Computer Vision EN.600.461
Fall 2016

Final Project Proposal
October 19, 2016

Joon Hyuck Choi
jchoi100

1. Topic: Optical document recognition

2. Group Members (1): Joon Hyuck Choi (jchoi100)

3. Motivation: Dynamics in college classrooms have changed drastically in the past decade. Just about ten years ago, rarely did students open their laptops in lecture to take notes—rarely did professors allow students to open their computers in class. These days, most students take notes using their laptops. Reasons range from the fact that many students these days have bad penmanship, they are used to typing on their computers, or electronic archives are easier to refer to in the future. However, some students still take notes using pen and paper. Personally, I take the traditional approach in the classroom so that I can easily draw figures presented in class, get less distracted on my computer, and just to have fun using my new ink pen. Nevertheless, it is true that sometimes, I wish I had taken notes using my computer. Can't I have the cake and eat it too?

4. Details: I will implement a program that can do mainly two things: 1) Input a photo of a document, convert it into a scanned version; 2) Input a scanned document (whether it be the output of this program or a scan from a photocopier), output a text or word document with the contents of the document typed out. Optionally: 1) recognize figures and drawings and save them separately; 2) translate document into different languages using existing APIs (e.g. Google Translate); 3) return search engine results for "important" terms that appear in the document as a list of web links.

5. Programming Language: Python or MATLAB

6. Tools: OpenCV, Google Translate API, etc.

7. Challenges: In many cases, handwritten notes have diagrams and figures that are not characters easily translatable into their electronic representations. Moreover, hand-scanned or photographed documents may have artifacts, noise, and features we are not interested in (e.g. table, hand, coffee mug, etc.). Furthermore, amongst all such noise, the document we are interested in may not be in perfect angles—we might need to apply affine transformations and/or solve for homographies.

8. Existing Applications:
- *Scanner Pro* (iPhone app): allows scanning of documents; corrects distortion and geometry; supports Dropbox, Google Drive, OneDrive, etc.; recognizes 21 different languages.
- *iScanner* (iPhone app): converts photos into scanned versions; removes noise and corrects distortion and geometry.