Skin Tone Enhancement and Background Change for Mobile Phones

Jong-Ho Lee, Jin Heo, Student Member, and Yo-Sung Ho, Senior Member, IEEE

Abstract — In this paper, we propose skin tone enhancement and background change algorithms for mobile phones. It is important to obtain natural results in skin tone enhancement and background change since only parts of an image, skin tone and background, are changed artificially. Moreover, since our target is a mobile application, the algorithm should be fast to satisfy mobile users and require small memory space to meet limited mobile conditions. From the experiments, we verified that the proposed algorithm generates the result images of people whose skin tone is enhanced to the preferred skin color and the original background is changed to other ones such that the foreground and the background are harmonized well. Furthermore, the proposed algorithm works fast, so it takes about 0.15 seconds with 360×480 images in our test conditions.

Index Terms — skin tone enhancement, background change, mobile phones.

I. INTRODUCTION

Nowadays, it is important to share and exchange information with others, and a mobile phone becomes an essential device for the current people by supplying this ability. As people always carry the mobile phone with them, the mobile phone is required to be more powerful to fulfill the various needs. People want to listen to music, watch TV or movie, manage their schedule, and play games with their mobile phones. That is, the mobile phone plays an important role as an entertainment device beyond a communication device.

Especially, people like to take pictures related with their lives and share them with others. For the needs for people to take better pictures, the camera embedded in the mobile phone is required to have high performance. Furthermore, the demands for the high-end model with a high-resolution camera and various entertainment functions have continuously increased [1]. In accordance with these trends, high-level entertainment technologies related with the mobile phone camera have been studied and developed widely so far.

Recently, since the number of people who take pictures of themselves with the mobile phone cameras and upload them to their homepages or blogs are increasing, various entertainment technologies related with the pictures have been developed. Face detection, image matting for extracting a foreground element from a background image, automatic panorama image production, eye flickering detection to prevent people from closing their eyes when they are taken a picture, anti-shaking for reducing blur of the image, and denoising from the pictures are the examples for those technologies. In addition to these technologies, skin tone enhancement and background change are also strongly demanded technologies for the mobile phone.

When the people picture is observed, the parts which attract observers’ attention most are the people’s faces. Especially, skin tone of the face is one of the important parts to evaluate beauty of the face. Several studies say that individual’s skin tone determines the attitudes and behaviors toward that individual [2]. Furthermore, researches on the preferences for skin tone in various races are also studied. According to their findings, a lightened skin tone is rated most attractive in African-Americans, and a dark-ended skin tone is found to be most attractive in Caucasians. In addition, women’s facial skin texture affects male judgment of facial attractiveness, and it was found that a slightly reddish skin is considered attractive and healthy [3]. Thus, converting skin tone to the preferred skin color is very valuable in that it can improve subjective quality of the picture by satisfying the observers.

When the unwanted objects are captured in the background of the picture, or the users do not want to expose their private lives in the background to others, it will be useful to change the background to the other ones. This background change technology in the mobile phone can provide privacy protection for the mobile phone users, and also give pleasure to them by converting the unwanted background to the preferred ones.

In this paper, we propose the skin tone enhancement and background change algorithms for mobile phones. For the skin tone enhancement, only the color of skin region is converted to the preferred one. Thus, we focus on the skin color conversion method reproducing the skin color in a way that unnatural difference does not occur around the skin region and the non-skin region. We also focus on the background change method that produces the natural output considering the luminance differences between the foreground and the background.

The paper is organized as follows. In Section II, the overall framework of our proposed algorithm is introduced. In Section III and IV, our proposed background change and skin
tone enhancement algorithms are explained, respectively. In Section V, the experimental results are given, and finally, conclusion is drawn in Section VI.

**II. OVERALL FRAMEWORK**

Fig. 1 represents the overall framework of the proposed algorithm. Skin tone enhancement and background change work as a unified algorithm. First, the foreground and the background are separated by initial label setting and segmentation. Skin tone enhancement is only performed on the foreground. Skin area is detected on the lighting-compensated foreground object, and the skin color is converted to the preferred one. This skin-tone-enhanced foreground is combined with other background. Finally, brightness compensation is conducted to produce a natural result.

