An automatic framework for example-based virtual makeup

Conference Paper · September 2013

DOI: 10.1109/ICIP.2013.6738660

Lin Xu
Intel

Yimin Zhang
Intel

AN AUTOMATIC FRAMEWORK FOR EXAMPLE-BASED VIRTUAL MAKEUP

Lin Xu, Yangzhou Du, Yimin Zhang

{lin.x.xu, yangzhou.du, yimin.zhang}@intel.com
Intel Labs China

ABSTRACT

Cosmetic makeup is a general event in our daily life, which improves women's beauties and attractions. But, it is difficult for ordinary users to make a wonderful makeup as the cover girls. Moreover, when you are in nude look and want to share better look with your friends, the fastest and easiest way is virtual makeup. However, current existing makeup software needs many user inputs to adjust face landmarks, which influence the user experience. And, it cannot remove the flaws on skin as good as the real cosmetic makeup. In this paper, we describe an automatic framework to apply a cosmetic makeup and skin beautification to your face, which can be selected from many example make-up face images. Our method detects the face landmarks with existing algorithm and adjusts the landmark with skin color Gaussian Mixture model based segmentation. Then, the skin color area is separated into three layers, and makeup is transferred with different method to different layers. The results look pretty good for some natural input face images.

Index Terms— virtual makeup, face landmark detection

1. INTRODUCTION

Cosmetic makeup is an ordinary technique to improve one's beauties and attractions. But, a delicate makeup needs many procedures: foundation, powder, eye shadow, lip gloss, blusher, etc. For some dummies, it maybe takes a long time for this whole procedure. Virtual makeup is a tool for the users to see the makeup results without actually paint cosmetic on their faces. This kind of tools can help the customers to select cosmetics on the face shops. Furthermore, it can also beautify the face in your photo before you share with your friends.

There are not many virtual makeup related works in previous studies. [1] is an example-based makeup transfer algorithm. But it deeply relies on "before-and-after" pairs, which is taken in the same position, expression and lighting environment. In the pre-processing steps, it needs many user inputs to select special areas, such as eye brows, moles, and freckles. [2] improves [1] with texture separation and blending example makeup with input face image. But it also needs users to adjust 88 face landmarks. The makeup procedure in our framework is similar to [2], but we improve it with automatic landmark adjustment.

In additional research on virtual makeup is on skin color enhancement, such as [5][6]. [5] also uses "before-and-after" pairs as examples for foundation synthesis. [6] utilizes a physical model to extract hemoglobin and melanin components. It also need user to select skin area which includes spots or moles. Furthermore, [4] is a makeup suggestion system to optimize the best makeup for user with a professional "before-and-after" pair dataset. However, since the limited dataset, the cosmetic effects are also limited.

There are also some studies on face beautification. [3] is an interesting study to improve face attractiveness with structure adjustment, which adjusts the input face structure with training lots of attractive faces. Other studies improve face attractiveness from textures [7][8][9], which remove flaws, moles and wrinkles from faces.

Some commercial softwares, such as Taaz [10] and Modi Face [11], provide users with virtual makeup on face photos by simulating the effects of specific cosmetics. But the manual adjustment of face landmark still limits the user experience in this kind of software usages.

In this paper, we propose an automatic framework for virtual makeup, which resolve two challenging problem: automatically detect face landmark and makeup synthesis between example and input face images. First, we utilize state-of-the-art ASM face landmark detection algorithm to locate key points in the face, and automatically adjust skin areas with Gaussian mixture models (GMM). Second, we use bilateral filtering to skin areas to remove flaws, wrinkles, etc.. At last, we use similar makeup technique in [2] to transfer the makeup.

The rest of the paper is organized as follows. In section 2, we introduce the face landmark detection method and GMM-based auto-adjustment of skin area. Skin smoothing and makeup transfer methods are introduced in the following section 3 and 4. In section 5, some results are presented. In the last section, we draw a conclusion.

