

Automatic Face Editing on a Morphable Model

Han Li

Department of Computer Sciences
University of Wisconsin-Madison

hli337@wisc.edu

Wentao Wu

Department of Computer Sciences
University of Wisconsin-Madison

wwu69@wisc.edu

Abstract

With the unprecedented and exponentially increasing of web images and videos, there is a growing concern of online iconographic privacy. Thus, editing and rewriting these image-based media resources is becoming an interesting and urgent task. In this project, we are planning to implement automatic face editing in a resource (image or video) based on face morphing and blending from an existing facial data collection. The anticipated procedure consists of three steps: (1) face alignment based on Active Shape Models and facial expression recognition, (2) face morphing by the 2D morphable model, and (3) face blending with color and illumination adjusting.

1. Introduction

With the unprecedented and exponentially increasing of high-definition web images and videos, online iconographic privacy is becoming a severe problem and concerned by more and more people. For example, online systems such as Google Street View allow users to browse photos of public images possibly containing many people who might not consent to be photographed. Thus, it becomes more and more urgent and popular to develop face editing techniques to alleviate this problem. In this project, we are planning to implement a face editing method based on a 2D morphable model for image and video facial information hiding. Given a set of facial image collection containing faces of different poses and expressions and a target resource (image or video), Fig. 1 shows the pipeline of our procedure: (1) face detection and alignment based on Active Shape Models integrating the facial expression recognition procedure; (2) face morphing by the 2D morphable model, and (3) face blending with possible color and illumination adjusting. Since the many of the previous methods haven't cared much about the facial expressions, the contribution of this project would be integrating the facial expression recognition into the face editing procedure and substituting a new face with a similar expression.

2. Related Work

Our project is most related to Min's work [5]. They proposed an automatic face replacement approach in video based on 2D Morphable model. This approach includes three main modules: face alignment, face morph, and face fusion. They use the Active Shape Models (ASM) for face alignment. They also consider color and lightning adjustment to keep consistent. Their approach is fully automatic without user interference, and generates natural and realistic results. Furthermore, as they use 2D model, their algorithm is also efficient. However, they don't consider facial expression. The tolerance to pose and expression variance is limited by ASM.

Liang [4] proposed a video face replacement system which allows replacing target human face from target video with source face in source video. For each target face in target video, they select the best candidate face with similar facial expression. Finally, they blend candidate replacement to target video. In their approach, however, they require the target video should have similar face expression and pose with the source video.

Similar to Liang's work, Dale [3] also proposed an algorithm to provides face replacement in target video from source video. They tracks both faces in target and source video using multilinear model. By the tracked 3D geometry, source face is warped to target face in every frame of video. But, their tracking algorithm is based on optical flow, so the light should change slowly in the video.

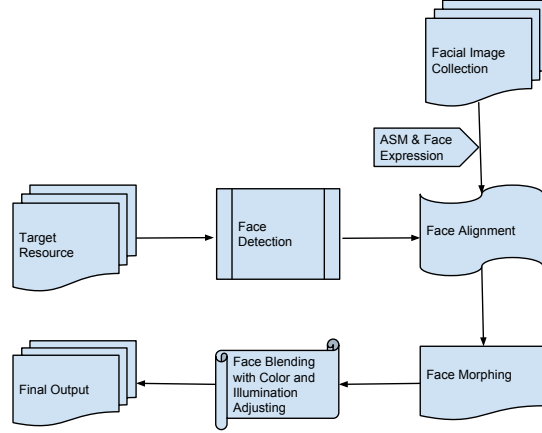


Figure 1. The pipeline of the face editing procedure.

Afifi [1] presented a system for video face replacement that requires only two videos of a source actor and a target actor using only a single digital camera. They could generate realistic results without using special equipment and 3D model. Also, they use a new face blending technique based on poisson blending. But their algorithm only works for fixed pose, i.e. front face.

Instead of using source video, Cheng [2] uses only two face pictures: one frontal view and one profile view. First, they use these two images to construct 3D face model. Then, they track the faces in the video and project the source face model to align and replacement the faces in target video. However, the tolerance to the pose is limited to robustness of their alignment algorithm.

A. Niswar [6] proposed a system where one image is required to replace the faces in target video. Also, the image is not limited to a specific pose. There are four steps in their approach: 3D face reconstruction, 3D face animation, feature points tracking, 2D projection with blending. Their approach, however, has high complexity and hence high time consuming.

3. Problem Description and Techniques to Be Used

The basic steps of the method have been outlined in Section 1. Given a face image collection¹ $F = \{f_1, f_2, \dots, f_n\}$ with n faces of different poses and facial expressions and a target resource with a list of face sequence $L_{F,old} = \langle l_1, l_2, \dots, l_m \rangle$ with m faces that would be recognized by the facial detection method, our goal is that for each $L_{F,old}(i)$, we would find the best face $f_{best(i)} \in F$ such that after performing morphing and blending, the new face $L_{F,new}(i) = Blend(Morph(L_{F,old}(i)))$ looks natural to human judgements. The methods and techniques used during the procedure would be face detection, alignment, morphing, blending model, and color and illumination adjusting methods.

4. Experimental Evaluation

We plan to focus our attention on image face editing firstly and then extend the method to edit videos. The source face images would be captured manually with many different poses and expressions, and the target resources would be captured from the Internet. Since there is no good quantitative evaluation method, we would evaluate the performance of our model qualitatively based on human judgements.

5. Conclusion and Discussion

Han Li: focus on face detection and alignment.

Wentao Wu: focus on face morphing and blending.

We plan to use one week for related methods understanding, two weeks for face editing implementation, and one week for debugging and improving. We would provide demos for the performance evaluation.

¹this set of images could be captured easily.

References

- [1] M. Afifi, K. F. Hussain, H. M. Ibrahim, and N. M. Omar. Video face replacement system using a modified poisson blending technique. In *Intelligent Signal Processing and Communication Systems (ISPACS), 2014 International Symposium on*, pages 205–210. IEEE, 2014. 2
- [2] Y.-T. Cheng, V. Tzeng, Y. Liang, C.-C. Wang, B.-Y. Chen, Y.-Y. Chuang, and M. Ouhyoung. 3d-model-based face replacement in video. In *SIGGRAPH'09: Posters*, page 29. ACM, 2009. 2
- [3] K. Dale, K. Sunkavalli, M. K. Johnson, D. Vlasic, W. Matusik, and H. Pfister. Video face replacement. *ACM Transactions on Graphics (TOG)*, 30(6):130, 2011. 1
- [4] Y. Liang, B.-Y. Chen, Y.-Y. Chuang, and M. Ouhyoung. Image based face replacement in video. 1
- [5] F. Min, N. Sang, and Z. Wang. Automatic face replacement in video based on 2d morphable model. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 2250–2253. IEEE, 2010. 1
- [6] A. Niswar, E. P. Ong, and Z. Huang. Face replacement in video from a single image. In *SIGGRAPH Asia 2012 Posters*, page 12. ACM, 2012. 2