12-01

15th IAPR International Conference on Machine Vision Applications (MVA)
Nagoya University, Nagoya, Japan, May 8-12, 2017.

Exemplar-Based Human Facial Features Cloning

Damon Shing-Min Liu
National Chung Cheng University
168 University Road
Chiayi, Taiwan
+886-5-2720411 EXT33118
damon@computer.org

Feng-Yi Lin
National Chung Cheng University
168 University Road
Chiayi, Taiwan
+886-5-2720411 EXT33153
nevergiveup79524@gmail.com

Abstract

In recent years the spread and easy availability of drug led to drug abuse; consequently drug trafficking gradually increased. In order to enable people to understand how their appearance becomes after having taken drugs, we propose a system on cloning human facial features based on example images. Our goal is to clone the facial features on an example image to a source image without changing the identity of such source image, making source and example image have the same visual effect. In our system, user first provides a source and an example image, and then we use Active Shape Model to distinguish between the skin and facial component areas, keeping only the skin area. Afterward, we apply seamless cloning to blend the facial features on example image to source image, making two images have the same facial features and visually look similar. Additionally, we also propose an automatic selection of a good example image to solve the problem if skin color between the source image and example image differs greatly, making the result image achieve better visual effects.

1. Introduction

Human face contains a wealth of information, including individual's identity, age, health condition, and emotion. During the last few decades with the vigorous development of computer vision and machine learning techniques, facial feature simulation, face detection and face recognition have been widely applied in real world, such as face aging simulation [1], face makeup simulation [2] and style transfer for headshot portraits [3]. Contrast to the aforementioned applications, we propose an application on drug addicts facial features simulation. In recent years, the spread and easy availability of drug led to drug abuse. In order to enable youth and students to understand how their appearance becomes after having taken drugs for a while so as to achieve a deterrent effect, we proposed a cloning human facial feature by examples technology. This technique of cloning the facial features, such as skin texture, wrinkle, fester and dark circle shown on example image, makes the faces on source image and the example image have a similar visual effect.

2. Related Work

Facial feature transfer techniques can be classified into two groups. The first group is using *training-based* techniques to obtain the facial feature transfer. The second group is adopting *exemplar-based* as facial feature transfer. In this section, we survey related works and discuss relevant topics on facial feature transfer. For training-based facial feature transfer, Suo et al. [1] presented a dynamic model to describe face aging process. They manually differentiated different age groups and labeled hairs, facial components and wrinkles in their dataset. After that, they performed a machine learning technique onto a large dataset for each individual age group to produce a dictionary. During facial feature transfer, they employed a high resolution grammatical face model [4] to determine the facial components of input image, and found a similar shape of facial component from dictionary. At last, they adopted a merging technique to transfer the facial features and facial components to the correct location. But they used a training approach to find the facial feature is time-consuming, which also needs a lot of sample images. For exemplar-based facial feature transfer, Guo et al. [2] presented a work on digital face make up by examples. They first decomposed both source image and example image into three layers: face structure layer, skin detail layer and skin color layer. Afterward, they performed operations for each layer on such source and example image. Finally, they combined three layers from source and example image, and get a result image. Their method can completely distinguish between face structure and skin area on example image, but it cannot process personal characteristics, such as beards and bangs, on skin area.

3. Cloning Facial Features

In this section, we introduce the process of cloning human facial features by examples. Our work is divided into three parts: (1) preprocessing of example images; (2) face alignment; and (3) image cloning. We elaborate our main techniques in a step-by-step manner.

3.1. Preprocessing of example images

Here we discuss the procedure of preprocessing of example images. We use techniques for object segmentation, facial components detection and skin color detection to acquire the area of facial features on the example image, and store the result image that consists of detected facial components and detected skin color to a data set.

First, we adopt Grabcut [5] to perform object segmentation, and allow user to extract the region of face from an example image.

Second, in order to preserve the region of facial features, we implement a facial component detector to find the human face region on example image, and remove the facial components, leaving only the facial skin texture. The steps of facial component detection are as follows. We first adopt Active Shape Model (ASM) [6] to find the facial feature points on an example image, and use Convex Hull [7] to find the outermost contour of such face. We show in Figure 1 the result after using ASM and Convex Hull.
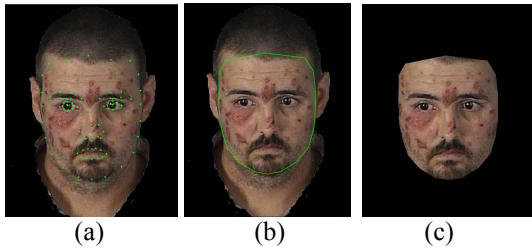

(a) (b) (c)

Figure 1. Result images after using Convex Hull and Active Shape Model (ASM). (a) The result after using ASM. (b) Adopting Convex Hull to find the points on face contour and connecting them together. (c) Excluding area outside the Convex Hull, retaining only the face area.

