Performance Improvement of Face Image Super-Resolution Processing by High-Precision Skin Color Detection

Keigo Kano, Tomio Goto, and Satoshi Hirano
Dept. of Computer Science, Nagoya Institute of Technology, Nagoya, Japan
Email: keigo430@splab.nitech.ac.jp, {t.goto, hirano}@nitech.ac.jp

***Abstract*—In recent years, opportunities to deal with digital images on the Internet have increased due to the information society, there is a great demand for techniques such as super-resolution processing to make images more beautiful. When super-resolution processing is performed on natural images such as scenery, edges are emphasized to obtain clear images. However, when super-resolution processing is applied to a facial image, the wrinkle and stains of the skin as well as the emphasis of hair and eyes are emphasized, so super-resolution processing on the skin part is not suitable. Therefore, in the previous study, we proposed a method to perform facial correction using non-linear filter on skin part, and tried to solve this problem. This method is composed of super-resolution processing and facial correction processing, and it was possible to realize a super-resolution processing with a sharp sense for facial images. However, we also confirmed that there was a problem that the image quality deteriorated according to the skin color detection accuracy at the time of image synthesis of each processed image. Therefore, in this paper, we study the skin color detection method and try to improve the image quality.**

***Index Terms*—super-resolution, facial correction, skin detection**

I. Introduction

Recently, digital cameras, smartphones and tablet terminals have been being increasingly used to handle very high-definition images. Especially, cameras on mobile phones have become highly functional year by year. With the appearance of smartphones, it is now possible to take photographs with a resolution of 10 million pixels or more in many camera functions. Also, opportunities to deal with digital images on the Internet are increasing due to social information. Processing such as sharper and more beautiful processing is done on images taken to be open to the public for many and unspecified people. There is a huge demand, especially face image processing technology is very demanding. Since it is becoming common to take pictures with smartphones and to process images on smartphones and to use them, techniques that make them easier to handle and more beautiful are required. As described above, the demand for facial correction becomes high, research is proceeding not only in engineering but also in a number of fields, and among them, super-resolution technology is a technique for increasing the sharpness of images. Linear interpolation techniques have been used conventionally when displaying low resolution images on high resolution screens. However, it is known that the image expanded by the linear interpolation technique becomes blurred. This blur occurs due to the lack of high frequency components. Super-resolution technology is a technique aimed at exceeding the limit of such a linear interpolation technique. Super-resolution processing is roughly divided into a method of using a plurality of input images to generate one high resolution image and a method of generating a high resolution image from one input image. In this paper, we deal with a method of generating a high resolution image from one input image, and we will examine facial image processing using super-resolution and aim to perform processing that makes digital image look beautiful. Finally, we aim to apply the technology to applications and distribute them to general users. The super-resolution processing is performed using a TV regularization filter, a shock filter, and a pulse emphasis filter. Introduction TV regularization decomposes the original image into structure components and texture components. Thereafter, after each component is linearly expanded, the structure component is processed by a shock filter to emphasize the edge. The texture component is processed by a pulse emphasis filter to strengthen the sharpness of the pulse component. A processed image by each filter is synthesized to obtain a super-resolved processed image. When super-resolution processing is performed on natural images such as scenery, edges are emphasized to obtain clear images. However, when super-resolution processing is applied to a facial image, the wrinkle of the skin is emphasized together with the emphasis of hair and eyes, so super-resolution processing on the skin part is not suitable. Therefore, this problem is solved by performing facial correction processing on the skin part. For the skin part, facial correction processing is performed, and super-resolution processing is performed on non-skin part. We will try to make the face image more beautiful by combining super-resolution processing and facial correction processing. The facial correction processing is

Manuscript received April 10, 2019; revised August 7, 2019.