![Fig. 1. Overall framework of the proposed algorithm.](image)

**III. BACKGROUND CHANGE**

There are many methods on interactive digital matting, the process of extracting a foreground object from an image based on limited user input [4], [5], [6]. Among them, we select the growcut method [7] considering the mobile condition: small memory space and fast processing time. Therefore, we design a background change algorithm based on the modified growcut method.

![Fig. 2. Flowchart of the background change algorithm.](image)

Fig. 2 shows the flowchart of the background change algorithm. The proposed background change algorithm consists of five stages. After user specifies the initial seeds, the segmentation is performed. Next, in order to reduce boundary noise, boundary refinement is applied, and then composition of the foreground and the background is performed. Finally, based on the difference in brightness between the foreground and the background, brightness of the background is compensated. Further details of the background change algorithm are described in the following subsections.

*A. Specifying the Initial Seeds*

In order to set the initial seeds called labels for the foreground and the background, user interaction is required. In most case of the segmentation work in photo editing, user interaction is needed. Fig. 3 shows the user-specified initial seed image. White and black lines correspond to the foreground and the background, respectively, as shown in Fig. 3 (b). By using user interaction, we make a label map where unlabeled pixels are grey, the foreground pixels are white, and the background pixels are black, as shown in Fig. 3 (c).

![Fig. 3. User interaction. (a) Original image (b) User interaction image (c) Label map.](image)

*B. Segmentation*

After the initial labels are specified, a segmentation process starts. In the segmentation process, we define three parameters: the current pixel label value $l_p$, strength $\theta_p$, and three dimensional feature vector $\mathbf{C}_p$ in RGB space. The flowchart of the segmentation algorithm is depicted in Fig. 4.

![Fig. 4. Flowchart of the segmentation algorithm.](image)
Before foreground extraction, we set the initial label values and the initial strength. The initial label values for the white, black, and grey pixels are determined by 1, -1, and 0, respectively. The initial strength for the white and black pixels is set by 1 and the initial strength for the grey pixel is set by 0.

Next, we copy the previous pixel state, the label value and the strength of the previous pixel, and then update the current pixel state. The current pixel label value \( l_p \) is calculated by the product of the similarity of three dimensional feature vectors between the current and neighboring pixels \( g(C_p - C_{q_i}) \) and the strength of the neighboring pixel \( \theta_{q_i} \), where \( i \) means the neighboring pixel position \( 1 \leq i \leq 8 \), as shown in Fig. 5. If the product of \( g(C_p - C_{q_i}) \) and \( \theta_{q_i} \) is larger than the current pixel strength \( \theta_p \), \( l_p \) and \( \theta_p \) are updated by the neighboring pixel label value \( l_{q_i} \) and the product of \( g(C_p - C_{q_i}) \) and \( \theta_{q_i} \), respectively; otherwise the update process is bypassed. The similarity value \( g(C_p - C_{q_i}) \) is calculated by:

\[
g(C_p - C_{q_i}) = 1 - \frac{\| C_p - C_{q_i} \|_2}{\max \| C \|_2}
\]

To refine the boundary, we use a morphological operation: the closing operation which operates the dilation operation first, and then does the erosion operation [8]. In general, the closing operation tends to smooth sections of contours; it fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour. Therefore, in the proposed algorithm, we perform the closing operation by using the 5×5 dilation and 5×5 erosion masks. The boundary refinement result is represented in Fig. 6 (c).

**D. Composition with Mean Filter**

When we compose the extracted foreground into another background image, we can find some unnatural discontinuity at the boundary between the foreground and background images. In order to reduce the discontinuity, we use mean filter with 5×5 size. With mean filtering the boundary regions of the composite image is much smoother, as shown in Fig. 7 (b).

**E. Brightness Compensation**

The composite image is not still natural due to the difference in brightness between the foreground and background regions, as shown in Fig. 8 (c). Therefore, we compensate brightness of the background in the composite image based on the difference between brightness of the foreground and background regions.
foreground are considered, correct brightness data cannot be obtained.

First, we select the skin region in the extracted foreground object; we will explain details on skin detection in Section IV-B. After converting the color space of the selected skin region from RGB to YCbCr, we calculate the average Y value in the region. Next, we also convert the color space of the background region which does not overlap with the foreground from RGB to YCbCr, and then calculate the average Y value in the region.

