Final Project Proposal: Face swapping between videos

Henil Satra, Prashanth S Iyer and Parth Sanghvi

November 5, 2022

1. Project Summary

For our final project, we as a team have decided to make a project which has the capability of swapping people's faces in videos. Face swapping refers to the task of swapping faces while maintaining the rest of the body and environment context. We plan on swapping faces of individuals between two separate videos. We will make use of Computer Vision concepts like Facial Landmarking, Facial Feature Extraction, Delaunay Triangulation, Warping, Affine Transformation and Image blending. The highlight of the project would be the preservation of the target face's emotions. We also need to take care of changes in exposure, lighting and shadows because we are aiming for a seamless swap between the faces in the video streams.

2. Goals and Objectives

Our primary goal is to successfully swap faces between any two given videos with one subject face in each video. We also want to keep the target video's facial emotions to remain the same even after the swap. We want the swap to be seamless so that there is not much difference between target face and source face's external conditions like lighting and exposure. One of the major challenges we will face is when the head's orientation changes. We will have to track the target face's position at all times and compensate for the features that are lost or added due to the face's movements.

3. Related Works

In [1], the authors explain how they use continuous interpolation of the face views based on reenactment, Delaunay Triangulation, and barycentric coordinates. They use a face blending network for seamless blending of the two faces while preserving target skin color and lighting conditions. Their network uses a novel Poisson blending loss which combines Poisson optimization with perceptual loss.

[2] describes a Faceshifter method that uses a two-stage framework for face swapping which is occlusion aware. They achieve this using an encoder for extracting multi-level target face attributes. To address the challenging facial occlusions, a second stage consisting of a novel Heuristic Error Acknowledging Refinement Network (HEAR-Net) is appended.

A group of researchers from Columbia University have a written a paper [3] where in they explain their approach for face swapping using Deep Learning and openCV methods.

We have already done projects on Image blending and Image morphing which we will heavily use in our final project. Figure.1 denotes our results for these projects.

Figure 1. Outputs of Image Blending and Image Morphing project

We will have to track the face position at all times and we can use optical flow to do that. We have also performed Poisson Image blending project where in we merge a target image with a given mask.

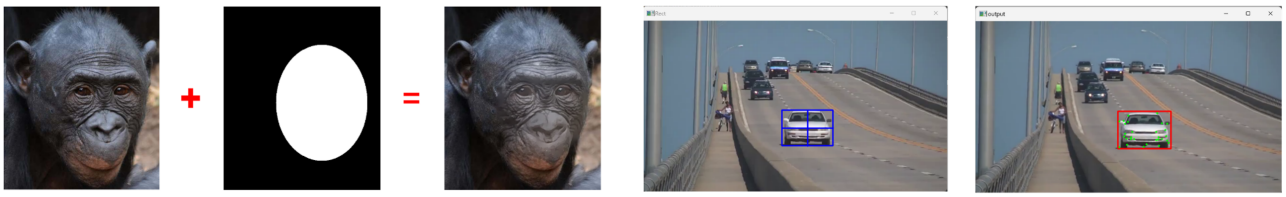

Figure 2. Outputs of Poisson Image Blending and Object Tracking project

4. Proposed approach

We will be creating a pipeline that will encompass the following steps one after the other such that the output of the previous step is the input for another.

- Select the source and target videos for swapping the two faces.
- Detect faces from both video streams and perform feature extraction. We plan to achieve this using the dlib deep learning based library which will localize and landmark the following features: eyes, eyebrows, nose, mouth, and jawline. We will then map the corresponding features between the faces in the video streams. Here, an important thing to note is that we will extract facial features only from the convex hull of the face. We do so to preserve the target face's emotions.
- Using Delaunay triangulation technique, we will select the corresponding triangles between the source and target face features and warp them so that the source frame triangles match exactly with the target frame triangles in shape and size.
- We perform blending of the faces using Seamless Cloning Technique so that the source face's properties like exposure and lighting match that of the target face. Here we also make sure that the blending is so seamless that there are no visible edges between the blend of the source image and the target frame.

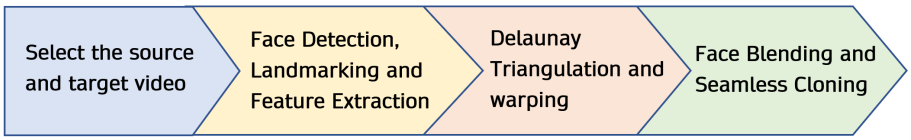

Figure 3. Pipeline depicting our project flow

5. References

[1] Nirkin, Yuval Keller, Yosi Hassner, Tal. (2019). FSGAN: Subject Agnostic Face Swapping and Reenactment.

[2] Lingzhi Li, Jianmin Bao, Hao Yang, Dong Chen, Fang. (2019). WenFaceShifter: Towards High Fidelity And Occlusion Aware Face Swapping.

[1] Bitouk, Dmitri Kumar, Neeraj Dhillon, Samreen Belhumeur, Peter Nayar, Shree. (2008). Face swapping: Automatically replacing faces in photographs.