doi: 10.18178/joig.7.3.82-89

performed using a non-linear filter. Perform a non-linear filtering process on the face image to eliminate wrinkles and stains. In this paper, TV regularization filter is used as a non-linear filter. By performing the above processing, a facial-corrected image is obtained. Thereafter, the skin part and the non-skin part subjected to each processing are synthesized. In order to separate the skin part and the non-skin part of the image, skin color detection needs to be performed. In previous research, HSV color conversion was used as skin color detection method. However, this method was insufficient to detect dark parts that were not exposed to light such as the neck. Therefore, we independently devise a color conversion method specialized for skin color and use it for skin color detection. Also, in the conventional method, skin color detection is performed using a fixed threshold. Therefore, in the proposed method, an algorithm that automatically adjusts the threshold for each image is devised, and skin color detection is performed. Thereafter, the facial-corrected image and the super-resolved image are combined based on the skin color detection result. In the skin color detection result, a facial-corrected image is applied to a part judged to be skin color, and a super-resolved image is applied to a part judged to be non-skin color, thereby obtaining a final result image. In the resultant image, the skin part can be cleared of the wrinkle by the facial correction processing, and the non-skin part can increase the texture of eyes and hair by super-resolution processing, and as a result, the more beautiful facial image can be obtained. In this paper, we will explain the super-resolution processing and the theory of each filter to be used in chapter 2. In chapter 3, we will explain the facial correction, and how to remove wrinkles / stains. In chapter 4, a skin color detection method will be described. In chapter 5, the image composition method after each processing will be described. In chapter 6, super-resolution processing and facial correction processing are performed, and the generated output results are evaluated and considered. We also conduct a comparative experiment between the conventional method and the proposed method. And we will conclude in chapter 7.

II. Super Resolution

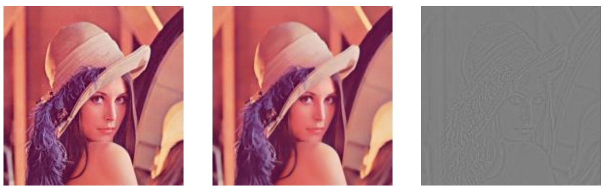
In this section, super-resolution processing using a non-linear filter will be described. By using a non-linear filter, it becomes possible to add high frequency components that can not be added by a linear filter. Through the super-resolution processing, each part of the image is enhanced. Specifically, Total Variation regularization is performed on the original image, and it is decomposed into a structure component and a texture component. After that, we emphasize the edge by applying a shock filter to the structure component, and apply a pulse emphasis filter to the texture component to strengthen the sharpness of the pulse component. Here, the processed images are synthesized to obtain a super-resolution processed image. In this section, three types of filters used in the super-resolution method will be explained. By using a non-linear filter, it becomes possible to add high frequency components that can not be added by a linear filter. Although the two non-linear filters have completely different characteristics, the image can be emphasized by decomposing the image into a structure / texture image by the Total Variation regularization method, and passing each component image through an appropriate filter. High-frequency components are added to the image after passing through the filter, and super-resolution is realized.

A. Total Variation

We will explain Total Variation (TV) regularization. The TV regularization method has various decomposition models [1]-[4]. In this paper, we use the ROF model. Total Variation is the sum total of luminance value fluctuations of adjacent pixels. The Total Variation *TV (u)* of the two-dimensional signal u is expressed in (1).

$$TV(u) = \int |\nabla u| dxdy \quad (1)$$

The ROF model of the Total Variation regularization is shown in (2). In (2), the first term on the right side is called a TV term and the second term is called a constraint term. The TV term is the sum total of fluctuation of the luminance value, and it increases as the vibration component increases. The constraint term is a difference between the observed image f and the updated image u, and is a term for determining to what extent it approximates to the observed image f. By minimizing the evaluation function $F(u)$, the vibration component is eliminated, and a signal composed of gentle change components and edges can be extracted. This signal u is called a structure image. In the ROF model, v can be derived from the difference between the original image f and the structure component u as shown in (3). This v is composed only of minute vibration components, and this is a texture component. The structure component is a part in which the luminance value is continuously violently fluctuated and removed, and the low frequency component is stored. In addition, it can be seen that the texture component is only the component whose luminance value fluctuates drastically. For the minimization problem of TV, A. Chambolle proposed a fast solution. In this algorithm [5], [6], the relative vector p shown in (4) is derived by iterative calculation. Next, the texture component v is obtained by calculating (5) using the obtained p. An example of structure / texture decomposition is shown in Fig. 1.



