

EE368 Spr 2015 Project Proposal

SUNet ID: [stevenfu]

Name: [Xiaofei Fu]

Email: [stevenfu@stanford.edu]

Collaborators: [NA]

Project: Mobile face recognition and audio representation for visually impaired users

Summary:

Real time mobile face detection and recognition, then representation of face/image with sound, for visually impaired people.

Project Goal:

- Build a iPhone App for visually impaired people which can help user detecting face in realtime by recording through back/front camera. Also have the capability of classify gender based on faces.
- Another aspect of the project is find a proper way to represent images/faces with sound that represent the colour of the point of touch location within images, so that visually impaired users can have a rough idea of the shape of camera captured objects, by scrubbing finger back and forth across screen, and listen to the modulated sound output.

Stretch Goals:

- Enable face recognition with faces which exists in the build-in model dataset.
- Enable user to train the model with new faces on mobile device, add new face data to the database overtime.
- Do above tasks locally on iPhone without need of any network connection or any kind of backend server.

Work items:

- Study different face/gender detection and recognition algorithms [ref.1], implement and compare their performances in term of speed and accuracy, then chose one that runs the best on iOS platform without requiring latest generation of hardware(acceptable speed on iPhone 5 or even 4).
- Try different image processing [ref.2, ref.3] and feature extraction methods to pre-process images before feeding it to synthesised audio representation, in hope of improving usability.
- Study different ways of representing images with sound [ref. 4], chose one that is best for mobile use.
- Build an iPhone App with the study result and maybe submit to AppStore.

Platform:

The project will be build natively on iPhone with iOS SDK and Objective-C, together with openCV framework and C++ code for some of the image processing tasks that openCV already have good support of.

References:

1. Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *Acm Computing Surveys (CSUR)*, 35(4), 399-458.
2. Joung-Youn Kim, Lee-Sup Kim, and Seung-Ho Hwang. An advanced contrast enhancement using partially overlapped sub-block histogram equalization. *Circuits and Systems for Video Technology, IEEE Transactions on*, 11(4):475484, 2001.
3. Qing Wang and R.K. Ward. Fast image/video contrast enhancement based on weighted thresholded histogram equalization. *Consumer Electronics, IEEE Transactions on*, 53(2):757764, 2007.
4. Roht, P.; Richoz,D; Petrucci, L; Pun, T. An audio-haptic tool for non-visual image representation. *Signal Processing and its Applications, Sixth International, Symposium on*. 2001 (Volume:1), 10.1109/ISSPA.2001.949776