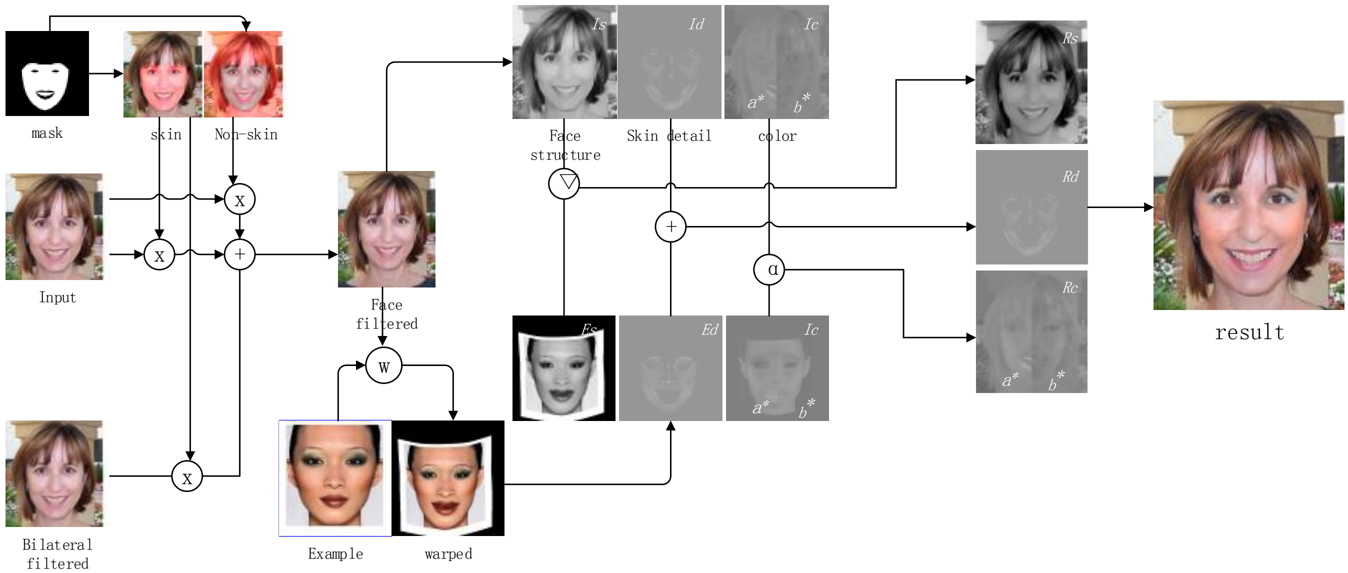


Figure 1. The framework of our face makeup system. The input face image and the filtered image are combined in skin region. In non-skin region, the original image is copied without any change. The combined image is decomposed into different layers: Face structure, skin detail and color. The makeup example image is warped (W) by combined image. ∇ denotes gradient editing. + denotes weight addition. ∝ denotes alpha blending.

2. DIGITAL MAKEUP FRAMEWORK

The framework of our virtual makeup system is illustrated in Figure 1. In the first step, an automatic face landmark location algorithm proceeds on input image *I*. In the next step, skin smoothing based on skin area mask is generated. And then, face alignment between the smoothed input image and the example image *E*, providing makeup example, is produced. Followed is layer decomposition. Both the input and example image is decomposed into three layers: face structure layer, skin detail layer, and the color layer. We utilize different transfer method in different layer to produce better effect. Finally, three resultant layers are composed together.

2.1. Automatic face features location

Face landmark detection is a traditional topic in computer vision. Active Shape Model (ASM) [10] is the most commonly used method for face landmark detection. Due to different lighting condition or color difference, ASM may not get accurate position of landmark points. So, many makeup softwares need user manually refine the position of control points. From our observation, most of the makeup transfer only processing on skin areas. Non-skin areas, such as eye and eyebrows, will not be modified. So, we adopt Gaussian Mixture Model (GMM) to refine the skin areas.

First, GMM needs data sampling on skin and non-skin area. For the input image, we adopt the data sampling method described in [9], which utilize the 66 face landmark points to select the skin area, described in Figure 2.

After data sampling, we model the pixel color distribution (in RGB format) for skin and non-skin data as GMM [11]. The input to GMM is a vector of three-dimension color pixel values. The initial *K* is set to 20.

With the GMM estimation, we segment the portrait into skin and non-skin area by a Bayesian segmentation algorithm. We do the segmentation using a multi scale random field (MSRF) in conjunction with a sequential maximum *a posteriori* (SMAP) estimator described in [12]. Figure 2 shows the segmentation results.

After the segmentation, we can obtain a more accurate boundary between skin and non-skin area. Then, we refine the face landmarks with the nearest-neighborhood point on the boundary line. For points on the lower jaw, we will not change the points since the skin-area is connected with neck. The points on the eyebrows will not be changed, either, since the boundary line is on hairs on forehead.

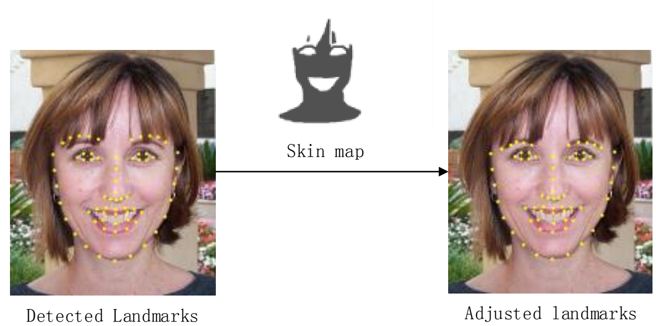


Figure 2. automatic face landmark adjustment.

