Final Project Overview and Topics List

Description

For the final project you will apply knowledge that you have learned in class plus additional research to build a computer vision system that achieves near state-of-the-art results in an area you select. The purpose of this document is to give you an idea of what is expected and the areas you can choose to work on. We have provided a limited list of final project topics in order to make grading easier while still allowing you to pick concepts you are familiar with or interested in learning more about. You must select a topic from one of the four listed in this document.

Please take time this week to read about the available topics so you have a better idea which to select when the final project is released.

Final Project Structure

The final project, regardless of the topic selected, will have the following 2 deliverables:

- Code:
  All code used to produce your results must be submitted with your report. Make sure it is setup to run as is with detailed instructions for the grader. In addition to the code used to create your report, you must submit instructions for running your code on an arbitrary image or video.

- Report (3-6 pages including images and references):
  Unless otherwise noted in the topic description, your report must contain

  1) A clear and concise description of the algorithms you implemented. This description should include references to recently published computer vision research and show a deep understanding of your chosen topic.
  2) Results from applying your algorithm to images or video. Both positive and negative results should be shown in the report and you should explain why your algorithm works on some images, but not others.

  You report should be written to show off your work and demonstrate a deep understanding or your chosen topic. The discussion in your report should be technical and quantitative wherever possible.

List of topics

- **Enhanced Augmented Reality:**
  For this topic you will explore techniques for inserting images and video into a scene without markers. Your final algorithm should identify a clear surface in a video and project an advertisement onto that

surface. The advertisement should remain in a fixed position in real-world coordinates as the camera is moved. Your algorithm should also handle the case where the advertisement location temporarily moves out of the camera's field of view.

For this topic only, your report should contain run-time benchmarking of your code and a discussion of the processing power required to run your code in real-time (30 frames / second) on HD video. Your code should be optimized to run in as close to real-time as possible. This replaces the requirement to benchmark your algorithms against a public imagery database.

Related Lectures (not exhaustive): 3A-3D, 4A-4C, 6A-6B, 7A-7D

- **Stereo Correspondence:**
  For this topic you will implement two algorithms. First, you will write code for a simple sum squared difference stereo correspondence algorithm as described in lecture. Second, you will write stereo correspondence code incorporating an advanced energy minimization or graph-cut method of your choice. If you choose this topic, you will need to perform additional research on stereo correspondence. You are expected to write code that takes in a pair of images and returns a disparity map.

  Sample Dataset: http://vision.middlebury.edu/stereo/data/scenes2014/
  Related Lectures (not exhaustive): 3A-3D, 4A-4C

- **Classification and Detection with Convolutional Neural Networks:**
  For this topic you will design a digit detection and recognition system which takes in a single image and returns any sequence of digits visible in that image. For example, if the input image contains a home address 123 Main Street, you algorithm should return "123".   One step in your processing pipeline must be a Convolutional Neural Network (CNN) implemented in TensorFlow or PyTorch. Digit classification should be performance separately from segmentation. If you choose this topic, you will need to perform additional research about CNNs. Note that the sequences of numbers may have varying scales, orientations, and fonts, and may be arbitrarily positioned in a noisy image.

  Sample Dataset: http://ufldl.stanford.edu/housenumbers/
  Related Lectures (not exhaustive): 8A-8C, 9A-9B

- **Activity Classification using MHI:**
  For this topic, you will implement the methods presented in lecture to create Motion History Images (MHIs) and use these images to perform activity classification in video. Your algorithm should take in a video of a human walking, jogging, running, boxing, waving or clapping, and correctly classify the behavior. Intermediate steps in your classifier might include generating features from motion history images and some form of machine learning on those features to classify the activity.

  Sample Dataset: http://www.nada.kth.se/cvap/actions/
  Related Lectures (not exhaustive): 8A-8D, 9A-9B

What Not To Do

Sometimes knowing what not to do on an assignment is just as useful as knowing what to do. What that in mind we've provide a short list. This list is not exhaustive!

- **Don't use external libraries for core functionality**
  You are encouraged to use libraries while writing code for your final report. However you will receive a low score if the main functionality of your code is provided via an external library. For example, if you choose the stereo correspondence topic, you should not use OpenCV's stereo functions.

- **Don't copy code from the internet**
  The course honor code is still in effect during the final project. All of the code you submit must be your own. You may consult tutorials for libraries you are unfamiliar with, but your final project submission must be your own work.

- **Don't use pre-trained machine learning models**
  If you choose a topic that requires the use of machine learning techniques, you are expected to do your own training. Downloading and submitting a pre-trained model is not acceptable for this assignment.

- **Don't rely on a single source**
  We want to see that you performed research on your chosen topic and incorporated ideas from multiple sources in your final results. Your project should not be based on a single research paper and definitely should not be based on a single online tutorial.