Face Replacement with Large-pose Differences

Yuan Lin[1], Qian Lin[2], Feng Tang[2], Shengjin Wang[1]
[1] Dept. of Electronic Engineering,Tsinghua University
[2] USA HP Labs, Multimedia Interaction and Understanding Lab
{linyuan,wsj}@ocrserv.ee.tsinghua.edu.cn, {qian.lin,feng.tang}@hp.com

ABSTRACT
In this paper, we present a novel face replacement system exchanging faces with large-pose differences. Traditional 2D image based face replacement can only replace faces with similar pose and appearance. This significantly limits the application of face replacement. In this paper, we propose to build a 3D head model from a single frontal face photo. The automatically constructed 3D head can be rendered under arbitrary poses and illuminations. This makes it possible to do swapping for faces with large pose variations. In the demo, the user captures a frontal face image using a capture device such as a webcam or a smartphone, and then the algorithm can automatically build the 3D model using feature detection, face alignment and reconstruction. This 3D model is used to swap to any other target face photo the user selects. While our system is automatic, we also provide interactive tools for the user to adjust the feature detection to enhance the results.

Categories and Subject Descriptors
I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism -*Virtual reality*; I.4.9 **[Image Processing and Computer Vision**]: Applications

General Terms
Design, Experimentation, Human Factors

Keywords
3D Head Reconstruction, Face replacement, Color Transfer, Multi-resolution Spline

1. INTRODUCTION
The face replacement, called face swapping as well, exchanges certain face of the target image with the input face image. It can be used in many applications, such as personal entertainment, face de-identification, replacement the substitute's face with that of actor, et al.. There are some websites for the parents to make their children personal storybooks by providing the children's photos. Some artists will replace the face of the character from the chosen story with that of the child.

These years many attentions have been given to the 2D face replacement techniques. These methods swap faces through image editing algorithms such as image warping. Some examples include [1] which swaps faces with most similar faces from an image database and unsupervised face alignment by robust nonrigid mapping [5]. These methods are based on mostly 2D image blending, the requirement is the source face and target face should be in similar pose. These restrictions limit the applications.

In this paper, we propose a system for face replacement using a 3D model based approach. A 3D head model is built by a single frontal face photo uploaded by the user. The 3D head can be rendered into virtual faces with any required pose. Thus, the pose restriction of the traditional face replacement can be eliminated. The previous exchanging face method based on 3D face model [5] generates the face model with 3D Morphable Model, which estimates shape, texture, pose and illumination condition simultaneously. The process is liable to trap into local minimum and quite time-consuming, which on the average takes 4.5 mins on a 2GHz Pentium 4 processor. Compared with [5], the 3D face reconstruction here is done by deforming a generic head model, which is quite fast. The reconstruction process on the average takes 1.8 seconds on 3.2GHz Intel Core i5 processor. After the virtual source face image with the target face pose is rendered, the 2D face replacement methods can be used for the replacement.

2. THE PROPOSED APPROACH
To exchange faces with large-pose variations, 3D face model is introduced. In our system, only one single frontal view image is required. The 3D reconstruction is done by deforming a textured generic 3D head model. The input frontal view image goes through face detection, face alignment, and face feature points matching with the 3D generic head model. The deformation tries to locate the 3D feature points with the same distribution as that of the input image. Then other non-feature vertices are refined by an interpolation method based on the feature points' displacements and their original distributions on the 3D generic head model. Radial basis function is used for the interpolation. After the shape reconstruction, the reconstructed head model is projected on the input image to extract the face region texture. Then this face texture will replace that of the generic texture. The color transfer [2] and multi-resolution spline technique [4] are adopted to ensure a seamless and satisfied image blending. Mapping the texture on the shape model constructs the final realistic head model. Details of the reconstruction are in [6].

As the first step of the system, pose estimation is applied for the target face. After that, the 3D head model with the estimated pose is rendered, named virtual source face image. At the same time, the mask of the whole head region on the image is generated as well. Three easily detected points: two eye center points and one mouth center point are selected as the fiducial points used as anchor points for the swapping. These three points on the generated virtual source face image can be obtained automatically by 3D head projection. These three pairs of points are used to map the virtual source face image and the target image into a common coordinate system. Here the virtual source face image is transformed into the target image coordinate by rigid

MM'12, October 29-November 2, 2012, Nara, Japan.
ACM 978-1-4503-1089-5/12/10.

transformation fitted from the three point correspondences. The source mask will be transformed in the same way as well. The mask of the target face needs to be labeled offline, but this process only needs to be done once for all the source images. Then the intersection of these two masks is selected as the region for swapping. Before the face replacement, the virtual source face image is color-adjusted with color transfer [2] to reduce the color and illumination differences between the inputs. To make the face replacement seamless, the multi-resolution spline technique [4] is adopted. More details are given in [6]. The flowchart of our proposed system is shown in Figure 1.

