Automatic Procedure to Fix Closed-Eyes Image

Hung Vu
University of California, Berkeley

Abstract

This paper proposes an automatic procedure to fix closed-eyes image by seamlessly copying the eyes from another images. First, it selects the matching features points in the input images. Then, it warps the shape of the opened-eyes face to the shape of the closed-eyes face. Next, it creates a binary mask for the eyes region. Finally, it blends the warped opened-eyes face with the original closed-eyes image. This procedure yields convincing results under different light conditions and small pose changing, as long as the copied eyes have similar color to those of the closed-eyes person.

Keywords: closed-eyes, face alignment, ellipse, blending

Concepts: Computing methodologies → Computational photography; Image processing; Object detection; Feature selection;

1 Introduction

Taking photo of people is a great interest for both amateur and professional photographers. However, it could be a challenging task, especially for group photo, since some of the subjects might accidentally close their eyes at the wrong moment due to flash light, anxiety, etc. Such unexpected moment could ruin an amazing photo and embarrass the subject without any notice until later on. There is a recent work on detecting eyes closeness from still images [Song et al. 2014], which help people eliminate these photo before sharing a huge trip album with friends and families. But fixing those photo is better than throwing them away as they might contain beautiful landscapes or unique memories. Therefore, this paper propose an automatic procedure to fix closed-eyes image by seamlessly copying the eyes from another image. The user only needs to find another opened-eyes image with similar skin and eye color, not necessary belong to the same person. Figure 1 shows the pipeline for this process:

1. Select the matching features from the input images.
2. Warp the shape of the opened-eyes face.
3. Create a binary mask corresponding to the eyes region.
4. Blend the warped image with the closed-eyes image.

Figure 1: Pipeline to Fix Closed-Eyes Image

2 Feature Points

Since the eyes are windows to the soul, they are usually the first places people look at when they see a photo. Thus, manually selecting the feature points is not feasible since it is time consuming and prone to error, which is easy to detect no matter how small it is. Figure 2 compares the results of manual selection and automatic selection, the first one is suboptimal with one eye significantly smaller than the other.

The automatic process first finds a rough rectangular region for the face using the Haar feature-based cascade classifier, it is a machine learning based approach where a cascade function is trained from a lot of positive and negative images [Viola and Jones 2001; Lienhart and Maydt 2002]. OpenCV library provides an implementation of this method along with the training procedure and a pre-trained classifier for the frontal face [] (see figure 3).

After getting the region of the face, it employs a state-of-art method
proposed by Kazemi and Sullivan, using an ensemble of regression trees, to estimate the face’s landmark positions directly from a sparse subset of pixel intensities \[2014\]. Dlib library supports this feature by providing a 68-landmarks predictor [] trained on the HELEN dataset [Sagonas et al. 2016]. Since the eyes in the opened-eyes and closed-eyes images are different, all the landmarks in the eyes region are removed, except the those corresponding to the edges of the eyes, as shown in figure 4.

This automatic process uses two publicly available libraries to select 60 pairs of corresponding feature points from the input images in real time.

3 Face Warping

Changing the appearance of the face is a canonical example of local deformations, where different amounts of motion are required in different parts of the image. Szeliski describes a mesh-based approach to warp face: first triangulate the set of feature points, then use an affine transformation model to warp each triangular region of the face [2011].

An affine transformation \(1\) has six degrees of freedom and preserves parallelism. Thus, two corresponding triangles with three pairs of corresponding vertices would yield six equations (horizontal and vertical coordinates) and uniquely define an affine matrix.

\[
\begin{bmatrix}
a & b & c \\
d & e & f \\
0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
1 \\
\end{bmatrix}
= 
\begin{bmatrix}
x' \\
y' \\
1 \\
\end{bmatrix}
\] (1)

After recovering the affine matrix, the triangular region in the opened-eyes face is warped to the corresponding region in the closed-eyes face using inverse warping and interpolation. Repeating this process for each region in the mesh would warp the whole face, as shown in figure 5.

