Syllabus for CS 677: Advanced Computer Vision, Fall 2024 (updated 8/5/24)

Instructor

Prof Ram Nevatia
Phone: N/A; Email : nevatia AT usc DOT edu
Office Hours: TBD; Office Location: TBD

Teaching Assistant: TBD

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. Currently, the course requires D-clearance for new enrollments.

Prerequisites

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

2. **Programming**: Ability to program in *Python* and significant programming experience: in particular, translate methods described in text and mathematical expressions to programs. The course will rarely present code or even pseudo-code.

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, students with weak math or programming skills may have difficulty keeping up with the pace of the course. Students with low GPA should self-assess their ability to successfully complete the course.

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 at **no cost** 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/.

"Foundations of Computer Vision", A. Torralba, P. Isola and W.T. Freeman, MIT Press, 2024

Grading Breakdown

There will be two exams: Exam1 and Exam2, each counting for 25% of the grade (for a total of 50%). Exam 1 will be in the middle of the term, Exam 2 will be at the end. The exact dates will be announced later. It is not planned to have a "final exam"; instead, a "term paper", will be due on December 12 and count for 10% of the grade.
It is anticipated that 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. 10% of the grade will be assigned to attendance (DEN students will automatically earn this 10%).

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:

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. 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 preparation, and new developments in the field. Different colors in the schedule indicate high-level groupings of the topics.

1. **Introduction (1 week)**
   Background, requirements and issues, human vision
2. **Image formation: geometry and photometry (1.5 weeks)**
   Geometry, photometry (brightness and color), quantization, camera calibration
3. **Image segmentation and Feature Extraction (.5 week)**
   Various methods of image segmentation, edge detection, SIFT features
4. **Multi-view Geometry (2 weeks)**
   Shape from stereo and motion, feature matching, surface fitting, Active ranging
5. **Object Recognition: Traditional methods (.5 week)**
   Image features, Various classifiers (Nearest Neighbor, Bayes, SVM)
6. Introduction to Neural Networks and Deep Learning (1.5 weeks)
   Neural networks, loss functions, optimization methods
7. Image Classification and Object Detection (2 weeks)
   LeNet, AlexNet, VGG, ResNet, Efficient Net
   RCNN, Faster RCNN, YOLO, SSD, FPN
8. Semantic Segmentation (1 week)
   Fully Convolutional Networks, Deep Lab, Mask RCN
9. Adversarial Attacks (0.5 week)
   Fast gradient sign method, projected gradient attack, poison attack
10. Activity Recognition (0.5 week)
    Classification networks: 2-stream, C3D, I3D, SlowFast
11. Vision Transformers (0.5 week)
    Transformer architecture, application to image classification and object detection
12. **Vision and Language (1 week)**
    Grounding, zero-shot classification, detection and segmentation, Visual Question Answering
13. **Neural Radiance Fields (.5 week)**
    3-D reconstruction and rendering