2.2. Skin smoothing

Image filtering technology can remove some flaws on face, but it also loss image details. Many studies improve the filter operator to achieve balance between those two aspects. Bilateral filter [11] is a common used operator for skin smoothing. In this paper, we apply bilateral filtering for the input image and obtain a smoothed image (in figure 3(b)). With the skin mask obtained in 2.1, we use a poison image editing method [13] to mosaic input and filtered images together. The result image removes the flaws in skin area and reserves the details in other areas (in figure 3(c)).

(a) input (b) Bilateral filtered (c) combined

Figure 3. Result of skin mask based face filtering.

2.3. Digital makeup

In the digital makeup, first face alignment should be done between the example and input faces. Then, we decomposed the images into three different layers: face structure, Skin detail and color. At last, color transfer is produced in different layers and combined together.

For face alignment, the example image E is warped to input image I with the Thin Plate Spline (TPS) [13]. The control points are the face landmark points determined in 2.1.

The input image I and the warped example image E are decomposed into color and lightness layer by converting them to CIE $L^*a^*b^*$ color space. The L^* channel is considered as lightness layer and a^* and b^* the color layer. The lightness layer is decomposed into face structure and skin detail layers. While several recent computational photography techniques can perform image decomposition, the weighted least squares (WLS) [14] is one of the state-of-the-art methods. It was originally proposed for general edge-preserving image decomposition.

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

Figure 4. (a) face landmarks points. (b) facial components: blue as C_1, red as C_2, *and* green as C_3. (c) unsmmothed β value. (d) β value computed by equation (4)

Let l, s and d, represent the lightness, face structure and skin detail layers respectively. The WLS method seeks to minimize the energy function,

$$E = \sum_p [(l_p - s_p)^2 + \lambda \cdot \Upsilon(\nabla l_p, \nabla s_p)] \tag{1}$$

Where p indexes the image pixel, the term $(l_p - s_p)^2$ is to make s as close as possible to l, and λ is a constant to balance the two terms. The regularization term $\Upsilon(\nabla l_p, \nabla s_p)$ is to make s as smooth as possible everywhere except across significant gradients in l using the formation,

$$\Upsilon(\nabla l_p, \nabla s_p) = \frac{(\frac{ds}{dx}(p))^2}{\left|\frac{dl}{dx}(p)\right|^\alpha + \varepsilon} + \frac{(\frac{ds}{dy}(p))^2}{\left|\frac{dl}{dy}(p)\right|^\alpha + \varepsilon} \tag{2}$$

To protect the facial parts, such as eyes and mouth, a weighting template $\beta(p)$ is introduced in [2] to get a new formulation,

$$E_1 = \sum_p [(l_p - s_p)^2 + \lambda \cdot \beta(p) \cdot \Upsilon(\nabla l_p, \nabla s_p)] \tag{3}$$

where $\beta(p)$ is defined by

$$\beta(p) = \min_q (1 - g(q) e^{-\frac{(q-p)^2}{2\sigma^2}}) \tag{4}$$

where q indexs the pixel over the image. $g(q)$ is 0 for skin area, 1 for other facial components. The value of σ^2 is set to *min*(height,width)/2.5. The result of $\beta(p)$ is shown in Figure 4.

The optimal solution of E_1 in (3) is the facial structure image s. Then the skin detail s can be obtained by:

$$d(p) = l(p) - s(p) \tag{5}$$

After decomposition of face layers, different transfer strategy is processed in different layer. In skin detail layer, we straightforward transfer example face with input:

$$R_d = \delta_I I_d + \delta_E E_d$$

where $0 \leq \delta_I, \delta_E \leq 1$. Since the physical makeup for foundation and loose powder is to cover the original skin detail. Thus, we set $\delta_I = 0$ and $\delta_E = 1$.

For color layer, we use alpha-blending of input and example image as the result color layer R_c, *i.e.*

$$R_c(p) = \begin{cases} (1-\gamma) I_c(p) + \gamma E_c(p) & p \epsilon C_3 \\ I_c(p) & otherwise \end{cases} \tag{6}$$

The value of γ is to control the blending effect of two color layers. In our experiment, we use 0.8.

