List of Course Project

1. Paranomic video stitching  (Contact Yilin Song ys1297@nyu.edu)
   Description: Generating a synthesis view from stationary camera located at retail stores (video data will be provided)
   Requirements: Basic knowledge of image processing
   Deliverable: Program that could run in real time for video stitching. Either in C++ or python.

2. Retail store usage analysis. (Contact Yilin Song ys1297@nyu.edu)
   Description: Estimating the utilization of store using single camera or stitched camera, via pedestrian detection or foreground background detection. (video data will be provided)
   Requirements: machine learning. Deep learning and optimization is a plus.
   References:

3. Image Deblurring with conventional approach or deep learning approach (Contact Yilin Song ys1297@nyu.edu)
   Description: Facial recognition or verification in a noisy environment is hard to tackle especially when there is blurring caused by motion. Students would explore facial deblurring with real image/video recording.
   Requirement: image processing, knowledge of convolutional neural networks, and programming skills in one of the deep learning packages is a plus
   References:

4. Facial depth map estimation and verification (Contact Chenge Li cl2840@nyu.edu)
   Description: Telling apart real people and faces appeared in painting using stereo video. (images from stereo camera will be provided)
   Requirements: Basic knowledge of image processing, machine learning. Deep learning and optimization is a plus.
   Reference:

5. Hashing design for Images/Videos in retrieval systems, (Contact Shervin Minaee, sm4841@nyu.edu )
Requirements: Basic knowledge of image processing, machine learning, and optimization.
References:

6- Image segmentation or denoising using deep learning (Contact Shervin Minaee, sm4841@nyu.edu )
Requirement: knowledge of image processing and convolutional neural nets, programming in one of the deep learning packages.
References:
3- D Eigen, D Krishnan, and R Fergus. "Restoring an image taken through a window covered with dirt or rain.", ICCV, 2013.

7- Human Action Recognition from Videos, ( Contact Shervin Minaee, sm4841@nyu.edu )
Requirement: basic knowledge of machine learning, such as classification, dimension reduction, etc. Programming skills in Python or Matlab. Knowledge of convolutional neural network is a bonus, but not required.
References:

8. Image Compression and Video Compression with deep learning. (Contact
Ran Wang rw1691@nyu.edu

Description: As traditional image coding have been beaten by deep learning approach by a large margin, video coding with deep learning is the next frontier. Students involved in this project would proposal innovative approaches to reshape video coding with deep learning.

Requirement: good understanding with compression and deep learning

Reference:

9. Motion Estimation and Foreground background subtraction (Contact Chenge Li cl2840@nyu.edu)

Description: Advanced motion estimation and foreground background subtraction with tensor decomposition and etc.

Requirement: Optimization, tensor decomposition and sparse modeling

Reference:

10 Image inpainting (Contact Yuan Wang yw1225@nyu.edu)

Description: TBA.

Requirement: Mathematical demanding. You need to be able to understand equation derivation and the meaning behind them. Optimization knowledge is a plus. Certain high-level programming experience is needed to implement algorithms, such as python or Matlab.

Reference:

11. Haze removal (Contact Yuan Wang yw1225@nyu.edu)

Description: Now days, haze strikes many places in the world, your hometown or your long-anticipated trip destinations. But still, you can take clear photos with
haze removal technique.

**Requirements:** basic image processing knowledge. High level programming skill is needed such as Python and matlab. Optimization knowledge is a plus.

**Reference:**


**12. Medical image segmentation** *(Contact Jenwei Kuo jwk299@nyu.edu)*

**Description:** Locate and segment the object of interest out of a 3D medical image (e.g. ultrasound, CT, MRI).

**Requirement:** Basic knowledge in 3D image segmentation methods (e.g. graph cuts, level-set based methods, active contour).

**13. Medical image registration** *(Contact Jenwei Kuo jwk299@nyu.edu)*

**Description:** Characterize and register multiple 3D volumes (e.g. Brain ventricle) or 3D medical images (e.g. ultrasound, CT, MRI).

