**CS188: Introduction to Computer Vision**

**Course Number**: CS 188.2

**Instructors**: Profs. Stefano Soatto, Fabien Scalzo, Achuta Kadambi

**Teaching Assistants**: Ali Hatamizadeh, Albert Zhao, Nikita Sivakumar, Alexandre Tiard

**Prerequisites**: Strong proficiency in linear algebra, statistics and basic probability required. Familiarity with Python.

**Grade Structure:** Letter grades will be based on performance in a midterm exam (35%), three homework assignments (30% each), final exam (35%). 4 units.

**Course Material:** There is no required textbook. Course material and recorded lectures will be made available online to registered students. Several titles will be held for consultation at the UCLA Library.

**Course Description**

Students will be given an introductory exposure to the “Four Rs” of Computer Vision:

* Reconstruction: Building a model of the 3D **geometry** of the scene from images
* Recognition: Inferring the **semantics** (identity and relations) among objects within the scene from images
* Rendering: Inferring **photometric** properties of scenes, separating material properties from illumination, and generation of synthetic images
* Regulation: Inferring the **dynamics** of objects in scenes, and using vision as a sensor to control mobile assets within the scene

Along the way, students will be led through the study of foundational models of image formation, physics-based vision, basic image processing, fundamentals of pattern recognition using Deep Learning, including convolutional neural networks (CNNs) and generative adversarial networks (GANs).

Homework, discussions, and forum will be done in Piazza.

**Objectives**

Students are led to develop analytical and computational tools to enable them to build basic visual recognition systems (written digit recognition, face discrimination), a basic visual reconstruction system (3D point cloud reconstruction from images) and foundational skills to enable them reading and understanding state-of-the-art literature published at modern Computer Vision and Machine Learning conferences. Students will conduct quantitative projects and are expected to develop critical skills to assess the quality of data, their annotation, the outcome of probabilistic modeling and processing, and the reporting of outcomes in a structured report, as part of three collaborative homework problems. Such homeworks are also designed to foster critical thinking and interaction with peers. Successful completion of the course curriculum gives students the necessary skills to operate in the areas of Augmented Reality (AR), Image-based Recognition, Robotics, and other areas under the broad umbrella of Artificial Intelligence (AI).

**Lecture Plan**

**Part I: Introduction**

**Lecture 1: 09/26:** Introduction, organization of the course, prerequisites, expectations, grading, policies. Overview of the course. Brief perspective of instructor’s activities in CV. Brief overview of prerequisite concepts (things we know you have seen, you probably forget, but you need to know to take this class).

Read: [http://www.gatsby.ucl.ac.uk/~pel/misc/stupidity.html](http://www.gatsby.ucl.ac.uk/~pel/misc/stupidity.html?fbclid=IwAR3MlahsDEYq1NARdm0kh6xy5W0_2_T6Zd2vwI-yj1sxtKmbpJYj6x_-tsk)

**Part II: CV from a Signal Processing Perspective**

**Lecture 2: 10/01:** Basic Image processing: Convolution, filtering, smoothing, Gaussian and Laplacian pyramids, edge detection.

**Lecture 3: 10/03:** Local features, scale-space, SIFT, introduction to basic classification.

**Lecture 4: 10/08:** Feature tracking, optical flow.

**Part III: CV from a Physics-based Perspective**

**Lecture 5: 10/10:** Singular value decomposition, 2D cameras, pinhole camera and the projective plane.

**Lectures 6: 10/15:** 2D image transformations, panoramas, RANSAC.

**Lecture 7: 10/17:** Euclidean geometry, rigid body motion.

**Lecture 8: 10/22:** Epipolar geometry, sparse 3D reconstruction.

**Lecture 9: 10/24:** 3D cameras and processing - stereo and structured light. 3D Registration and ICP.

**Lecture 10: 10/29:** MIDTERM

**PART IV: CV from a Machine Learning Perspective**

**Lecture 11: 10/31:** Basics of statistical decision theory and pattern recognition. Linear classification. Risk, loss, regularization

**Lecture 12: 11/05:** Introduction to Deep Learning: Convolutional Neural Networks, auto-encoding, Stochastic Gradient Descent.

**Lecture 13: 11/07:** Deep Learning continued, Large-scale Image Classification.

**Lecture 14: 11/12:** Object Detection: SSD, Yolo. Semantic Segmentation.

**Lectures 15: 11/14:** Generative models. Encoder-decoder architectures, GANs.

**PART V: Applications**

**Lecture 16: 11/19**: Applications I – Deep Learning for Medical Imaging.

**Lecture 17: 11/21:** Applications II – Autonomous Navigation (self-driving cars, mobile robots).

**Lecture 18: 11/26:** Guest Lecture by Dr. Nikhil Naik.

**Lecture 19: 12/03:** Recursive 3D reconstruction and pose estimation, AR/VR.

**Lecture 20: 12/05:** Recap.

**Final Exam:** Thursday, December 12, 2019. 11:30 AM - 2:30 PM

**Reading list**

**Course lectures** available through SEASNET two days after lecture day.

**MASKS:** Y. Ma et al*., An Introduction to 3D Vision,* Springer 2004 (on-line pre-print available for free at <https://www.eecis.udel.edu/~cer/arv/readings/old_mkss.pdf)>

**Tutorial compendium** on Github: <https://github.com/sunglok/3dv_tutorial>

**F&P:** Forsyth and Ponce: *Computer Vision, a Modern Approach*, MIT Press (on-line pre-print available for free at <http://cmuems.com/excap/readings/forsyth-ponce-computer-vision-a-modern-approach.pdf)>

**DDL:** A. Zhang et al., *A Dive into Deep Learning.* <http://d2l.ai/> interactive community-based book project with step-by-step code examples.

**MatConvNet:** Matlab-based tutorial on convolutional neural networks: <http://www.vlfeat.org/matconvnet/>

**VLFeat:** <http://www.vlfeat.org/> including tutorial on low-level features.

**GluonCV:** <https://gluon-cv.mxnet.io/> a python-based tutorial and library.

**Smola:** A. Smola et al., *Introduction to Machine Learning* (on-line pre-print available for free at <http://alex.smola.org/drafts/thebook.pdf)>