However, we cannot adopt this same way to find the facial component contours for eyes, nose and mouth; thus, we use other alternative to solve this problem. We recognize the characteristics of ASM that feature points are ordered in a fixed sequence; we therefore find the serial numbers for facial components feature points and connect those points together. Afterward, we determine on which pixel sitting on the facial component contour, and remove those pixels inside the contour, making the result image retain only the facial skin texture. Figure 2 shows the result images.
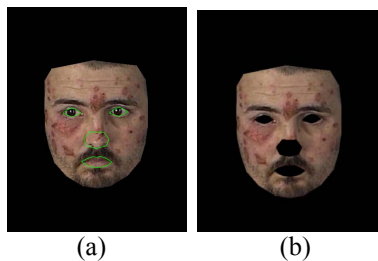

(a) (b)

Figure 2. The result images of facial component detection. (a) Adopting ASM to find the facial

components contour. (b) Excluding the facial components, leaving only skin area.

Third step is a skin color detection for excluding possible non-skin color regions, such as eyebrows, bangs and beards, and for preserving facial skin texture on facial skin area. In our approach, we fully exploit three different color spaces, including RGB, YCbCr and HSV color space, to decide on the facial skin color region. In RGB color space, we refer to the human skin color threshold previously set by Peer et al. [8] and alter their threshold value for skin color in order to retain facial features of *dark* color (i.e., usually the color for eyebrows, bangs and beards). These modified RGB facial skin color thresholds at uniform daylight illumination are defined as follows:

$$(R > 85) \text{ AND } (G > 30) \text{ AND } (B > 10) \text{ AND}$$
$$(\max\{R, G, B\} - \min\{R, G, B\} > 15) \quad \text{AND} \qquad (1)$$
$$(|R - G| <= 15) \text{ AND } (R > G) \text{ AND } (R > B)$$

while modified RGB facial skin color thresholds under flashlight or daylight lateral illumination are given as follows:

$$(R > 210) \text{ AND } (G > 200) \text{ AND } (B > 160) \text{ AND}$$
$$(|R - G| <= 15) \text{ AND } (R > B) \text{ AND } (G > B) \qquad (2)$$

All threshold values were set through our empirical studies and findings. We use a logical OR operator to combine Equation (1) and (2), and obtain the desired threshold values for RGB skin color.

In YCbCr color space, we extract Y value from skin color model [9] and use it to strengthen the RGB skin color thresholds. In order to preserve the facial features of low brightness (i.e., *dark* color), we decrease the value on Y. The value of Y threshold is defined as follows:

$$Y \geq 70 \qquad (3)$$

In HSV color space, we adopt the work that Phung et al. [10] proposed for skin color threshold to enhance the skin color determination. They defined the hue and saturation value as follows:

$$0° \leq H \leq 50°$$
$$0.23 \leq S \leq 0.68 \qquad (4)$$

At last, we use a logical AND operator to combine the thresholds in three color spaces, i.e., RGB, YCbCr, HSV, and generate a binary mask after using such thresholds for skin color detection. Figure 3 shows the result after skin color detection.
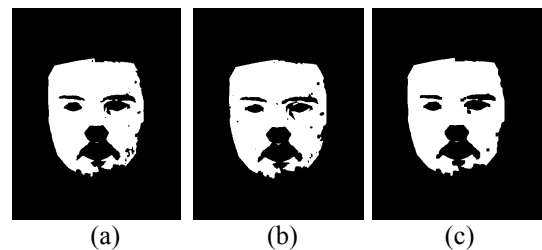

(a) (b) (c)

Figure 3. The result images of skin color detection. The last one was observed to get the best

result. (a) Using only RGB skin color threshold. (b) Using both RGB and YCbCr skin color thresholds. (c) Using all RGB, YCbCr and H-S skin color thresholds.

3.2. Face alignment

For face alignment, we use ASM to find the facial feature points on both source image and example image, and implement Delaunay triangulation algorithm [11] to build the triangular meshes. At last, we adopt affine transformation [12] to compute the transformation matrix of each triangle so as to perform image *warping*, making each facial component on example image located in a near position with respect to that on source image. The result of face alignment is shown in Figure 4.
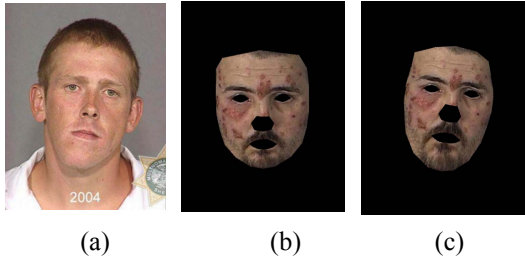

(a)   (b)   (c)

Figure 4. Result of face alignment from example to source image. (a) Source image (b) Example image (c) Result image.

