CS 280, Computer Vision

Instructors: Alyosha Efros, Jitendra Malik, Stella Yu
GSIs: Pulkit Agrawal, Yi Wu
{efros, malik, stellayu, pulkitag, jxwuyi} @ berkeley.edu

January 19, 2018

Introduction

Computer vision seeks to develop algorithms that replicate one of the most amazing capabilities of the human brain – inferring properties of the external world purely by means of the light reflected from various objects to the eyes - and using this information to control actions in the world. We can determine how far away objects are, how they are oriented with respect to us, and in relationship to various other objects. We reliably guess their colors and textures, and we can recognize them - this is a chair, this is my dog Fido, this is a picture of Bill Clinton smiling. We can segment out regions of space corresponding to particular objects and track them over time, such as a basketball player weaving through the court. We can use the information we extract from images or video to manipulate objects in the world and navigate in environments while avoiding obstacles.

In this course, we will study the concepts and algorithms behind some of the remarkable successes of computer vision – capabilities such as face detection, reconstructing three-dimensional models of cities, automated monitoring of activities, segmenting out organs or tissues in biological images, and sensing for control of robots. We will build this up from fundamentals – geometry and radiometry of image formation and statistical machine learning techniques such as neural networks.

Topics

- Introduction – The Three R's – Recognition, Reconstruction, Reorganization
- Static perspective – the pinhole camera model
- Transformations – rotation,translation,affine,projective
- Dynamic perspective & optical flow
- Radiometry of image formation
- Basic image processing operations – filters, features

- Biological visual processing – retina, V1 and beyond
- The feedforward model of visual processing – convolutional networks
- Useful patterns for neural networks - residuals, dilations, feedback
- 3D reconstruction from multiple images
- Short range and long range correspondence, optical flow
- Large scale reconstruction, SLAM
- Applications in Augmented and Virtual Reality
- Perceptual organization - grouping and figure/ground. Objects.
- Contour detection and bottom-up segmentation
- Object Detection and Instance segmentation using region-based CNN and variants
- Inferring shape and spatial layout using learning-based approaches
- Image synthesis
- Perceiving humans - pose, activity recognition
- Face recognition
- Visual navigation
- Visual control of manipulation

Administrivia

The course has a website on `http://inst.eecs.berkeley.edu/cs280/sp18/` (will be live shortly). We will use it to post lecture notes, papers to read, homework assignments etc. Piazza will be used for discussions related to the assignments and general announcements.

The course will have a midterm examination and a final project. In addition, there will be 5 or so homework assignments, which will typically involve Matlab/Python programming (you are free to use alternative packages). Matlab runs on all the Instructional Windows and UNIX systems. Instructions and toolkits are described in

`http://inst.eecs.berkeley.edu/cgi-bin/pub.cgi?file=matlab.help`

CS280 students can use their existing EECS Windows accounts in EECS instructional labs, and they can request new accounts (for non-majors) or additional access to Instructional resources by following the instructions about 'named' accounts in

`http://inst.eecs.berkeley.edu/connecting.html#accounts`

Students can logon remotely and run it on some of our servers:

`http://inst.eecs.berkeley.edu/connecting.html#labs`

For the neural network component of the class, we will be using pytorch `http://pytorch.org/`. Knowledge of *python* will be assumed. We are in the process of finalizing the GPU resources and will announce the instructions on how to use GPUs for training neural networks shortly.

Webcast of lecture videos will be available at:
`https://www.youtube.com/c/CalESG/live`
Unlisted playlist for archived videos:
`https://www.youtube.com/playlist?list=PLkFD6_40KJIxge3DKNXckQ8YMvBWH70dm`
Please do not post the videos to public forums – they are for Berkeley consumption only.

Notes and powerpoint slides will be available for some parts of the course. Forsyth and Ponce's *Computer Vision : A Modern Approach* and Szeliski's *Computer Vision: Algorithms and Applications* books cover much of the same material, and may be useful as a reference for many parts of the course. Szeliski's book is available online for free.

We hope you will enjoy the course!