(a) Original Image (b) Structure Component (c) Texture Component

Figure 1. Structure / Texture decomposition

$$F(u) = \int |\nabla u| dxdy + \lambda \int |f - u|^2 dxdy \quad (2)$$

$$v = f - u \tag{3}$$

$$P_{i,j}^{n+1} = \frac{P_{i,j}^{n} + (\tau/\lambda)\nabla(f + \lambda\nabla \cdot P_{i,j}^{n})}{\max\{1, \left|P_{i,j}^{n} + \left(\frac{\tau}{\lambda}\right)\nabla\left(f + \lambda\nabla \cdot P_{i,j}^{n}\right)\right|\}} \tag{4}$$

$$v = \lambda \nabla \cdot P_{i,j}^{n+1} \tag{5}$$

B. Shock Filter

Shock filter is a non-linear edge emphasis filter proposed by Osher, Rudin. It is based on the idea of partial differential equations and restores or emphasizes the edges of the signal by simple computation. Processing is shown in (6)-(9).

$$u^{(t+1)} = u^{(t)} - sign(\Delta u^{(t)})|\nabla u^{(t)}|dt \tag{6}$$

$$|\nabla u| = \sqrt{u_x^2 + u_y^2} \tag{7}$$

$$\Delta u = u_{xx} + u_{yy} \tag{8}$$

$$\mathrm{sign}(x) = \begin{cases} 1\ (x > 0) \\ 0\ (x = 0) \\ -1\ (x < 0) \end{cases} \tag{9}$$

u is an operation output image, and an enlarged structure image is given as an initial value. u_t is the partial differentiation of *u* by *t* at time *t*, and *dt* is the step width of update operation of shock filter. The left term sign(*x*) is a sign function and is defined as in (9). By performing the calculation of (6), the maximum value and the minimum value in the local region are adjacent to each other at the zero crossing point of the Laplacian Δu. Therefore, a steep edge is formed near the zero-crossing point. From this fact, Laplacian acts as an edge detector which specifies a position to reconstruct steep edges. Also, since shock filter does not cause Total Variation of images and range change of maximum and minimum values, ringing distortion occurring in general edge sharpening processing does not occur. Also, since steep edges can be reconstructed only with one simple update operation, it is suitable for high speed processing. On the other hand, since various image quality deterioration occurs as compensation of edge sharpening, there are many problems to apply to general natural images. In order to prevent deterioration of image quality of shock filter, each arithmetic expression, gradient strength ∇u, Laplacian Δu is defined as follows. By performing the updating operation in (10), it is possible to construct steep edges while suppressing degradation of the natural image. In the equation, $K\sigma$ is a low-pass filter, and * is a convolution operator. By adding a smoothing effect to the edge detector, a smooth outline is constructed in the edge configuration [7]-[10]

$$u^{(t+1)} = u^{(t)} - sign(\Delta(K_\sigma * u^{(t)}))|\nabla u^{(t)}|dt \tag{10}$$

$$\begin{aligned}\Delta \mathrm{u} = {} & \partial_x^+\partial_x^- \mathrm{u} \cdot (\partial_x u)^2 + \partial_x^+\partial_x^- \mathrm{u} \cdot (\partial_y u)^2 \\ & +(\partial_x^-\partial_x^- \mathrm{u} + \partial_x^+\partial_x^+ \mathrm{u}) \\ & \cdot \partial_x u \cdot \partial_y u\end{aligned} \tag{11}$$

$$|\nabla \mathrm{u}| = \sqrt{(\partial_x u)^2 + (\partial_y u)^2} \tag{12}$$

$$\partial_x u = m(\partial_x^+ \mathrm{u}, \partial_x^- \mathrm{u}) \tag{13}$$

$$\partial_y u = m(\partial_y^+ \mathrm{u}, \partial_y^- \mathrm{u}) \tag{14}$$

$$\partial_x^\pm u(x,y) = \pm(u(x \pm 1, y) - u(x,y)) \tag{15}$$

$$\partial_y^\pm u(x,y) = \pm(u(x, y \pm 1) - u(x,y)) \tag{16}$$

$$\mathrm{m(x,y)} = \begin{cases} sign(x) \cdot \min(|x|,|y|)\ \ (xy > 0) \\ 0\ \ (xy \le 0) \end{cases} \tag{17}$$

It has been confirmed that image quality is greatly deteriorated if a component whose luminance value changes smoothly is passed through shock filter. This is thought to be due to the strong flattening action of shock filter. An edge is formed at a stationary point where the second order differential value, that is, the Laplacian value becomes zero, even in the smooth luminance value changing part. The image quality deterioration caused by this can be dealt with by lowering the sensitivity of the edge detector of shock filter. Specifically, threshold processing is incorporated in the sign function sign(*x*) as shown in (18). As a result, the edge reconstruction processing of shock filter is not performed in the smooth luminance value change portion where the Laplacian value becomes a minute value, and the processing can be performed only on the edge portion. As a result, it is possible to suppress deterioration of the smooth luminance value changing portion and to accurately process the edge portion.