3.3. Image cloning

We use seamless cloning algorithm [13] to clone the facial features on the example into face on the source image. The result is shown in Figure 5.
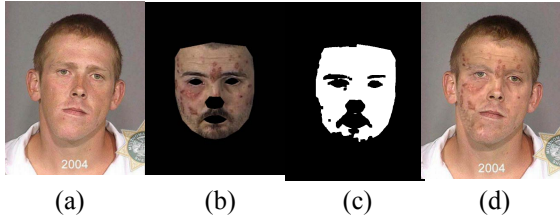

(a)   (b)   (c)   (d)

Figure 5. The result of image cloning. (a) Source image (b) Example image (c) Binary mask (d) Result image.

We also take care of the cases in which a face on a source image has a beard. Under such case, we must perform skin color detection for source image, and acquire a new binary mask (in Figure 6 (c)) which excludes the beard region.
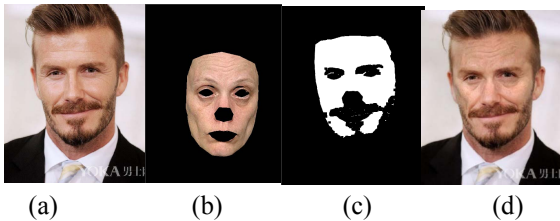

(a)   (b)   (c)   (d)

Figure 6. Result with a beard existing on the source image. (a) Source image (b) Example image (c) A new binary mask (d) Result image.

3.4. Additional processing

In this section, we introduce two additional techniques, first is an eyebrow detection, second is an automatic selection of an example image in data set, and describe our motivation and approach.

We observe that some eyebrow color on an example image is often too light to be detected, making the eyebrow on such example image accidently be transferred to the source image. Therefore, we adopt Moreira et al. [14] proposed technique to find the approximate region of eyebrows, and define a new threshold to determine the location of eyebrow.

The threshold range is shown as follows:

$$150 \leq Y \leq 255$$
$$84 \leq Cb \leq 136 \qquad (5)$$
$$139 \leq Cr \leq 181$$

Figure 7 illustrates the result of eyebrow detection.

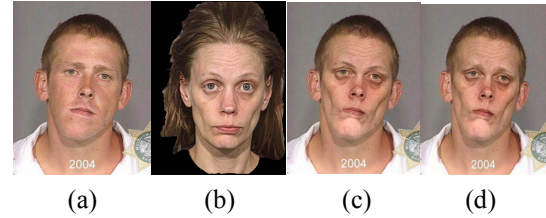(a)   (b)   (c)   (d)

Figure 7. Eyebrow detection. Last one was observed to get a better result. (a) Source image (b) Example image (c) Result image without eyebrow detection (d) Result image with eyebrow detection.

In data set, we often have a lot of example images, each with a different skin color. A good example image candidate should have a similar facial skin color compared to that on the source image. Therefore, we provide an algorithm for automatic selection of a suitable example image with a similar facial skin color. First, we convert from RGB to YCbCr color space for both source and example images; then we transfer the values of Hue and Saturation to compute histogram. Finally, we adopt Bhattacharyya distance [15] to compare the H-S histogram between the source image and example image. The formula of Bhattacharyya distance is shown as below:

$$d(H_1, H_2) = \sqrt{1 - \frac{\sum_i \sqrt{H_1(i) * H_2(i)}}{\sqrt{\sum_i H_1(i) * \sum_i H_2(i)}}}$$

where H_1 and H_2 represent the histogram on source image and example image, respectively; i represents the value of i-th histogram, and d represents the similarity between two histograms; d's range is [0, 1]. A smaller d value implies that facial skin color between two images is more similar.

4. Experimental Results

Our system performed an image cloning algorithm to

clone the facial feature area in the example image onto the source image. We applied skin color detection to generate the mask needed by image cloning, and used facial component detection to find the facial skin area in example image. Additionally, if face on the source image has beards, we also implemented skin color detection for source image, and integrated the result with the mask of an example image. In this section, we display the results yielded from several cases. The major aim is to verify the necessity of using skin color detection for an image whose face has beards on it. Figure 8 to Figure 10 show the image cloning results from different examples and source images.
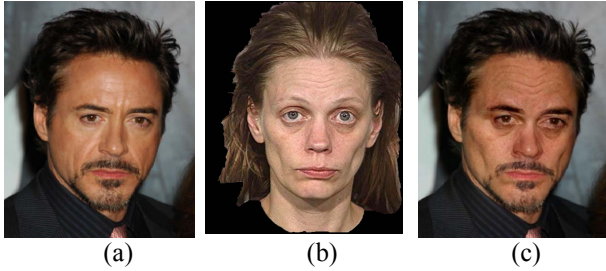

