Face and photograph augmentation based on a desired theme

Orly Liba

Department of Electrical Engineering at Stanford University

Introduction

Face and photograph augmentation are a form of expression, communication and entertainment. Applying a mask on one's face is surprising and fun and may also be considered artistic. **This poster presents a technique to make any photographed or animated face into a mask and an automated algorithm to apply masks on top of faces.** To complete the resulting image, it is possible to apply gradual blur and vignetting and change the color temperature of the photograph. **Nine masks were created for this project, which are also available for a live demonstration.**

Related work

Face morphing was demonstrated more than 20 years ago, most notably in Michael Jackson's black and white music video. More recently, companies such as Google and Baidu have created tools for automated photograph and video augmentation. These are mainly used for live chats by continuously applying a mask on one's face. Automation was enabled in recent years owing to advanced machine learning algorithms which allow exact detection of dense facial landmarks.

It is reasonable to assume that these products use similar steps to the ones demonstrated in this project, which are:

1) Facial landmark detection (and/or tracking)
2) Warping of a mask to fit the face
3) Blending of the mask with the original image

a) Images showing the morphing done in Michael Jackson's clip [1]. Right and left are the originals and center is their mixture.
b) Google hangout mask
c) Baidu FaceYou morphing application
d) Google's halloweenify (was available in 2014)

Technique

Mask Preparation

We start from an ordinary photograph, for example, a cat. Two gray scale maps are created from the photograph, using paint or Photoshop. These maps indicate the amount of blending of the mask onto the target photograph. The maps can be smoothed to allow gradual blending. Next, feature points are marked on the mask. This can be done automatically for people, but has to be done manually in other cases. For this project, Clandmark [2,3] facial landmarks were used (68 points). For placing ears (or a hat), we record a reference point (between the eyes) for positioning and the width of the face for scaling.

Processing the target photograph

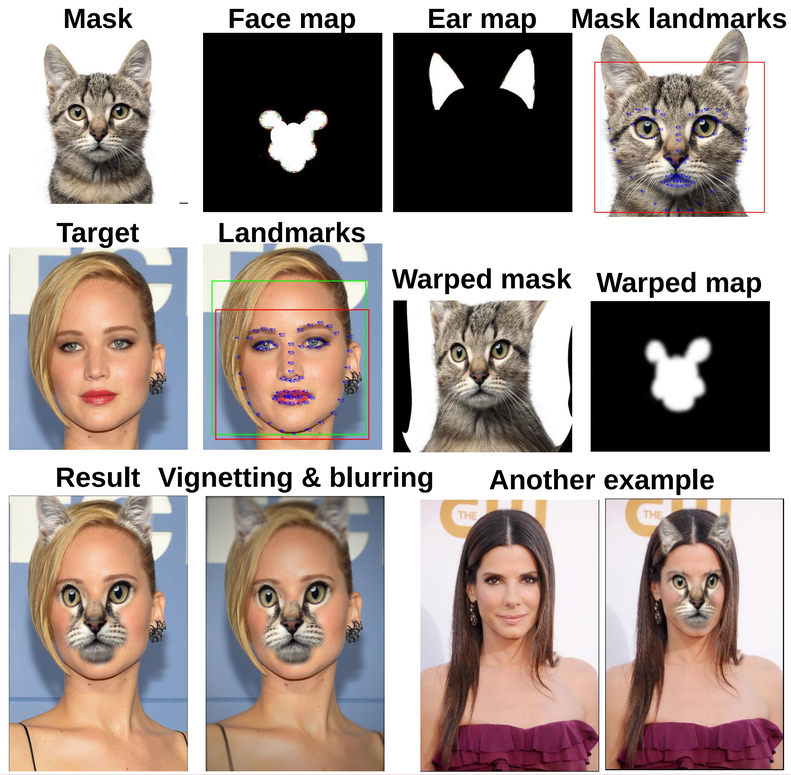
First, faces are detected in the photograph using Haar Cascades with OpenCV. For each face, facial landmarks are detected with Clandmark. Next, the mask and the face-map are scaled and warped according to the facial landmarks. Warping was done with locally weighted means [1]. The warped mask is alpha-blended with the target photograph according to the weights of the warped face-map. For some of the masks, a map was created dynamically to allow "holes" for the eyes. This was done in the skeleton and clown masks.

In order to place the ears (or hat) correctly, the position, scale and angle of the target face were calculated from its landmarks. The ear-map, after scaling and rotation, was used to alpha-blend the ears onto the target image.

Additional effects

To complete the photograph, several effects can be applied: vignetting, gradient blur and changing the color temperature of the photo. These effects are commonly used in services such as Instagram.

The full process takes 2-10 seconds, depending on the size of image and number of faces.



Results

In general, the algorithm works very well and very quickly. However, it fails when the facial landmarks are detected incorrectly on the target image. This happens when the head is at a large angle and when the main features of the face are obscured. In addition, the resulting image may have artifacts when warping is very uneven.

Bellow are examples of the nine masks that were created for the project:

Summary & Discussion

Face and photograph augmentation were demonstrated on a variety of images and masks. As a next step, this technique could be implemented on a mobile device. Increasing the speed of the processing would allow real-time augmentation for video and chat. While facial landmark detection is usually very exact, additional improvement should be done to detect landmarks on rotated and partially obscured faces, such as by a beard or bangs.

Acknowledgements: A big thank you to Michal Uřičář for making Clandmark an open source project. Additionally I would like to thank prof. Gordon Wetzstein and the EE368 course staff.

References: [1] T. Beier and S. Neely, Feature-based image metamorphosis, ACM SIGGRAPH Computer Graphics (1992). [2] M. Uřičář , V. Franc, V. Hlavac , Facial Landmark Tracking by Tree-based Deformable Part Model Based Detector, Proceedings of IEEE International Conference on Computer Vision (2015). [3] M. Uřičář , et. al., Real-time multi-view facial landmark detector learned by the structured output SVM, Biometrics in the wild (2015).