Computer Vision (16-720 A) Spring 2018 [REV09]

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

Introduction

Date	Topic	Due
Jan-16	Introduction, class overview	

Visual Features

Jan-18	Filtering, Image Pyramids, Image Gradients and Lines	
Jan-23	Hough Transforms	
Jan-25	Quadratics, Harris Corners, Multi-Scale	HW1 Released
Jan-30	Feature Descriptors (GIST, HOG, SIFT, MOPS)	

Visual Recognition

Feb-1	Bag of Visual Words, K-means, Nearest Neighbor	
Feb-6	Probability, Naive Bayes, SVM, Perceptron	
Feb-8	MLP, SGD, Backpropagation	HW1 Due, HW2 Released

Deep Learning

Feb-13	LeNet, AlexNet (Response Normalization, Dropout, Augmentation)	
Feb-15	VGG, ResNet, GoogleLeNet (Batch Normalization)	
Feb-20	RCNN, Fast RCNN, Faster RCNN, SSD, YOLO	

Feb-22	NO CLASS (IARPA Site Visit)	HW2 Due, HW3 Released
Feb-27	NO CLASS (UMD)	

Image Formation

Mar-1	Pinhole Camera, Exposure, Focus, Digital Photo	

Image Geometry

Mar-6	2D Transforms, LLS, DLT, RANSAC, Homography	
Mar-8	Single view 3D	HW3 Due, HW 4 Released

Mar-13	SPRING BREAK	
Mar-15	SPRING BREAK	

Structure from Motion

Mar-20	Camera Matrix, Pose Estimation, Triangulation	
Mar-22	Guest Lecture: Human Pose Estimation	
Mar-27	Epipolar Geometry, E,F, 8-Point, Reconstruction	HW4 Due, HW 5 Released

Stereo, Blending

Mar-29	Image Rectification, Stereo Block Matching	
Apr-3	Multi-band blending, Poisson Blending	

Alignment, Tracking

Apr-5	Brightness Constancy, Optical Flow (LK, Horn and Shunck)	
Apr-10	Image Registration (Additive/Inverse)	HW 5 Due, HW 6 Released
Apr-12	KLT, Mean-Shift Tracking	

SLAM

Apr-17	State Estimation, Bayesian Inference, Temporal State models, HMM	
Apr-19	NO CLASS	
Apr-24	Kalman Filtering, EKF, MonoSLAM	

Advanced Topics

Apr-26	3D Reconstruction, Panoptic Studio (Hanbyul)	HW 6 Due
May-1	Action Recognition (Gunnar)	
May-3	Activity Forecasting (Nick), Low-Shot Learning (Yuxiong)	