$$\mathrm{sign}(x) = \begin{cases} 1\ (x > \varepsilon) \\ 0\ (x = \varepsilon) \\ -1\ (x < \varepsilon) \end{cases} \tag{18}$$

C. Pulse Emphasis Filter

A pulse emphasis filter is a filter that refines pulse components and strengthens their sharpness. Convex parts of the signal are detected and non-linear emphasis processing is applied to each pixel. The difference between a target pixel and the previous and following pixels is calculated to detect a block boundary, and the same processing is applied to the entire image. A flag between a block end and the start of a following block is set as a convex part from the processing described above, and it is possible to detect convex parts and to apply non-linear emphasis processing [11], [12].

III. Facial Correction

When super-resolution processing is performed on face images, deterioration of their images will be seen. This is due to the fact that wrinkles and stains on faces are emphasized by super-resolution processing. Therefore, it is necessary to perform facial correction processing on skin parts without super-resolution processing. An example of image degradation by super-resolution processing is shown in Fig. 2. A block diagram of a combination of super-resolution processing and facial correction processing is shown in Fig. 3. Super-resolution processing and facial correction processing are separately performed on an original image. Thereafter, in accordance with the skin color detection result, a facial corrected image is applied to the skin part, and a super-resolution image is applied to the non-skin part. In this paper, we try to remove wrinkles and stains in facial images by using non-linear filter. Before performing wrinkle and stain correction processing, we define them. The wrinkles are grooved irregularities formed by bending the skin. Therefore, it is assumed that the shape of the wrinkle exists as a generally straight line or a gentle curve. The stain is a part where melanin pigment is deposited on the skin and it turns black or red than the surrounding skin. The shape is assumed to be a small circle or a lump. Since wrinkles are shadows caused by the irregularities of the skin, it is considered that this is a sufficiently low or sufficiently high parts of the luminance value as compared with the average luminance of the peripheral region. Since stains are changes in color of skin, it is considered that it is included in the part where the luminance value is lower than the average luminance of the peripheral region on face images [13], [14].

(a) Original Image (b) Super-Resolution

Figure 2. Degradation due to super-resolution processing

A. Non-Linear Filter

We will explain a non-linear filter for reducing wrinkles and stains. By performing non-linear filter processing on the part detected in skin color detection, it is possible to generate a natural image that does not cause changes in the non-skin part. In this paper, we use the TV regularization filter as a non-linear filter.

IV. Skin Detection

Finally, super-resolution processing image and facial correction processing image are synthesized. At that time, the skin parts use the facial corrected image and the non-skin parts use the super-resolution image. Therefore, it is necessary to know the skin parts and the non-skin parts of the image. Therefore, we perform skin color detection. That is, the part to perform the super-resolution processing and the part to perform facial correction processing are determined according to the skin color detection result. If skin color detection is not correctly performed, the super-resolution processing may be performed on the part where face correction processing should be performed originally. In that case, wrinkles and stains to be removed are inversely emphasized, and the image is deteriorated. Therefore, skin color detection plays a considerably important role in the process. An example in which the final result image is deteriorated due to the failure of skin color detection is shown in Fig. 4.

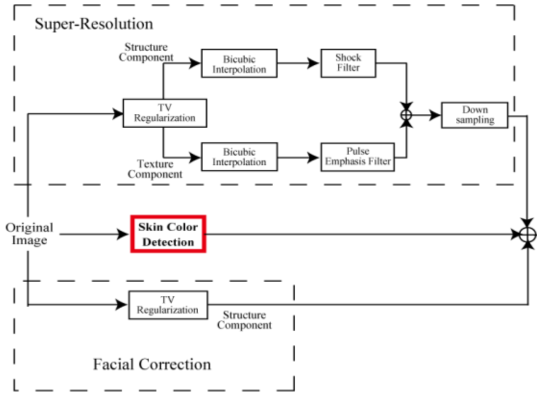
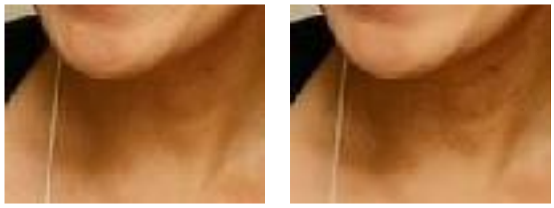