After calculating the difference of the average Y values between the selected skin region and the background region, we compensate the difference Y value to the background region.

F. New Label Map Generation

In order to increase both the accuracy and the speed of segmentation processing, we modify the initial label map. In general, the foreground of a picture taken with the mobile camera is located at the center. Therefore, we can assume that the region below the white line is the foreground and the region above the black line is the background in Fig. 9 (a). Based on this assumption, we can modify the original label map to the new label map, as shown in Fig. 9 (b).

Before segmentation, we can definitely know that the white and black regions are the foreground and background regions, respectively, in the new label map. As a result, the segmentation process is only performed at the grey region. Thus, use of the new label map has advantages as follows. First, we can reduce the complexity of segmentation processing. Second, we can remove the holes which appear inside the foreground object under segmentation processing, as shown in Fig. 6 (a). Therefore, we can expect higher accuracy and faster processing speed in the segmentation stage.

IV. SKIN TONE ENHANCEMENT

Skin tone of the foreground is converted to the preferred skin color. Fig. 10 shows the overall framework for the skin tone enhancement algorithm. First, lighting compensation is performed on the foreground for the successful skin detection. Next, skin region is extracted from the lighting-compensated foreground. Based on the extracted skin region, the global rate and the local rate which represent the rate of skin color conversion globally and locally respectively are calculated. Finally, the color of skin region is reproduced to the preferred skin color based on the calculated global rate and the local rate. Further details of the skin tone enhancement algorithm are described in the following subsections.

A. Lighting Compensation

The skin-color distributions of the same person under different lighting conditions differ, and this causes degradation of skin detection performance. There are many methods to reduce the effect of illumination variation. Among them, we used the method combining grey world and retinex theory [9]. This method is suitable for the mobile application since it is simple, and fast enough to enable real-time processing. Furthermore, it does not require training and large memory space. Nevertheless, it shows great performance for lighting compensation.

This method combines two common methods, grey world and retinex theory. The grey world assumption argues that for a typical scene, the average intensity of the red, green, and blue channels should be equal. The retinex theory argues that the perceived white is associated with the maximum cone signals of the human visual system.

According to the method combining grey world and retinex theory, the red and blue channels can be adjusted with the following quadratic equation keeping the green channel unchanged.

$$I^* = \mu I^2 + \nu I$$

where $I^*$ is the adjusted intensity of red, or blue channel, $\mu$ and $\nu$ are the parameters for the adjustment which are obtained with the following equation.

$$\begin{bmatrix} \sum \sum I^2 & \sum \sum I \\ \max I^2 & \max I \end{bmatrix} \begin{bmatrix} \mu \\ \nu \end{bmatrix} = \begin{bmatrix} \sum \sum I \\ \max I \end{bmatrix}$$

(2) and (3) are constructed such that the grey world and the retinex theory are satisfied. Fig. 11 represents the effect of lighting compensation for skin detection.
Fig. 11. Effect of lighting compensation in skin detection. (a) Without lighting compensation (b) With lighting compensation.

B. Skin Detection

Skin region is extracted from the lighting-compensated foreground image. Like lighting compensation, there are many methods for the skin detection. Considering the complexity and required memory space for mobile application, we use a simple thresholding method for skin detection proposed by Hsu [10]. Although the skin region is detected by straightforward transformation of the given pixel and thresholding, this method shows better result than the other thresholding methods which use only chrominance components since it uses luminance component additionally.

The given pixel of the image is considered to be in the skin region if it is located inside the ellipse defined by (4).

\[
\frac{(x-ec_x)^2}{a^2} + \frac{(y-ec_y)^2}{b^2} < 1
\]  

(4)

where \(x\) and \(y\) are transformed chrominance components considering the luminance, and \(ec_x\), \(ec_y\), \(a\), and \(b\) are the constants. We use the values given in [10] for the constants. Fig. 12 represents the skin detection result on the foreground obtained in section III.

