Computer Vision (16-720 A) Spring 2018 [REV00]

Days: TR
Room: WEH 7500
Time: 12:00 - 1:30 PM
Lecturer: Kris Kitani
TAs: Leonid Keselman, Mohit Sharma, Arjun Sharma, Rawal Khirodkar, Aashi Manglik, Tanya Marwah
Class Discussion, Slides and Office Hours: `piazza.com/cmu/spring2018/16720a/home`
Assignments and Grades: `canvas.cmu.edu`

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 A of 16-720 is intended for students with no prior knowledge of computer vision and minimal exposure to machine learning. Undergraduate students should take 16-385 which is the undergraduate version of the class. Those with intermediate exposure to computer vision or machine learning should take the more advanced B 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) Convolutional Neural Networks (18%)
(4) Homography (18%)
(5) 3D Reconstruction (18%)
(6) LK Image Alignment and Meanshift Tracking (18%)

Late Submissions
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).

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.

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
	Introduction	
Jan-16	Introduction, class overview	
	Image Processing	
Jan-18	Filtering, Image Pyramids	
Jan-23	Image Gradients and Lines	
Jan-25	Hough Transforms	HW1 Released
	Visual Features	
Jan-30	Quadratics, Harris Corners, Multi-Scale	
Feb-1	Feature Descriptors (GIST, HOG, SIFT, MOPS)	HW1 Graded
	Visual Recognition	
Feb-6	Overview: Probability, Bag of Visual Words, Sliding window, Part Models	
Feb-8	Basic ML: K-means, Nearest Neighbor, Naive Bayes, SVM	HW2 Released
	Deep Learning	
Feb-13	Perceptron: Gradient Descent, BackProp	
Feb-15	Classification: LeNet, AlexNet, VGG, ResNet, GoogLeNet, DenseNet	HW2 Graded
Feb-20	Detection: RCNN, Fast RCNN, Faster RCNN	
	Image Transformations	
Feb-22	2D Transforms and Alignment (LLS)	
Feb-27	2D Alignment (DLT, RANSAC)	HW3 Released
Mar-1	TBD	
	Multi-view Geometry	
Mar-6	Camera Matrix, Pose Estimation, Triangulation	HW3 Graded
Mar-8	Epipolar Geometry, Essential Matrix	
Mar-13	SPRING BREAK	
Mar-15	SPRING BREAK	
Mar-20	Fundamental Matrix, 8 Point Algorithm	
Mar-22	Structure from Motion	HW 4 Released
	Image Registration	
Mar-27	Optical Flow (LK, Horn and Shunck)	
Mar-29	LK Image Registration (Additive/Inverse)	HW 4 Graded
	Stereo and Segmentation	
Apr-3	Motion Segmentation *	HW 5 Released
Apr-5	Stereo Rectification, Block Matching	
Apr-10	Normalized Cuts, Fully Convolutional Networks, Mask RCNN	HW5 Graded
	Tracking	
Apr-12	State Estimation, Bayesian Inference	
Apr-17	Tracking (KLT, Mean-Shift, GoTurn, MDNet)	HW6 Released
Apr-19	NO CLASS (Spring Carnival Weekend)	
Apr-24	Probability, Temporal State models, HMM, LSTMs*	HW6 Graded
Apr-26	Kalman Filtering, EKF, MonoSLAM	
	Special Topics	
May-1	TBD	
May-3	TBD	