

Instant Camera Translation and Voicing for Signs

Shanshan Xu: xuss@stanford.edu

Kaifeng Chen: kfchen@stanford.edu

Yao Zhou: yaozhou@stanford.edu

Goals:

Develop an Android app for signs translation from English to Chinese or other foreign languages with auto voicing. Useful for sign recognition at airport or on road, menu translation while dining at a restaurant, guide translation at tourist attractions, text translation while reading a book. Help foreigners to guide and enjoy themselves when travelling in the states without the hassle of opening a digital dictionary and typing in the text. Sometimes, when typing we could easily mistake one letter or another and we want to avoid any mistake by recognizing the words directly. Just focus your camera and get the translation popping out immediately! In addition, we will also test the performance of our app on the database of signs. We'll also conduct field test of the app at airports and inside buildings.

Plans:

- Clean the image and segment the text part of signs. We'll mainly focus on sign texts which are sparse and have a fix font and distinct color difference from the background. We'll not work on the signs with missing letters.
- Pre-processing the text to suit further recognition requirement.
- Recognize the text.
- Use online recourses such as Google translate to get the translation and also obtain the voice information.
- Show translation below the original text.
- Perform various tests on our app and present statistics by evaluating our app on database and field tests.
- Iterate over the above steps by using different algorithms and improve the accuracy of the app.

Challenges:

- The texts might be hard to recognize because of unexpected reasons such as reflection from the sign plate, or tilted camera angle.
- The texts might be of different fonts or colors.

We will use an Android device. We will try to make the app easy to use. Also, to make the camera translation robust, we would test at different scenes and translation should work at a variety of situations.

References:

- 1 <https://developers.google.com/vision/text-overview>
- 2 COCO-Text: Dataset and Benchmark for Text Detection and Recognition in Natural Images, <https://arxiv.org/pdf/1601.07140v2.pdf>
- 3 Robust Scene Text Recognition with Automatic Rectification, <https://arxiv.org/abs/1603.03915>
- 4 A survey of modern optical character recognition techniques, <https://arxiv.org/abs/1412.4183v1>