Lip Flip Application

Group Members:
Trisha Lian (tlian@stanford.edu)
Kyle Chiang (kchiang@stanford.edu)

Project Goal

In the Tonight Show with Jimmy Fallon, Fallon has a sketch segment where he swaps the mouth of him and his guest. However, the methods employed are very simple. They merely crop out a fixed portion of two camera images and swap them. Both he and the guest must keep their mouths centered in this cropped region. This often results in some ridiculous looking faces and mouths improperly cropped. The goal of this project is to create an entertaining, real-time, application which swaps the mouths of two users, each behind a webcam, in a believable manner.

Project Plan

The first step is to be able to detect and track the mouth region of each face. Our first iteration of our application will use the Viola-Jones detection framework [1], implemented in OpenCV, to find a bounding box for the face. Next, we will estimate the mouth region with respect to the face (e.g. divide the face into quadrants, chose the quadrant that most likely contains the mouth.) If this method does not produce believable results, we could also try to directly detect the mouth. One way to achieve this is to train a Haar cascade with mouth images, or we could try a lip segmentation method using color [2]. However, since we only need the mouth region, this might introduce detail that would complicate our swapping. Once we swap the mouths, we need to be able to blend the mouth realistically into the opposing user's face. We plan to try Poisson blending [3] or Laplacian blending and evaluate each techniques speed and ability to perform real time. In order to achieve more realism, we will also attempt to match the skin color of the opposing user through histogram matching. If the above steps are achieved successfully, we can attempt to implement a simple head pose estimation technique ([4] is a simple, geometric method that requires us to detect the location of several facial features, [5] is an overview of other techniques we might consider.) By treating the mouth as a plane, we can apply a perspective distortion to match the lips to the rotation of the users head.

- This project will be implemented using C++
- No Droid Camera Phone Needed

References

[1] Viola, Paul, and Michael J. Jones. "Robust real-time face detection." *International journal of computer vision* 57.2 (2004): 137-154.

[2] Eveno, Nicolas, Alice Caplier, and Pierre-Yves Coulon. "New color transformation for lips segmentation." *Multimedia Signal Processing, 2001 IEEE Fourth Workshop on*. IEEE, 2001.

[3] Pérez, Patrick, Michel Gangnet, and Andrew Blake. "Poisson image editing."*ACM Transactions on Graphics (TOG)*. Vol. 22. No. 3. ACM, 2003.

[4] Gee, Andrew, and Roberto Cipolla. "Determining the gaze of faces in images."*Image and Vision Computing* 12.10 (1994): 639-647

[5] Murphy-Chutorian, Erik, and Mohan M. Trivedi. "Head pose estimation in computer vision: A survey." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31.4 (2009): 607-626.