FALL 2020

16-720B Computer Vision

updated 11.22.2020

Days MW
Mode IRR: In-person (Rotation) + Remote
Room CUC MCCNMY
Time 11:40 - 1:00 PM
Units 12
Lecturer Kris Kitani
TAs Nadine Chang, Rawal Khirodkar, Zhengyi Luo, Alireza Golestaneh, Anand Bhoraskar
Class Discussion and Slides `https://piazza.com/cmu/fall2020/16720b/home`

Description
This course introduces the fundamental techniques used in computer vision, that is, the analysis of patterns in visual images to reconstruct and understand the objects and scenes that generated them. Topics covered include image formation and representation, camera geometry, and calibration, computational imaging, multi-view geometry, stereo, 3D reconstruction from images, motion analysis, physics-based vision, image segmentation and object recognition.

Version
Version B of 16-720 is intended for students with prior knowledge of computer vision and prior exposure to machine learning. Undergraduate students should take 16-385 which is the undergraduate version of the class. Those with no exposure to computer vision or machine learning should take the A version of the class. Those with advance experience in computer vision should take the 800 level computer vision courses.

Prerequisites
Linear Algebra, Multivariate Calculus, Probability theory, Programming

Grading
Programming Assignments 100% (6 assignments total).
(1) Hough Transform (10%)
(2) Bag of Visual Words (18%)
(3) Neural Networks (18%)
(4) Homography (18%)
(5) 3D Reconstruction (18%)
(6) LK Image Alignment and Meanshift Tracking (18%)

Grades determined on an absolute scale. Typically 90% and above is A, 80% - 89% is B, 70% - 79% is C, 60% - 69% D, 59% or below is R. There will be extra credit opportunities for students who want to go deeper into the material. Extra credit for class participation will be capped at 3%.

Homework submission and Regrades
5 late days for the entire semester. Use up to 2 late days on one assignment. No credit for assignments submitted after using all late days (to prevent delay of grading). All regrade requests must be made within 2 weeks of the deadline.

Educational Outcomes
(1) Implement the Hough Transform to detect lines in an image
(2) Extract SIFT features to build a Bag-of-Words representation of an image for classification
(3) Perform object recognition using a convolutional neural network
(4) Detect Harris Corners and implement the RANSAC algorithm to find the homography between two images
(5) Perform 3D reconstruction and stereo rectification to implement stereo block matching using two images
(6) Implement a gradient descent based image alignment algorithm to track objects in a video
(7) Apply the mean-shift tracking algorithm to track a colored object

Academic Integrity
All encouraged to work together BUT you must do your own work (code and write up). If you work with someone, please include their name in your write up and inside any code that has been discussed. If we find highly identical write-ups or code without proper accreditation of collaborators, we will take action according to university policies. For scribe notes, you are encouraged to find pre-existing presentation material to aid your preparation. If you use someone's material, please given them credit by including a citation where necessary.

In-Person or Remote

Students in the in-person section will attend physically on a rotational schedule. The rotation schedule will be updated on this syllabus. Please make sure that you have the latest version of the syllabus.

Take Care of Yourself
Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress. All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful. If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at `http://www.cmu.edu/counseling/`. Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

Class Schedule

Date	Topic	Due
Aug-31	Introductions, Policies, Grading, Applications of CV	
	Image Basics	
Sept-2	Filtering, Image Pyramids, Image Gradients	
Sept-7	NO CLASS (Labor Day)	
	Feature Representations	
Sep-9	Lines, Hough Transforms	HW1 Released
Sep-14	Quadratics, Harris Corners	
Sep-16	Feature Descriptors (SIFT, HOG)	
	Visual Learning	
Sep-21	Bag of Visual Words, Classification	HW1 Due, HW2 Released
Sep-23	Support Vector Machine, Neural Networks	
Sep-28	Optimization*	HW1 Graded
	Deep Learning	
Sep-30	Classification: LeNet, AlexNet,	
Oct-5	Classification: VGG, GoogleNet, ResNet, Transformer	HW2 Due, HW3 Released
Oct-7	Detection: RCNN, Fast RCNN, Faster RCNN, YOLO, SSD	
Oct-12	Generation: Generative Adversarial Networks, VAE*	HW2 Graded
	Physic-based Vision	
Oct-14	Guest Lecture: Reflectance and Illumination*	
Oct-19	Guest Lecture: Photometric Stereo*	HW3 Due, HW 4 Released
	Multi-view Processing	
Oct-21	SV: 2D Transformations (Euclidean, similarity, affine, projective)	
Oct-26	SV: Homography estimation, Single View Geometry	HW3 Graded
Oct-28	SV: Camera matrix, Camera Models, Pose Estimation	
Nov-2	2V: Triangulation, Epipolar Geometry	HW 4 Due, HW 5 Released
Nov-4	2V: Fundamental Matrix, 8 Pt, Reconstruction	
Nov-9	Guest Lecture: Semantic Segmentation*	HW4 Graded
Nov-11	2V: Bundle Adjustment, Stereo Rectification, Stereo Matching	
	Motion, Tracking and State Estimation	
Nov-16	Brightness Constancy, Optical Flow	
Nov-18	Horn and Shunck, Lucas-Kanade Image Alignment	HW5 Due (19th), HW6 Released (20th)
Nov-23	Forward/Inverse Compositional Image, KLT Alignment	
Nov-25	NO CLASS (THANKSGIVING)	
Nov-30	Bayesian Filtering, Kalman Filtering, EKF, MonoSLAM	HW5 Graded
	Advanced Topics	
Dec-2	Guest Lecture: Perception for Autonomous Driving*	
Dec-7	Guest Lecture: Deep 3D Object Pose Estimation*	HW6 Due
Dec-9	Guest Lecture: Ethics and Bias in Computer Vision*	