Updated Syllabus for CS 677: Advanced Computer Vision, Fall 2022

Instructor

Prof Ram Nevatia
Email : nevatia AT usc.edu
Office Hours : TBD

Teaching Assistant: TBD

Class Modality: The class will be held in-person, on-campus. Being a DEN class, lecture videos will also be available to watch later.

Brief Course Description

The course will provide an overview of the challenges of vision, the common approaches and current techniques. While specific examples and applications may be used to illustrate, the focus will be on fundamental techniques and algorithms. We assume no prior knowledge of computer vision but still aim to study many modern, state-of-art techniques.

Course Availability

CS677 is available for CS PhD credit but is also open to CS and ECE MS students. In exceptional cases, undergraduate students may also be admitted. PhD students will be given priority in enrollment; however, MS students are expected to be the large majority of enrollment.

Prerequisites

1. **Mathematics**: Knowledge of and ability to use *calculus, analytical geometry, linear algebra and probability theory*.

2. **Programming**: Ability to program in *Python*.

3. **Other Courses**: There are no specific pre-requisite courses. In particular, courses in AI, Machine Learning, Deep Learning, Computer Vision and Image Processing are *not required*.

4. Entrance Exam: *No exam* will be given to assess pre-requisites. However, *GPA* may be used to screen students for preparedness.

Textbooks

There is, unfortunately, not a single, modern textbook available to cover the topics in this course. We will use published papers and tutorials extensively. Nonetheless, following books will be helpful for study.

Required:
"Computer Vision: A Modern Approach", D. Forsyth and J. Ponce, 2010.

"Deep Learning: Algorithms and Applications", I. Goodfellow, Y. Bengio and A. Courville, 2017 (online version available at no cost for personal use).

"A Guide to Convolutional Neural Networks for Computer Vision", S. Khan, H. Rahmani, S. Shah and M. Bennamoun, 2018 (online version available from a USC account).

Recommended:

"Computer Vision: Algorithms and Applications", Richard Szeliski, Second Edition, 2021; online version available at no cost for personal use at https://szeliski.org/Book/.

Grading Breakdown (Updated)

There will be two exams: Exam1 and Exam2, each counting for 30% of the grade (for a total of 50%). There will be one mathematical assignment and five or six programming assignments. Large scale "projects" are not planned. Total assignments will count for 30% of the grade. Lastly, an end of the term, "term paper", will count for the remaining 10% of the grade. We aim to waive the last requirement but this requires university-level approval; if granted, no term paper will be required and the weights given to the two exams will increase to 35% each.

Programming Assignments

The assignments must be completed using the Python language. We will use OpenCV library for the traditional part of the course and PyTorch for the deep learning component. It is expected that some cloud resources will be made available for assignments requiring use of GPUs; students are not required to have GPU-enabled personal computers of their own.

Detailed Course Syllabus (Updated)

The topic of computer vision is evolving very rapidly. Recent advances have come largely from "data-driven" deep learning and neural network approaches. However, traditional, "model-based" methods continue to be of interest and use in practice and continues to be taught at major universities active in vision research. This course will cover both traditional and deep-learning approaches with an emphasis on the latter category.

Following is a list of topics expected to be covered, in anticipated order, and with expected time to be spent on them. However, this list should be taken as being only indicative and actual topics, the order and the time devoted to them may vary depending on various factors including student interests and new developments in the field.

- Introduction (2 classes)

  - Background, requirements and issues, human vision.
- Image formation: geometry and photometry (3 classes)
  - Geometry, brightness, quantization, camera calibration, photometry
- Image segmentation (2 classes)
  - Region segmentation, Edge detection, Point features (SIFT)
- Multi-view Geometry (3 classes)
  - Shape from stereo and motion, feature matching, Active ranging
- Object Recognition: Traditional Methods (2 classes)
  - Bag of Words, Bayes classifiers, Linear classifiers
- Neural Network Basics (3 classes)
  - Neural nets, CNNs, Backpropagation, SGD, Batch Normalization, COLO
- Image Classification (2 classes)
  - Variety of networks, possibly some coverage of semi-supervised methods
- Object Detection and Semantic Segmentation (4 classes)
  - Variety of approaches (Faster/Mask RCNN, YOLO and related); Human pose estimation, Face ID
- Adversarial Attacks and Defense (1 class)
- Motion Analysis and activity Recognition (2 classes)
  - Optical flow, motion features, classification network
- Selected Topics (2 classes)
  - Vision and language, Vision Transformers…