EE368 Project Proposal: Mobile Chinese Translator

Team Members: Ray Chen (rchen2@stanford.edu), Sabrina Liao (liaos@stanford.edu)

Uses Android Camera Phone: Yes

Description

With the advent of more processing power in smartphones and the increasing popularity of augmented reality, technologies such as Google's Project Glass do not seem far away. At the same time, because of globalization, newer generations of mobile technology attempt to break down language barriers by implementing text translators as one of their staple apps. Current applications such as World Lens and Google Goggles do a good job in certain areas, but do not address a number of use cases.

For example, many applications require an internet connection to process images on a server because processing & translating on the phone is too slow [1]. This is especially the case with languages that use non-Roman alphabets (apps like World Lens are limited to using English, French, and Spanish). We've tried using running the Tesseract OCR Engine purely on an Android phone for Chinese characters [2], and this is in fact the case; the accuracy was low and text often took over 10 seconds to extract. Other problems include requiring the user to select the location of the text and/or having limited camera orientation.

The task of mobile text translation is usually divided into three main steps: detection, extraction, and translation[3]. In this project, we are mainly concerned with detection and preliminary filtering in preparation for extraction (the actual extraction step can be handled by open source OCR libraries such as Tesseract, and translation is an almost trivial problem with services like Google Translate). Our ultimate goal is to have a Simplified Chinese translator system which completely runs on a smartphone without requiring an internet connection. Of course, we feel that this is a bit ambitious given our time frame, so we plan to focus on the following problems (in the order of importance):

1. Implementing text detection & localization specifically for Chinese text. This is a problem due to the nature of Chinese characters, especially since the set contains over 20k characters [4]. For this, we plan to use OpenCV and to explore a variety of algorithms such as edge detection, clustering, simple template matching, etc [1,5,6,7].
2. Improve text detection with minimal required user input. This includes automatically detecting text across the entire camera image (and be robust enough for a variety of backgrounds) and detecting the orientation of the text (such as in cases where the perspective of the text is tilted) [7].
3. Speed up text detection and extraction enough to enable real time translated text overlay (i.e. without making the user take a picture and having to wait 10 seconds).

References

[1] T. Tinoco De Rubira. "Mobile Sign Translator for the Thai Language." Stanford University EE368 Spring 2011 Project.

[2] Tesseract OCR Engine. http://code.google.com/p/tesseract-ocr/

[3] V. Fragoso, J. Kleban, S. Gauglitz, S. Zamora. TranslateAR: A Mobile Augmented Reality Translator on the Nokia N900. http://vision.ece.ucsb.edu/~kleban/papers/translatAR.pdf

[4] Y. Liu. Printed Chinese Character Recognition. Computer Science Department at Massey University. http://www.massey.ac.nz/~albarcza/ResearchFiles/YuanLiu_Hon_2009.pdf

[5] X. Wang, W. Lai. "Edge detection for Chinese text image based on novel differential operator." Computer and Information Application (ICCIA), 2010 International Conference on, pp.44-47.

[6] I. Haritaoglu. Scene Text Extraction and translation for Handheld Devices. IBM Almaden Research. http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=00990990

[7] P. Shivakumara, T. Quy Phan, C. Lim Tan. A Laplacian approach to multi-Oriented Text Detection in Video. http://www.comp.nus.edu.sg/~phanquyt/papers/pami11.pdf