Face structure layer contains the highlight and shading effects on the face. Since those effects are important in cosmetic makeup, we adopt a gradient-based editing method, which can preserve the illumination of I and transfer the effects. The gradient of R_s is defined as:

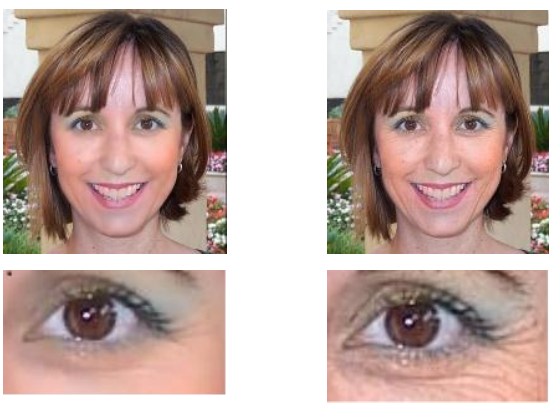
$$\nabla R_s(p) = \begin{cases} \nabla \varepsilon_s(p) & if\ \beta(p)\|\nabla \varepsilon_s(p)\| > \|\nabla I_s(p)\| \\ \nabla I_s(p) & otherwise \end{cases} \quad (7)$$

3. RESULTS

Our framework can automatic beautify skin area and manipulate makeup effects on input face images. The example makeup images are selected from [15], which is a professional makeup handbook. The input images can be an ordinary photo. In this paper, we choose some faces from Caltech face database [16].

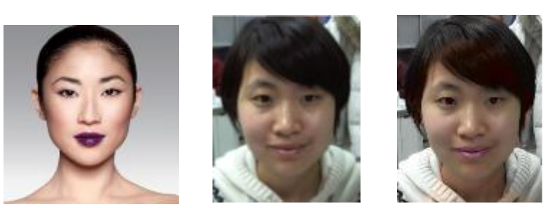
Because skin smoothing is produced before makeup transfer, compared to tradition virtual makeup softwares, our method shows better effects on the skin flaws, such as wrinkles on corner of the eye (shown in figure 5.(b)).

Figure 6 shows another makeup result used by our method. The spots on the input face can be removed by filtering on skin areas.



(a) (b)

Figure 5. makeup result comparison between filtered (a) and original images (b).



(a) (b) (c)

Figure 6. Makeup transfer result: (a) example image; (b) input image; (c) makeup result

4. CONCLUSIONS

In this paper, we presented a framework to beautify face area and transfer cosmetic makeup by an example makeup image. The whole procedure is automatically with skin region mask, which makes face beautification more convenient for the end-user, especially on smartphones. We also apply a skin smoothing before makeup transfer, which is more like a concealer step in real makeup. The results show that our framework can produce better effects on the input faces.

11. REFERENCES

[1] W.S. Tong, C.K. Tang, M.S. Brown and Y.Q. Xu, "Example-Based Cosmetic Transfer," *15th Pacific Conference on Computer Graphics and Applications*, 2007.

[2] D. Guo and T. Sim, "Digital Face Makeup by Example", In Proc. *CVPR*, 2009.

[3] T, Leyvand, D. Cohen-Or, G. Dror, and D. Lischinski. "Data-driven enhancement of facial attractiveness". *ACM Trans. Graphics*, 27(3):1-9, 2008.

[4] K. Schebaum, T. Ritschel, M. Hullin, T. Thormahlen, V.Blanz, H.P. Seidel, "Computer-Suggested Facial Makeup", In Proc. *EUROGRAPHICS*, 2011, Volum 30(2011).

[5] N. Ojima, K. Yoshiba, O.Osanai, and S. Akasaki. "Image synthesis of cosmetic applied skin based on optical properties of foundation layers", In Proc. *ICIS*, pages 467-468, 1999.

[6] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, "Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin", *ACM Trans. Graphics*, 22(3):770-779, 2003.

[7] M. Brand and P. Pletscher. "A confitional random field for automatic photo editing". In *Proc. CVPR*, 2008.

[8] G. Guo, "Digital Anti-Aging in Face Images", In *Proc. ICCV*, 2011.

[9] C. Lee, M.T. Schramm, M. Boutin, J.P. Allebach, "An Algorithm for Automatic Skin Smoothing in Digital Portraits", In *Proc. ICIP*, 2009.

[10] S. Milborrow and F. Nicolls. "locating facial features with an extended active shape model", In *Proc. ECCV*, 2008.

[11] C. A. Bouman, "CLUSTER: an unsupervised algorithm for modeling Gaussian mixtures," https://engineering.pudue.edu/~bouman/software/cluster/manual.pdf.

[12] C.A. Bouman and M. Shapiro, "A multiscale random field model for Bayesian image segmentation," *IEEE Trans. Image Processing*, vol. 3, no. 2, pp.162-177, March 1994.

[13] F.Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(6):567-585, 1989.

[14] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski. ‘‘Edge-preserving decompositions for multi-scale tone and detail manipulation’’, *ACM Trans. Graphics*, 27(3):1-10, 2008.

[15] F.Nars. *Makeup your Mind*. PowerHouse Books, 2004.

[16] ‘‘Caltech faces 1999 (front),’’ http://www.vision.caltech.edu/html-files/archive.html.