ESE 358 COMPUTER VISION
Stony Brook University, Electrical and Computer Engg., Fall 2019, 3 credits.

Instructor: Prof. Murali Subbarao

DRAFT VERSION 1.0. This is subject to some changes.

Pre-requisites: Basic background in Linear algebra, Claculus, Probability, and Programming. Projects will be in MATLAB. If you have prior programming experience (as in ESE 224), then you will need 8 hours to learn enough MATLAB for this course.

Text book:

References
Many online resources.
Some examples:
https://www.cc.gatech.edu/~hays/compvision/
http://vision.stanford.edu/teaching/cs131_fall1617/schedule.html
http://www.cs.cmu.edu/~16385/
http://cs.brown.edu/courses/csci1430/#schedule
http://www.cs.cmu.edu/~16385/s17/
http://www.cs.cmu.edu/afs/cs/academic/class/15385-s12/www/
https://cs.brown.edu/courses/csci1430/proj4/
https://colab.research.google.com/notebooks/welcome.ipynb
http://inst.eecs.berkeley.edu/~cs280/sp15/

Part I Image Formation Models and Image Processing

1. Introduction: Introduction, Overview, and applications.
2. Digital images for representing 2D, 3D, and moving objects. Human eye and digital camera models.
3. MATLAB tutorial for computational vision, and Linear algebra overview. (vectors, points, lines, planes, surfaces, matrices). Other CV tools: Python, numpy, OpenCV, Tensor flow, etc.
4. Image recognition paradigm, Quantitative vision for robotics and industry, and qualitative vision for object recognition (e.g. face recognition).
7. Image filtering: gray-level transformations, histograms, convolution, noise reduction, spatial and Fourier domain filtering and convolution, Gaussian filtering, and image resolution pyramids.
Part II Image Features: detection and matching


Mid-term test 1.

10. RANSAC, Hough transform.
11. SIFT vector, image stitching.

Part III 3D Imaging, 3D Motion, Medical imaging.

13. Three-dimensional shape recovery: 3D from Stereo Images; Stereo Camera model, calibration, matching, rectification.
15. 3D Motion from Video, optical flow, other shape-from-x methods (texture, shading, focus/defocus, Optical flow, etc). Machine and robot vision applications and self-driving cars.
16. Medical Imaging: Modes of medical imaging, X-ray Computed Tomography, image reconstruction algorithms.

Mid-term test 2.

Part IV High-level Vision: Machine Learning, Neural Nets, and Artificial Intelligence

19. Neural Nets, Convolution Neural Nets,
20. Deep learning, AI.

Final Quiz (10%. Final exam will be a 30 minute quiz, with questions having short answers).

There will be around 4 programming projects using MATLAB. Each project may take around 7 hours for completion.

Project 1: 2D and 3D Geometric transforms, imaging in a pin-hole camera.
Project 2: Image filtering, local image features, and model fitting.
Project 3: Pattern Classification: Image segmentation and clustering.
Project 4TBD
GRADING: Grading is based on absolute total score.

Mid-term Test 1: 25% (1 hr 20 mins)
Mid-term Test 2: 25% (1 hr 20 mins)
Final Quiz: 10% (30 mins)
Projects: 30%
Homeworks: 10%
Bonus: 5%: Class participation 2% (answering questions during class)

Grading Policy
Grades are assigned based on absolute percentage of total marks as below.
A: 91--100  A-: 86--90
B+: 81--85  B: 76--80  B-: 71--75
C+: 68--70  C: 64--67  C-: 61--63
D+: 56--60  D: 51--55  F: 0--50

Attendance: <50%: 0%, 50-65%: 1%, 65-80%: 2%, >85%: 3%.