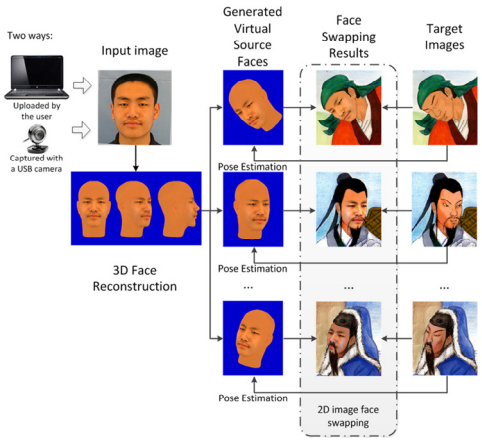

Figure 1. The framework of our proposed system.

3. SYSTEM DEMONSTRATION

Figure 2 shows the user interface of the proposed system. In part A, our user can either upload one frontal view image or capture one from a capture system such as a webcam or a smartphone, and then automatic face alignment is applied to locate the feature points. A user assistance face alignment refinement modul allows the user adjust these points. The face alignment and face feature point matching results are shown in part D. The reconstructed 3D head model is shown in part B. Then the user will choose one target image, shown in part C, and pose estimation of the target face is applied. The 3D head is transformed to the estimated pose by user assistance. By clicking the virtual image generating button, the virtual source face image, the related mask and the fiducial points file are generated simultaneously. Then the face replacement is done automatically. The result is shown in part E.

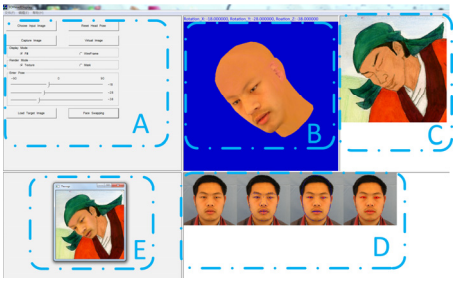

Figure 2. The user interface of our proposed system.

Since the ground truth shape of the input face is not available, only subjective evaluation results can be given. We compared our results with a commercial software by reconstructing ten different subjects based on his/her frontal face image. Ten testers are asked to select the favorite 3D head generated by either our method or the software. The statistic of the user rating is given in Figure 3 (left). For 80% cases, our system performs better than or equal to the software. Ten face swapping results are evaluated subjectively by asking fourteen testers to score them with 5, 4, 3, 2, 1. 5 points stands for the wonderful performance, and 1 point stands for the poor performance. The average score for every result is shown in Fig. 3 (right). The total average score for all the results is 3.8 points, which shows our system is quite impressive.

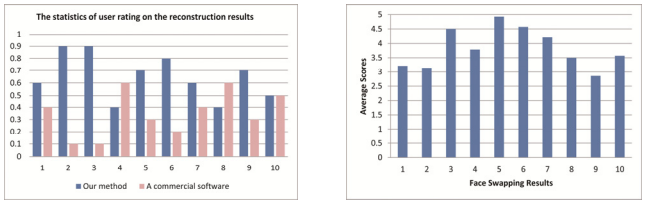

Figure 3. The subjective user evaluations for the 3D face reconstruction and the proposed face swapping system.

4. CONCLUSIONS

In this demo, we present a face replacement system using one single frontal image generated 3D head model to exchange input faces with large-pose differences. By detecting face features of the source face, a generic model is deformed to reconstruct the 3D head model of the subject. With the aid of the 3D head model, the target faces with large-pose variations from the source image can be exchanged as well. It cannot be realized by the traditional 2D based face replacement. The adoption of the 3D head model eliminates the pose similarity restriction, which extends the face replacement applications.

5. ACKNOWLEDGMENTS

This work was supported by USA Hewlett-Packard Development Company and the National High Technology Research and Development Program of China (863 program) under Grant No. 2009AA11Z214 and the National Natural Science Foundation of China under Grant Nos. 61071135 and the Doctoral Fund of Ministry of Education of China under Grant No. 20090002110077.

6. REFERENCES

[1] D. Bitouk, N. Kumar, S. Dhillon, P. Belhumeur, and S. K. Nayar. Face swapping: automatically replacing faces in photographs. In *Proc. of ACM SIGGRAPH*, 2008.

[2] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley. Color transfer between images. *Computer Graphics and Applications, IEEE*, 21(5):34 –41, 2001.

[3] J. Zhu, L. V. Gool, and S. Hoi. Unsupervised face alignment by robust nonrigid mapping. In *Proc. of ICCV* 2009.

[4] P. J. Burt and E. H. Adelson. A multiresolution spline with application to image mosaics. *ACM Trans. Graph*., 2:217–236, 1983.

[5] V. Blanz, K. Scherbaum, T. Vetter, and H.-P. Seidel. Exchanging faces in images. *Comput. Graph. Forum*, 23(3):669–676, 2004.

[6] Y. Lin, Q. Lin, F. Tang, and S. Wang. Face swapping between inputs with large pose variations. In *ICME*, 2012