Figure 3. Block diagram of processing procedure



(a) Original Image (b) Super-Resolution

Figure 4. Image degradation due to skin color detection failure

A. Conventional Method (HSV Color Model)

HSV is a color model composed of three elements of hue (Hue), saturation (Saturation), and lightness (Value). When people routinely handle colors with applications etc., in most cases they use the RGB color model. However, in the RGB color region, the overall RGB value decreases in the dark places and conversely becomes high overall in the bright part, so that there arises a problem that the range is widened. In skin color detection, it is important to handle finer color regions, so conversion from the RGB color region to the HSV color region was done in the conventional method. Compared to RGB, HSV does not color mixture, it can designate colors with a sense close to selecting a color from a color sample, and it can be said that it is a color model that makes it easy to designate colors intuitively from the viewpoint of human beings. First, each value will be described. H component is a hue and has a value from 0 to 360. Human skin will be in the range of 0 to 30, but range may be finely shifted depending on shooting situation such as light.

S component is saturation and a value indicating vividness, and ranges from 0 to 255. It is thought that it is unnecessary to specify the range because the skin color part reflects light and becomes vivid.

V component is lightness, which is a numerical value indicating the brightness and darkness of color, and is in the range of 0 to 255 similarly to the saturation. This value has a value that greatly affects discrimination of hair and background in real images. However, if this range of values is taken too much, wrinkles or stains that you want to reduce on skin color parts will also be out of detection, so careful parameter setting will be required. In the conventional method, skin color detection was performed using a fixed threshold for setting the HSV parameters in advance [15], [16].

B. Proposed Method

In the HSV color conversion, although the skin parts, in which light is irradiated, can be properly detected, there is a problem that detection of a dark skin parts does not hit by light can not be performed well even with the same skin parts. Therefore, we propose a novel color conversion method specialized for skin color independently and use it for skin color detection. Also, in the conventional method, skin color detection is performed using a fixed threshold. Therefore, in the proposed method, an algorithm that automatically adjusts the threshold for each image is proposed, and skin color detection is performed.

1) Skin color axis conversion

We propose a skin color axis transformation from skin color trajectory. A total of 336 skin colors from seven kinds of facial images were plotted in RGB space, and their trajectories were created as a skin color axis. And, two axes orthogonal to each other on a plane perpendicular to this axis are produced. S, a and b axes are obtained as new detection parameters. Since S axis is a parameter specialized for skin color, skin color detection with higher accuracy will be expected to become possible by using this axis. The transform equation from RGB space to the skin color axis is shown in (19), and the three-dimensional plot and the appearance of the created axis are shown in Fig. 5.

$$\begin{pmatrix} S \\ a \\ b \end{pmatrix} = \begin{pmatrix} 0.4054 & 0.3524 & 0.3903 \\ 0.5639 & -0.2383 & -0.4840 \\ -0.1644 & 0.6125 & -0.5029 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} -61.75 \\ -23.63 \\ 3.264 \end{pmatrix} \tag{19}$$

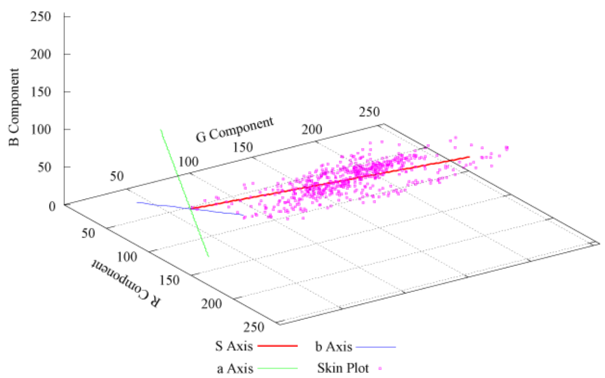


Figure 5. Plot of skin color and calculation of skin axis

2) Skin color detection algorithm

When a fixed threshold like the conventional method is used at the time of skin color detection, it takes long time and labor since it is necessary to set an appropriate threshold for each image on its own. Therefore, we propose an algorithm that automatically decides appropriate threshold for each image. The processing procedure of the algorithm is shown in Fig. 6.

