentation
- Courses in ML or DL are **not** required though could be helpful
 - We will cover the needed amounts of ML/DL in the course but don't expect depth of treatment found in such courses

Pre-requisites: Mathematics

- Calculus
 - Derivatives, partial derivatives, Integration
- Geometry
 - Coordinate systems (Cartesian, spherical, cylindrical...)
 - Equations of entities such as points, lines, planes, circles, spheres...;
 - Computing relations between entities (distance, intersection, angles....)
- Linear Algebra
 - Linear transformations
 - Matrix representation, inversion, eigenvectors....
 - Solutions of systems of equations
- Elementary probability theory
 - Discrete/continuous random variables
 - Joint probability distribution/density functions, conditional probabilities
 - Bayes' theorem
- **We are not conducting a pre-req exam but please take the math requirements seriously!**

Books

- Required:
 - “Computer vision” by Forsyth and Ponce, **Second Edition**, 2012
 - Deep Learning: Algorithms and Applications, I. Goodfellow, Y. Bengio and A. Courville, 2017. Free online version available at <http://www.deeplearningbook.org>
 - “A Guide to Convolutional Neural Networks for Computer Vision”, S.Khan, H. Rahmani, S. Shah and M. Bennamoun, Springer 2018. Free,online version available from SpringerLink if you connect via USC IP address computer.
- Recommended:
 - “Computer Vision: Algorithms and Applications” by Richard Szeliski; electronic copy available for personal use at <http://szeliski.org/Book>
- Additional reading material will be distributed electronically or available in form of tutorials and research papers online

Assignments

- Exact assignment and due dates are not available in advance but expect an assignment ~ every 2 weeks
- Following are expectations based on previous experience
 - One “math” assignment, focusing on imaging geometry
 - Two assignments on “traditional” methods, using OpenCV
 - One will be just to learn how to use library functions
 - Second will be more substantial and require some coding
 - Three DL assignments using PyTorch
 - Image Classification
 - Semantic segmentation
 - TBD: Possibly detection or adversarial attacks
 - Assignments pending suitable cloud support
- Preferred language for programming is Python; if you use another language, support from TAs may be limited or unavailable

Exams and Grading

- There will be two in-class, closed book exams
 - Exam 1, 7th or 8th week of classes (will be announced >1 week in advance)
 - Exam 2, Nov 24, last class day (this is not a cumulative “final” exam)
 - Term paper due on December 3, 2:00 PM (university requirement to have a “summative” experience)
 - We have requested approval to drop this requirement
- Grading weights
 - Assignments 30%; Exam1: 25%, Exam2: 25%; Term paper: 10%, Class attendance 10% (DEN students will be assumed to have perfect attendance)
 - If term paper requirement is removed, Exam1 and Exam2 will count for 30% each

Academic Integrity

- Assignments and Exams are to be completed *individually*, unless otherwise specified.
- Discussions among students is encouraged but *not copying* of the answers.
- The class will be conducted on the basis of mutual trust and respect; for USC ethics code, see <http://web-app.usc.edu/scampus/principles-of-community/>
- Unfortunately, there are occasional instances of gross abuse; these instances will be referred for action according to the USC Student Conduct Code, which can be found at this link:
 - <http://web-app.usc.edu/scampus/university-student-conduct-code/>
- Grades can only be assigned based on performance, not need

Related Courses

- Machine Learning (cs576)
 - Computer vision makes heavy use of these techniques; this course may be helpful but is not a pre-requisite
- Deep Learning
 - CV is now dominated by DL; we will cover what is needed to know about DL for CV in this course but 677 is not a course in DL.
 - cs 566 offers more depth in DL and may have significant
- Image Processing (ee 569)
 - In traditional image processing, goal is usual to enhance, compress or modify images for better human perception or transmission
 - Recent versions of the course have included material with significant overlap with 677
- Miscellaneous machine learning courses in CS and EE

Two Major Components of the Objectives

- Infer 3-D scene geometry
 - Needed for navigation and manipulation
 - May be helpful for object/activity recognition
 - How can we infer 3-D info from a single 2-D image?
 - Can we use multiple images to simplify the problem?
 - Can we measure 3-D directly (and bypass some basic vision problems)?
 - Above problems relatively well understood, many working systems
- Semantic understanding
 - Recognition of objects, relations, activities...
 - Difficult to formulate mathematically
 - Very active area of research: methods have changed from “intuitive” to “statistical” to “deep learning”

Topics to be studied in this class

- **Introduction (1 week)**
Background, requirements and issues, human vision.
- **Image formation: geometry and photometry (2 weeks)**
Geometry, brightness, quantization, camera calibration, photometry (brightness and color)
- **Image segmentation (1 week)**
Region segmentation, Edge and line finding
- **Multi-view Geometry (2 weeks)**
Shape from stereo and motion, feature matching, surface fitting, Active ranging
- **Object Recognition: Traditional Methods (2 weeks)**
HoG/SIFT features, Bayes classifiers, SVM classifiers
- **Object Recognition: Deep Learning Methods (3 weeks)**
Deep neural networks, classification networks, object proposal networks
- **Motion Analysis and activity Recognition (1 week)**
- **Selected Topics (1 Week)**
Face Identification, Vision and language ...

Next Class

- Read ch. 1 of Forsyth/Ponce book
 - Sections 1.1, 1.2.