

EE 368 Digital Image Processing

Sonia A. Bhaskar, Nicholas Lavassar, Scott Green
Final Project Spring 2010

1 Basic Description

Our project will utilize the visual capabilities of the Android mobile phone to extract information from a business card. We will use the video and camera features of the Android to capture data. As further extensions, we would like to use this information to run a Google search which would provide relevant information. If we have additional time, we would like to try to implement photoshop editing on the business card for fun, for example, in real time, make the business card grayscale, or photoshop various features of the business card out of the frame.

Extracting information from the business card would require accurate recognition of the text of the business card, as well as recognition of the accompanying logos in order to run a Google search. The business card would be subject to several environmental conditions, such as lighting, reflection, rotation, scaling (we would desire the same data to be extracted from the business card regardless of the distance from the camera), among others. Our project will involve working with optical character recognition engines, most likely Tesseract, since it is the most accurate one currently available. We could either try to optimize the environmental conditions to make it conducive to using Tesseract, or we could try to develop an optical character recognition more suitable for our task based off of Tesseract.

2 Implementation

2.1 MATLAB

We will start out our implementation in MATLAB, since MATLAB has an easily accessible library of mathematical functions. We will begin with the simplest test objects, simple computer-generated squares with text on them, where the text and edges of the object will be easily recognizable. We'll then move to subject the object to different conditions, such as variable lighting and finer edges, and then finally move to actual images of business cards that we've uploaded from the Android phone.

2.2 DROID

Once we are successful with our MATLAB implementation, we will then move to trying to implement our project on the DROID mobile phone and try to develop the mathematical functions that may not yet exist in the Android library. This will involve beginning with still frames and then finally with moving frames, that is, capturing the business card using the video feature on the Android phone.

References

- [1] H. Li and D.S. Doermann. *Text Enhancement in Digital Video Using Multiple Frame Intergration*. ACM - Multimedia 99, Orlando, Florida, pages 19-22, 1999.
- [2] G. Zhu and D. Doermann. "Logo Matching for Document Image Retrieval," *International Conference on Document Analysis and Recognition (ICDAR 2009)*, pp. 606-610, 2009.
- [3] J. Liang, et. al. "Geometric Rectification of Camera-captured Document Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 591-605, July 2006.