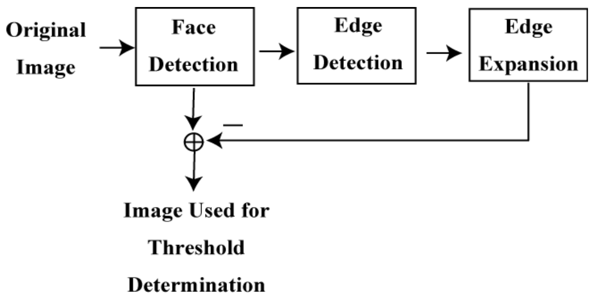


Figure 6. Block diagram of skin color detection procedure

First of all, face detection is performed, and the detection range is made relatively small. Doing this will make it less likely to contain non-skin parts (eyes, mouth etc.). Thereafter, edge detection is performed on the face detection image. In addition, expansion processing is performed on the edge detection result image. The edge parts thus obtained are subtracted from the face detection image. By the above processing, non-skin parts (eyes, mouth, etc.) in the face detection image can be removed. The pixel values of the face detection image from which the non-skin parts have been removed are read, and these are used as threshold for skin color detection. Through the above processing, it is possible to obtain an appropriate skin color detection threshold for each image. However, it is impossible to correctly detect parts of different brightness on skin. Therefore, a raster scan is performed on the mask image of the skin color detection result. In the case of a pixel, which is interest, is determined to be a skin part, the S element of the pixel of interest and the 8 neighboring pixels are compared. If both S elements are similar, adjacent pixels are considered as skin parts. Such processing is performed for all the pixels, and it is repeated until there is no change. The outline of this algorithm is shown in Fig. 7. By this processing, skin color parts can be enlarged, and skin color parts that can not be detected by skin color detection using threshold can be detected. We expect to detect skin colors more accurate compared as the conventional methods.

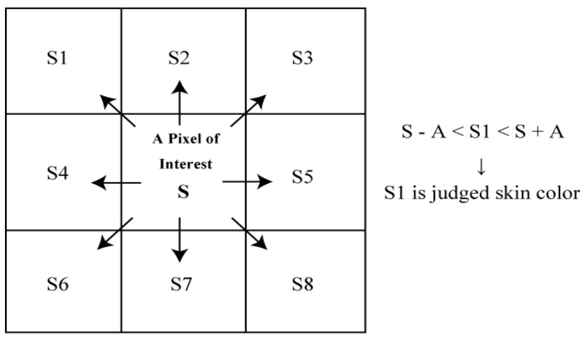


Figure 7. Outline of proposed method

V. IMAGE SYNTHESIS

The super-resolution processing image and the facial correction processing image are synthesized. The outline of synthesis method is shown in Fig. 8. Based on the skin color detection result, the facial correction processing image is used for the part detected as skin color, and the super-resolution processing image is used for the part detected as non-skin color. With this synthesis method, wrinkle or stains are removed from the skin part, and images with clearer eyes, hair, etc. can be obtained.

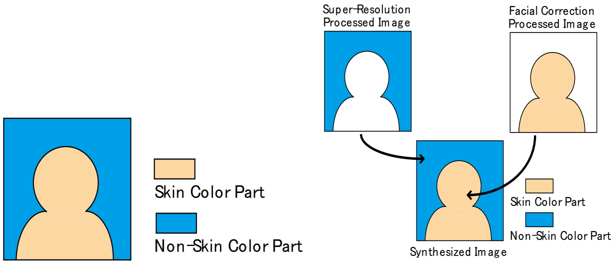


Figure 8. Outline of image synthesis

VI. Experimental Results

A. *Super-resolution Processing and Facial Correction Processing for Facial Images*

Experiments were conducted to confirm the performance of super-resolution processing and facial correction processing on facial images. Experimental conditions are shown in Table I and Table II. A linearly enlarged image, a super-resolution enlarged image, and a facial-corrected image are shown in Fig. 9. From Fig. 9, it can be seen that the super-resolution enlarged image of (b) has a sharper eye outline than the linear enlarged image of (a). At the same time, however, it is clear that the skin stains are also more clear. On the other hand, in the image obtained by combining the super-resolution processing of (c) with the facial correction processing, while maintaining the sharpness of the facial parts (eyes etc.), the skin stains which were conspicuous in (b) successful removal. This experiment showed that by combining super-resolution processing with facial correction processing, sharpness can be imparted to facial parts, and wrinkles and stains on the skin can be removed.