**Requirement:** Basic knowledge in 3D shape descriptor or 3D image registration.

**14. Kinect Project for pose cluster** *(Contact Andy Chiang atc327@nyu.edu)*

**Description:** According motion camera array and Kinect data to correct and cluster users’ pose.

**Requirement:** Basic knowledge of computer vision, image process. Programming skill in Python.

**Reference:**

**15. Kinect Project for pose reconstruct** *(Contact Andy Chiang atc327@nyu.edu)*

**Description:** According motion camera array and Kinect data to detect, correct and reconstruct users’ behavior.

**Requirement:** Basic knowledge of computer vision, image process. Programming skill in Python.

**Reference:**
16. **360 Degree Video Stitching**  (Contact Fanyi Duanmu / fanyi.duanmu@nyu.edu)

**Description:** In this project, the student(s) will conduct literature survey on existing 360 degree image/video stitching algorithms and develop their customized algorithm(s) for 360 degree video stitching and compare with benchmark solutions (e.g., rendering artifacts, etc).

**Team:** up to 3 students

**Prerequisites:** Basic image processing and computer graphics knowledge required. Previous coding experience in MATLAB, OpenCV, OpenGL is a plus but not required.

**Deliverable:** A literature summary report for existing 360 video stitching and projection algorithms. A demo program that can process input images from 360 camera set and populate the stitched image. (GUI recommended but not required.)

17. **360 Degree Video on-demand Streaming System Design**  (Contact Fanyi Duanmu / fanyi.duanmu@nyu.edu)

**Description:** In this project, the student(s) will study the video streaming basics, set up an end-to-end 360 degree video streaming platform (similar to Youtube and Facebook but with additional functionalities). The students will learn how to set up web server, host video contents, stream on-demand videos and develop WebUI player.

**Team:** up to 3 students

**Prerequisites:** Previous experience with Javascript, HTML required or a fast learner of web development languages. Web design experience is a big plus. Previous experience with OpenGL, WebGL libraries are preferred but not required.

**Deliverable:** A literature study report on 360 video streaming (guided by PhD). A DASH-based web-server platform. This system can stream adaptive 360 videos according to client bandwidth and buffer status. The client can subscribe on-demand 360 videos from smartphone or PC. Additional functionalities to be extended based on progress.

18. **360 Degree Video View Prediction Study**  (Contact Fanyi Duanmu / fanyi.duanmu@nyu.edu)

**Description:** In this project, the student(s) will conduct statistical study on the
perceptual behaviors over 360 video contents and understand how human visual system (HVS) reacts to different kinds of 360 video contents. They student(s) will need to develop a webUI to collect client head motion traces (or equivalently viewing angle traces). Later, the student(s) will conduct machine learning studies over collected traces and derive a view prediction model based on visual and audio clues.

Team: 3 students. (1 opening left)

Prerequisites: Previous experience with web development in JavaScript, HTML, PHP or similar. Machine learning knowledge is a big plus but can be learned gradually. A Fast learner and team player.

Deliverable: A literature study report on existing view prediction solutions. A WebUI collecting viewing angle traces. A machine learning model that can predict future view direction based on incoming frame contents and viewing history. A statistical report evaluating the accuracy of the proposed machine learning model.

19 Action/Scene recognition using both color and motion information through deep learning (Contact Yilin Song ys1297@nyu.edu)

Description: Students should be able to adapt the 3D convolution neural network framework to optical flow streams. And use joint feature from RGB streams and optical flow streams to perform action recognition. Students can investigate how to do this in one GOP only and then integrate GOP results following approach of C3D paper. But eventually, should also consider how to integrate GOP features (either using the attention weighted average, or through bag of words idea, which each word is a quantized version of the original features through clustering analysis).

Reference:

20 Video prediction using deep learning in wavelet domain (Contact Yilin Song ys1297@nyu.edu and Ran Wang rw1691@nyu.edu)

Description: Decompose a video into wavelet sub bands using 3D wavelet transform, and using one of the learning based prediction approach (LSTM or 3D convolutional net, with only auto-encoder approach or with adversarial formulation, probably to start just use auto-encoder formulation) to do prediction for each wavelet band independently. (later on, also compare with predicting each band using reconstructed signals from lower scale as well)

Reference:

### 21. Painting with deep artificial neural network
(Contact Yuan Wang yw1225@nyu.edu)

**Description:** Come and meet the beauty: the joint of cutting-edge image processing technology and art. Now, artificial neural network knows how to abstract art style from paintings, and applies it to your own content. We will survey what has been done and what we can do beyond it. Make sure to check the fascinating examples in the reference.

**Requirement:** Programming demanding in terms of learning new tools. Python, Tensorflow and possible AWS will be used. If you have had experience with them, it makes things much easier. However, if you are really interested in it and feel confident at learning new tools, you can catch up with it.

**Reference**