C. Preferred Skin Color

Kim [11] et al. tested preference on the preferred skin colors of Sanger [12], Kim [13], and Kim [14]. They made experiments with 24 images and the score for each image with different preferred skin color is given by the observers. According to their paper, Kim [13]’s preferred skin color shows the best grade, so we use the preferred skin color of Kim [13] for our skin tone enhancement algorithm. Actually, our skin tone enhancement algorithm is independent of the preferred skin color, so if required, other preferred skin colors can be used. The preferred skin color we used is as follows.

<table>
<thead>
<tr>
<th>Y</th>
<th>Cb</th>
<th>Cr</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.51</td>
<td>-0.07</td>
<td>0.1</td>
</tr>
</tbody>
</table>

D. Global Rate Calculation

It is important to control the rate of change in skin tone enhancement. For example, if a picture is taken in a rainy day, the skin tone in the picture will be dark. If the color of the skin is converted to the preferred skin color, unnatural result will be obtained due to the large color difference between the skin and the non skin regions. The global rate is to control the degree of skin color conversion globally in each component.

Skin color conversion is performed in HSV domain, and the global rates for \(H\), \(S\), and \(V\) component are calculated by:

\[
GR_H = e^{-\frac{H_{offset}}{\sigma_H}}
\]  

(5)

\[
GR_S = e^{-\frac{S_{offset}}{\sigma_S}}
\]  

(6)

\[
GR_V = 1 - \alpha V_{offset}
\]  

(7)

where \(GR_H\), \(GR_S\), and \(GR_V\) are the global rates for \(H\), \(S\), and \(V\) component respectively. \(H_{offset}\), \(S_{offset}\), and \(V_{offset}\) are the differences between the preferred value and the average value of the detected skin of \(H\), \(S\), and \(V\) component respectively.

If the difference between the preferred value and the average value of the skin region is large, the global rates for \(H\) and \(S\) components are decreased with Gaussian functions, and that for \(V\) component is decreased with a linear function. This leads to weak skin color conversion if the difference between the preferred skin color and the detected skin color is large.

Fig. 12. Skin detection. (a) Input image (b) Foreground (c) Skin mask.

Fig. 13. Effect of global rate in skin tone enhancement. (a) Without global rate (b) Original image (c) With global rate.
With these global rates, more natural enhancement of skin color can be achieved. The details of skin color conversion with these global rates will be explained in Section IV-F. We used $a_r=8.33 \times 10^{-3}, a_i=1.67 \times 10^{-2}, a_v=0.7$ for the experiments. Fig. 13 shows the effect of global rate in skin tone enhancement.

### E. Local Rate Calculation

The decision for skin or non-skin region is made by the ellipse defined by (4) and represented with the solid red line in Fig. 14 (a). Since $p_1$ is inside the ellipse, it is considered to be in the skin region, and is converted to the preferred value. However, although $p_2$ is very close to $p_1$, it stays unchanged since it is outside the ellipse.

That is, two pixels which have very similar colors in the original image come to have very different colors after skin color conversion, and this leads to contouring artifact as in Fig. 15 (b).

Local rate is to reduce this contouring artifact. It controls the degree of skin color conversion based on the location of the color of pixel in the skin ellipse so that abrupt changes cannot happen at the ellipse boundary. Local rate, $LR$, is calculated as follows.

$$LR = \begin{cases} 1 & 0 \leq L < T_i \\ \frac{(L-T_i)^2}{\sigma^2} & T_i \leq L \leq 1 \end{cases} \quad (8)$$

where $T_i=0.8, \sigma_i=0.01$, and $L$ is defined as:

$$L = \frac{(x-e_{cx})^2}{a^2} + \frac{(y-e_{cy})^2}{b^2} \quad (9)$$

That is, if the pixel inside the skin region is located near the skin boundary, $T_i \leq L \leq 1$, the degree of skin color conversion is reduced according to the Gaussian function defined in (8), as shown in Fig. 14 (b).

The effect of local rate is represented in Fig. 15. Without local rate consideration, contouring artifacts occur as in Fig. 15 (b) after skin color conversion. However, contouring artifacts are reduced with local rate in skin color conversion as in Fig. 15 (c).