(a) Original image

(b) Super-resolution (c) Super-resolution + facial correction

Figure 9. Super-resolution and facial correction for facial image

TABLE I. Experimental Conditions (Super-Resolution)

TV Filter	
Strength of Restraint	0.06
Repeat Count	30
Step Width	0.125
Shock Filter	
Step Width	1.2
Threshold	3
LPF	Moving Average Filter 5×5

TABLE II. Experimental Conditions (Skin Detection)

HSV Condition	
Hue	0—20
Saturation	85—255
Value	170—255

B. *Comparison between Conventional Method and Proposed Method*

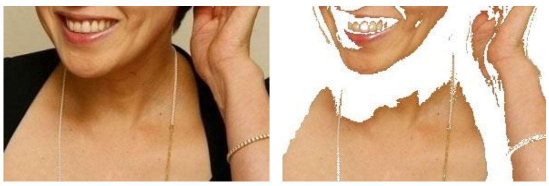
We compare the skin color detection results and facial super-resolution processing of the conventional method and the proposed method.

1) *Conventional method (HSV color model)*

First, the skin color detection result of the conventional method is shown in Fig. 10. The experimental conditions are shown in Table III. Based on the experimental results, the skin color part of the face can be generally detected. However, we can see that we can not detect the neck part, which is dark.

TABLE III. Experimental Conditions (Facial Correction)

TV Filter	
Strength of Restraint	0.06
Repeat Count	100
Step Width	0.125
HSV Condition	
Hue	0—20
Saturation	85—255
Value	170—255



(a) Original image (b) Skin detection

Figure 10. Skin detection (conventional method)

2) *Proposed method*

Next, the images before and after the raster scanning process of the skin color detection results of the proposed method are shown in Fig. 11. In the case of the proposed method, the threshold is automatically determined from the image, so it is not necessary to set it in advance. Experimental results show that the skin color part of the face can be generally detected in the image before the raster scan, but the detection of the neck part is incomplete. On the other hand, in the image after the raster scan, the detection of the neck part is complete.

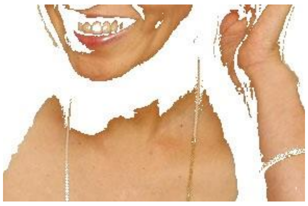
3) Comparison of results (skin detection)

(a) Original image

(b) Skin detection (before the raster scanning) (c) Skin detection (after the raster scanning)

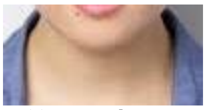
Figure 11. Skin detection (proposed method)

We compare the skin color detection results of the conventional method and the proposed method. The skin color detection results are shown in Fig. 12. From the results, there is a big difference in the detection of the neck part between the conventional method and the proposed method, and the superiority of the proposed method can be seen. Moreover, the proposed method is superior to the conventional method also in that no threshold setting is made in advance. The results in another image are shown in Fig. 13. Also in this image we can see that the proposed method can detect skin color with higher accuracy than the conventional method.



(a) Conventional method (b) Proposed method

Figure 12. Comparison of results 1 (skin detection)



(a) Original image

(b) Conventional method (c) Proposed method

Figure 13. Comparison of results 2 (skin detection)

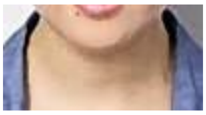
4) Comparison of results (facial super-resolution processing)

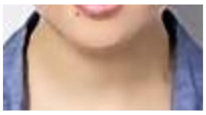
We compare the results of facial super-resolution processing of conventional method and proposed method. The results of performing facial super-resolution processing using each skin color detection results is shown in Fig. 14. Experimental results show that the conventional method has applied super-resolution processing to the neck part and wrinkles in the neck are conspicuous. On the other hand, in the proposed method, it is possible to perform facial correction processing on the neck part, which shows that wrinkles can be removed. A comparison of the facial super-resolution processing results for another image is shown in Fig. 15. It can be seen that a more beautiful result image could be obtained by highly accurate skin color detection even in this image.

By the above experiments, it was possible to demonstrate the superiority of the proposed method.

