3D-Model-based Face Replacement in Video

Virginia Tzeng*, Yi-Ting Cheng*, Yu Liang*, Chuan-Chang Wang[†], Bing-Yu Chen*, Yung-Yu Chuang*, Ming Ouhyoung*
*National Taiwan University   [†]Next Media Animation Limited

1. Introduction

The development in digital technologies and the widespread Web 2.0 concept have made many digital videos accessible. Editing and modifying digital videos have become an interesting and important topic. In this paper, we present a system for face replacement in video. Most digital processing software can perform face replacement only when the poses for the source and target faces are similar, and the manipulation process with those software is often time-consuming and labor-intensive. While previous work [Blanz et al. 2004] focuses on image face replacement, our system performs face replacement in video by constructing 3D models for both target and source faces and swapping them accordingly. 3D face models are created by fitting 3D morphable models [Blanz et al. 1999] and the input is reduced to two pictures for the face to be placed in.

The system is useful in many ways: for example, one may replace a leading character's face in the movie with her/his own. In addition, the face model can be altered in user's favor and then replace the target face in video.
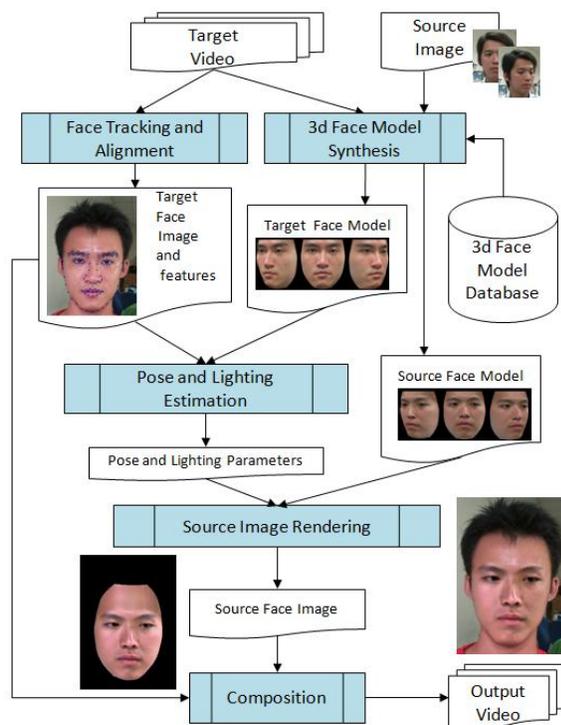
2. System Overview and Implementation

Figure 1: System overview for our "face replacement in video".

Our system has four main stages (see Figure 1). First, we create 3D face models for the source and target faces separately. 3D face models are created from a couple of images by adopting 3D morphable model fitting method [Blanz et al. 1999]. The source face model takes two pictures as input: one is a frontal view, and the other is a profile view. For the target face model, we pick an appropriate frame from target video for constructing the target model.

Second, we track and align the target face in the video to get feature points of the face, such as eye corners, mouth corners, etc. The feature points along with target face model are used to estimate the face pose parameters. By using the pose parameters and the normal vector of face model, we estimate nine spherical harmonics parameters as lighting parameters for the environment of the target video.

Third, we use the estimated pose and lighting parameters to rotate and relight the source 3D model accordingly and to render the corresponding frames.

In the final stage, we use feature points to determine the replacement region on the face. Then, we replace the region of the target video with the one from corresponding source frame. Although we have adjusted the lighting of source face, it is still possible that the colors of source and target faces are slightly different. Therefore, we use Poisson blending to obtain better compositing results.

3. Conclusion and future work

We have created a system for face replacement in video. Given the target video and two pictures of the face to replace with, we can replace the target face in video by a small amount of user inputs: setting the initial guess for face alignment in the first frame (translate and scale a face template to roughly align the face) and setting 11 feature points for both source and target faces (it is only required for three images: two source image and one frame from the target video). In terms of computation, it takes about 2 minutes to construct the 3D model for target and source faces and 6 seconds to replace the face for each frame. Results are shown in our demo video.

Currently, our system only works on neutral face in video due to the limitation on 3D face model database. In the future, we plan to capture or create a richer 3D face dataset of facial expressions, so that we can estimate expression parameters and replace the face with an expression similar to that of the target face.

Reference
Blanz V, Scherbaum K, Vetter T, and Seidel H.P, "Exchanging Faces in Images," *Proc. EUROGRAPHICS*, pp. 669-676, 2004.

BLANZ V., VETTER T.: A morphable model for the synthesis of 3D faces. In *Computer Graphics Proc.SIGGRAPH'99* (Los Angeles, 1999), pp. 187–194.