EE 368 Project Proposal

Restaurant Menu Expert

Chung Yu Wang (chungyuw@stanford.edu)
Yu-Sheng Chen (yusheng@stanford.edu)
Ziyu Lang (zylang@stanford.edu)

Motivation:
When visiting foreign restaurants, people are often unfamiliar with the names on the menu. This is not only due to language barrier, but more importantly due to cultural differences. In this project, we would like to help people lower such hurdle by presenting a translated name of the dish along with the image of the dish on the side of the original dish name on the menu. Via our proposed pipeline, people may further extend the application to showing the information on a wearable VR device.

Project Description:
In this project, we plan to tackle the problem of identifying dish names in menus to provide useful contexts, such as an example picture of the dish and the translated dish name, for the diner.

We plan to implement several techniques learned in class, including denoising, binarizing images, keypoint detections, keypoint descriptions and keypoint matching. We are very curious

about the advantages and disadvantages of various keypoint detection and description techniques learned from class and research papers. As such, we will perform detection techniques based on SIFT, SURF, Shape Context and various other approaches and compare the results. After identifying dish names, to provide useful contexts, we may search for information about the dish online, but this heavily depends on the internet availability. As such, for the purpose of this project, we plan on building a database for a predefined set of dishes that are commonly seen from various parts of the world.

Pipeline:
1. For a given input image of a menu, we first conduct denoising, sharpness enhancement and thresholding to obtain a binary image of the text on the menu.
2. Conduct feature detection and generate keypoint descriptions for each character on test image.
3. Conduct keypoint matching on input characters with the training set of characters.
4. Implement commonly used natural language processing techniques, such as shortest Levenshtein distance, to resolve mismatched characters within a word.
5. Identify the translated dish name and the image for the dish.
6. Display the image and the translated dish name on the menu.

Platform:
We plan on using MATLAB and possibly OpenCV in C++.

References:
[1] Y. Amit; D. Geman, and K. Wilder. Joint induction of shape features and tree classifiers. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1997.
[2] S. Belongie and J. Malik. Matching with shape contexts. In *IEEE Workshop on Content-based Access of Image and Video Libraries*, 2000.
[3] S Belongie, J Malik, J Puzicha. Shape context: A new descriptor for shape matching and object recognition. *Annual Conference on Neural Information Processing Systems*, 2000.
[4] H Fujisawa, C.-L. Liu. Directional Pattern Matching for Character Recognition Revisited. *IEEE conference. Document Analysis and Recognition*, 2003