4 Eyes Mask

Manually create a good binary mask for the eyes region is a quick and simple task. However, in order to automate the whole pipeline, this paper proposes to create the mask by fitting an ellipse (figure 6) to the edges of the eyes \(\vec{p}_1\) and \(\vec{p}_2\), found in the earlier step.

\[
\vec{p}_c = \frac{1}{2} (\vec{p}_1 + \vec{p}_2) \\
2a = \|\vec{p}_2 - \vec{p}_1\| \\
2b = (2a) \times r \\
\theta = \text{atan2} (y_2 - y_1, x_2 - x_1) \\
\]

The binary mask (figure 7) is created by testing whether each point on the face is inside the ellipses, using the following inequation \(6\) [Kalman 2008], assuming that \(\vec{p}_c = (h, k)\).

\[
\frac{(x - h) \cos \theta + (y - k) \sin \theta)^2}{a^2} + \frac{(x - h) \sin \theta - (y - k) \cos \theta)^2}{b^2} \leq 1 \\
\]

\[
((x - h) \cos \theta + (y - k) \sin \theta)^2 \\
\]
5 Images Blending

Three decades ago, Burt and Adelson have proposed the multiresolution blending technique using Laplacian pyramid [1983]. It is based on the idea that the low-frequency color variations should be smoothly blended, while the higher-frequency textures should be blended more quickly to avoid the "ghosting" effect. The source images are first decomposed into Laplacian pyramids. Then each band is multiplied by a smooth weighting function, generated by constructing a Gaussian pyramid of the binary mask. The sum of these two weighted pyramids (equation 7) is then used to construct the blended image.

\[
LS(i, j) = GR(i, j) \times LA(i, j) + (1 - GR(i, j)) \times LB(i, j)
\]  

However, this method would yield suboptimal result in the case the input images have different brightness (figure 8). On the other hand, the Poisson blending technique [Pérez et al. 2003] does not have this problem since it tries to preserve the gradient while blending. In fact, it uses least square approach to solve for the pixel values \(\vec{v}\) in the following optimization problem (8): minimizing the difference between the gradients of the foreground region and the gradients of the target region.

\[
\vec{v} = \arg \min_{\vec{v}} \sum_{i \in S, j \in N(i) \cap S} ((v_i - v_j) - (s_i - s_j))^2 
+ \sum_{i \in S, j \in N(i) \cap \neg S} ((v_i - t_j) - (s_i - s_j))^2
\]  

6 Results

This procedure yields excellent results with different light conditions and small pose changing (figure 9). However, it would fail when the pose is too far from the frontal pose or the eyeglasses reflect light because the automatic techniques of selecting feature points would return incorrect points in those cases (figure 10).

An advantage of this procedure is that it does not require the opened-eyes image from the same person as the closed-eyes image. Since it only copies a small region around the eyes, it would yield plausible result as long as the eyes and skin color are not too different from the target person. Figure 11 shows some examples of using the author’s opened-eyes image to fix images in the Closed Eyes in the Wild dataset [Song et al. 2014].

7 Summary

This paper has proposed an automatic procedure to fix closed-eyes image using the eyes from another images with opened-eyes. First, it selects the matching features points in the input images using OpenCV and Dlib libraries. Then, it warps the shape of the opened-eyes face to the shape of the closed-eyes face with the mesh-based affine transformation approach. Next, it creates a binary mask for the eyes region by fitting ellipses to the edges of the eyes. Finally, it blends the warped opened-eyes face with the original closed-eyes image using the Poisson blending technique.

This procedure yields convincing results under different light conditions and small pose changing, as long as the copied eyes have similar color to those of the closed-eyes person. Even though not shown in the paper, this procedure can be easily extended to fix multiple closed-eyes people in the same image.

References


OPENCV. Haar feature-based cascade classifier for object detection. [http://docs.opencv.org/2.4/modules/objdetect/doc/cascade_classification.html](http://docs.opencv.org/2.4/modules/objdetect/doc/cascade_classification.html).