Figure 8. Image cloning using a source image with breads [16]. (a) Source image (b) Example image (c) Result image.
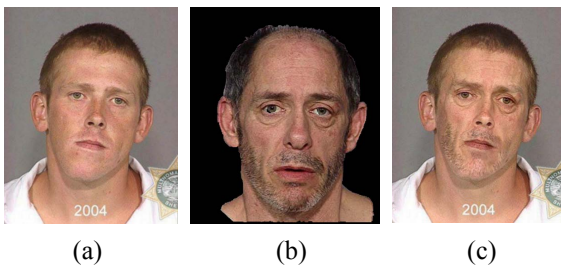

Figure 9. Image cloning using an example image with beards [16]. (a) Source image (b) Example image (c) Result image.
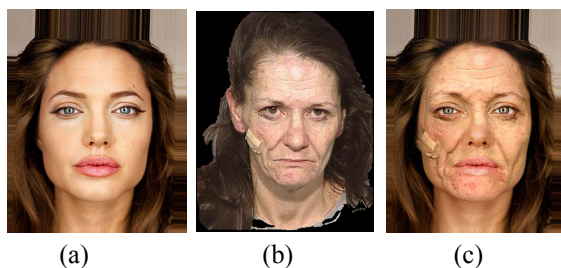

Figure 10. Image cloning with one other example image [16]. (a) Source image (b) Example image (c) Result image.

5.   Conclusion and Future Work

In our study, we proposed a system on interactively cloning human facial features by examples. We allow a user to choose an example image and offer additional information via user interface. Through our system, we can clone facial features from example image to source image, making two images have a similar visual effect.

Through this system, we can easily clone the facial

features on drug addicts to the faces on other images, and have more facial features on drug addicts can be selected.

In the future work, we hope that the system can be applied to other different domains, such as smokers, and alcoholics.

References

[1] J. Suo, F. Min, S. Zhu, S. Shan, X. Chen: "Multi-Resolution Dynamic Model for Face Aging Simulation," *IEEE Conference on Computer Vision and Pattern Recognition,* pp. 1-8, 2007.

[2] D. Guo, T. Sim: "Digital Face Makeup by Example," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pp. 73-79, 2009.

[3] Y. Shih, S. Paris, C. Barnes, W.T. Freeman, F. Durand: "Style Transfer for Headshot Portraits," *ACM Transactions on Graphics,* vol. 33, no.4, pp.1-148, 2014.

[4] Z.J. Xu, H. Chen, S.C. Zhu: "A High Resolution Grammatical Model for Face Representation and Sketching," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pp. 470-477, 2005.

[5] C. Rother, V. Kolmogorov, A. Blake: "Grabcut: Interactive Foreground Extraction using Iterated Graph Cuts," *ACM Transactions on Graphics,* vol. 23, no. 3, pp. 309-314, 2004.

[6] S. Milborrow, F. Nicolls: "Locating Facial Features with an Extended Active Shape Model," *IEEE Conference on Computer Vision and Pattern Recognition,* pp. 504-513, 2008.

[7] R.L. Graham, F.F. Yao: "Finding the Convex Hull of a Simple Polygon," *Journal of Algorithms,* vol. 4, no. 4, pp. 324-331, 1983.

[8] P. Peer, J. Kovac, F. Solina: "Human Skin Color Clustering for Face Detection," *IEEE Conference on Computer as a Tool,* pp. 144-148, 2003.

[9] N. Rahman, K. Wei, J. See: "RGB-H-CbCr Skin Color Model for Human Face Detection," *MMU International Symposium on Information & Communications Technologies,* 2006.

[10] S.L. Phung, A. Bouzerdoum, D. Chai: "Skin Segmentation using Color Pixel Classification: Analysis and Comparison," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 148-154, 2005.

[11] B. Delaunay: "Sur la Spe`re Vide," *Bull. Acad. Sci. USSR (VII)*, pp. 793-800, 1934.

[12] G. Wolberg: Digital Image Warping, *IEEE Computer Society Press*, 1990.

[13] P. Prez, M. Gangnet, A. Blake: "Poisson Image Editing," *ACM Transactions on Graphics,* vol. 22, no. 3, pp. 313-318, 2003.

[14] J.L. Moreira, A. Braun, S.R. Musse: "Eyes and Eyebrows Detection for Performance Driven Animation," *IEEE Conference on Graphics, Patterns and Images*, pp. 17-24. 2010.

[15] Bhattacharyya Distance:
https://en.wikipedia.org/wiki/Bhattacharyya_distance.

[16] Style Transfer for Headshot Portraits:
https://people.csail.mit.edu/yichangshih/portrait_web/