### F. Skin Color Conversion

Finally, the skin color is converted to the preferred one with the global rates and the local rates which are calculated with (5) to (8) as follows.

$$H' = H + GR_H \times LR \times H_{offset} \quad (10)$$

$$S' = S + GR_S \times LR \times S_{offset} \quad (11)$$

Fig. 16 shows the final result of skin tone enhancement.

V. EXPERIMENTAL RESULTS

In order to evaluate performance of the proposed algorithm, we used several test images, original and background images taken by the camera of commercial mobile phone. The size of all test images is 360×480. We implemented our proposed algorithm by using C programming. The experimental environments are shown in Table II.

<table>
<thead>
<tr>
<th>Device</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>Quad <a href="mailto:CPU@2.40GHz">CPU@2.40GHz</a></td>
</tr>
<tr>
<td>RAM</td>
<td>3 GB</td>
</tr>
</tbody>
</table>
Fig. 16 shows the experimental results. We tested our proposed algorithm with four test images, as shown in Fig. 16 (a). First, the foreground object is extracted, as shown in Fig. 16 (b) and the skin tone of the extracted foreground element is converted to the preferred skin color, as shown in Fig. 16 (c). The skin-tone-enhanced foreground is combined with the new background image in Fig. 16 (d) to produce the final result in Fig. 16 (e). The pale skin tones in the test images A, B, and C were converted to healthy skin tones in the upper three results in Fig. 16 (c), and the reddish skin tone in the test image D was converted to bright skin tone, as shown in the last result in (c). Furthermore, the luminance of the combined
background image in Fig. 16 (d) is adjusted to have similar luminance as the enhanced foreground in Fig. 16 (c), so natural results are obtained, as shown in Fig. 16 (e).

Table III shows the execution time of the proposed algorithm. It takes about 0.06 seconds for the skin tone enhancement, and about 0.1 seconds for the background change, so totally, less than 0.2 seconds in our test conditions with 360×480 images.

<table>
<thead>
<tr>
<th>Original image</th>
<th>Skin tone enhancement (seconds)</th>
<th>Background change (seconds)</th>
<th>Total time (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.062</td>
<td>0.109</td>
<td>0.171</td>
</tr>
<tr>
<td>B</td>
<td>0.062</td>
<td>0.093</td>
<td>0.155</td>
</tr>
<tr>
<td>C</td>
<td>0.062</td>
<td>0.093</td>
<td>0.155</td>
</tr>
<tr>
<td>D</td>
<td>0.062</td>
<td>0.093</td>
<td>0.155</td>
</tr>
</tbody>
</table>

VI. Conclusions

In this paper, we proposed the skin tone enhancement and background change algorithms for the mobile devices. For skin tone enhancement, the color of the skin region is reproduced based on the preferred skin color. The degree of skin color conversion is controlled to get satisfying results according to the difference between the preferred skin color and the input skin color, and the location of the input pixel in the skin region on the skin ellipse. For background change, the unwanted background is replaced with the preferred background. Natural results were obtained by compensating the luminance difference between the foreground and the background. From the experiment, we demonstrated that the proposed algorithm can generate the picture of people whose skin tone is enhanced and background is changed naturally. In addition, the proposed algorithm works fast with about 0.16 seconds with 360×480 images.

REFERENCES


BIOGRAPHIES

Jong-Ho Lee received the B.S. degree in electronic engineering from Ajou University, Gyeonggi-do, Korea, in 2006. He is currently pursuing the M.S. degree from the Department of Information and Communication Engineering from the Gwangju Institute of Science and Technology, Gwangju, Korea. His research interests include digital signal and image processing, and image deblurring.

Yo-Sung Ho (SM’06) received the B.S. and M.S. degrees in electronic engineering from Seoul National University, Seoul, Korea, in 1981 and 1983, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, in 1990. He joined the Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea, in 1983. From 1990 to 1993, he was with Philips Laboratories, Briarcliff Manor, NY, where he was involved in development of the advanced digital high-definition television system. In 1993, he rejoined the Technical Staff of ETRI and was involved in development of the Korea direct broadcast satellite digital television and high-definition television systems. Since 1995, he has been with the Gwangju Institute of Science and Technology, Gwangju, Korea, where he is currently a Professor in the Department of Information and Communications. His research interests include digital image and video coding, multiview video coding, and high performance video coding technologies.