



Final Project  
Computer Vision, EN. 600.461/661  
Professor Austin Reiter  
Fall 2016

This course will have a final project component to it. The best way to *learn* computer vision is to *do* computer vision. So let's do just that!

**Overview:**

The final project will makeup 35% of your final grade, and will consist of: a project proposal write-up (due: 10/20/16), an in-class poster presentation (will take place during final class period on 12/08/16, 12:00-1:15 pm), and a final report (due: on 12/08/16 - **NO EXCEPTIONS!**)

You may work in teams, up to 4 people (no more than 4 people unless otherwise approved by me), and if you'd rather work alone that's allowed. Your project is designed to be open-ended, but it must be approved by me and the TAs in the project proposal stage. **If you are working in a team, I have higher expectations that those who choose to work alone!** Please come talk to me if you're unsure if your project is acceptable. I don't want it to be too easy, and I don't want it to be too hard. The point is to get some good hands-on computer vision experience with an actual project using the concepts we learn throughout the semester. Students in 461 and 661 can work as a team but 661 projects will have higher expectations.

**Programming Language:**

You may use any programming language you want here (e.g., MATLAB, C, C++, Java, Python etc). You may also use the Image Processing Toolbox provided by MATLAB as well as openly available toolkits on the web. I will suggest OpenCV as a great toolbox as well as Point Clouds Library (PCL) if you are dealing with 3D data.

## **Final Report:**

The report should be in the form of a typical conference paper. This is a great skill to have, and as we will be reading several conference papers throughout the semester you should have a good feel about the format. We will go over this in class as well.

## **Potential Project Ideas:**

Here are some sample final project ideas (you are not limited to these as they are merely suggestions as to the types of projects I'm looking for):

- A surveillance video tracking system. Perform a visual analysis of people or “things” moving in the video stream and determine: (1) “suspicious” activity, (2) recognize faces for security purposes, and/or (3) collect statistics for use, say, in a retail store.
- General object recognition system in “regular, home environments”. Think of the vision skills a personal home robot would need.
- 3D reconstruction system using stereo cameras of entire objects or entire rooms
- Real-time image mosaics from a video sequence. Think about implementing such a system on a smartphone (like an iPhone or Android), and the app would build real-time mosaics. Implement blending to “hide” artifacts.
- Finger detection/tracking (using skin analysis of pixels) in a stereo vision system. Implement some automated hand gesture detection.
- Photometric stereo setup for surface normal and depth estimation of entire objects
- Implement a face recognition system (including detection). Imagine scanning all of your Facebook photos to detect all faces and determining who each of your friends is in the photos. This might require training on samples of your friends faces.
- Similar to the previous face recognition in Facebook task, try detecting general attributes in images. Imagine I'm a Marketing Firm and I hire you to write some vision software which scans users' Facebook photo albums and automatically detects particular attributes about that user for targeted advertising purposes. Can you figure out the likes of that user with vision alone?
- Automatic detection of abnormalities in medical imagery (e.g., white light images, CT, MRI, etc, but you must obtain the data on your own AND LEGALLY!). For example, can you automatically find tumors in an image?
- Google Image Searcher. Given an image template, find other images which also contains this “thing”. Consider the speeds at which Google searches for text in the same way.
- Optical Character Recognition system: scan documents and automatically detect all of the text. Can you run this through a translator, say Google Translate, and translate text in real-time between arbitrary languages?
- Implement your own structured light system. Can you make it real-time? How many images do you need to reconstruct something in 3D?