(a) Conventional method (b) Proposed method

Figure 14. Comparison of results 1 (facial super-resolution processing)





(a) Conventional method (b) Proposed method

Figure 15. Comparison of results 2 (facial super-resolution processing)

VII. Conclusion

In this paper, we proposed a facial correction process such as elimination of wrinkles and stains, emphasis of eyes and hair by using super-resolution technology and TV regularization filter for facial images.

In the experimental results, it was possible to impart sharpness to eyes, hair, etc. and to emphasize it by super-resolution processing. At the same time, however, it was confirmed that the wrinkles of the skin were also emphasized. Therefore, wrinkles and stains on the skin color part were removed by facial correction processing, and hair and eyes were emphasized by super-resolution processing in the non-skin color part. By combining super-resolution processing and facial correction processing, image quality could be improved.

In addition, since the processing performed in the skin color part and the non-skin color part of the face image are different from each other, skin color detection is performed. At that time, in the conventional method, skin color detection was performed using the HSV color space. However, with this method, it was difficult to simultaneously detect both the part to which the light of the face is hit and the part not hit. Therefore, we proposed the proprietary color expression method specialized for skin color and the method to automatically determine thresholds. With this method, it was also possible to detect dark parts that were incompletely detected in the conventional method. In addition, there is no need to set thresholds suitable for the processed image beforehand. In addition, we succeeded in improving the image quality of facial super-resolution processing by making it possible to detect skin color with higher accuracy. Future tasks include increasing the processing speed.

REFERENCES

[1] S. J. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, pp. 259–268, 1992.

[2] Y. Meyer, "Oscillating patterns in image processing and nonlinear evolution equation," *Fifteenth Dean Jacqueline B. Lewis Memorial Lectures, University Lecture Series*, vol. 22, 1992.

[3] S. J. Osher and L. A. Vese, "Modeling textures with total variation minimization and oscillating patterns in image processing," *Journal of Scientific Computing*, vol. 19, no. 5, p. 145, 2003.

[4] L. Blanc-Feraud J. F. Aujol, G. Aubert, and A. Chambolle, *European Conference on Computer Vision (ECCV)*, Sep. 2014, pp. 783–798.

[5] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1, pp. 89–97, 2004.

[6] A. Chambolle, "Total variation minimization and a class of binary MRF models," in *Proc. International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, 2005, vol. 60, pp. 136–152.

[7] S. J. Osher and L. I. Rudin, "Feature-Oriented image enhancement using shock filters," *SIAM Journal on Numerical Analysis*, pp. 310–340, 1990.

[8] L. Alvarez and L. Mazorra, "Signal and image restoration using shock filters and anisotropic diffusion," *SIAM Journal on Numerical Analysis*, pp. 590–605, 1994.

[9] M. Watanabe, T. Goto, S. Hirano, M. Sakurai, and Y. Sakuta, "Super-Resolution through non-linear enhancement filters," in *Proc. IEEE International Conference on Image Processing*, Sep. 2013, pp. 854–858.

[10] J. G. M. Schavemaker, M. J. T. Reinders, J. J. Gerbrands, and E. Backer, "Image sharpening by morphological filtering," *Elsevier Pattern Recognition*, vol. 33, no. 6, pp. 997–1012, Jun. 2000.

[11] T. Munezawa, T. Goto, and S. Hirano, "Super-resolution method for enlarged images utilizing non-linear filtering," in *Proc. IEEE 6th Global Conference on Consumer Electronics*, 2017, pp. 1–2.

[12] K. Goto, F. Nagashima, T. Goto, S. Hirano, and M. Sakurai, "Super-resolution for high-resolution displays," in *Proc. IEEE 3rd Global Conference on Consumer Electronics*, 2014, pp. 309–310.

[13] T. Goto, *et al.*, "Face image processing by TV filter and super-resolution," in *Proc. of VCIP*, 2014, pp. 245–248.

[14] T. Goto, S. Hirano, and M. Sakurai, "Face image processing by TV filter and super-resolution," in *Proc. IEEE Visual Communications and Image Processing Conference*, 2014, pp. 245–248.

[15] B. Muhammad and S. A. R. Abu-Bakar, "A hybrid skin color detection using HSV and YCgCr color space for face detection," in *Proc. IEEE International Conference on Signal and Image Processing Applications*, 2015, pp. 95–98.

[16] M. A. Rahman, I. K. E. Purnama, and M. H. Purnomo, "Simple method of human skin detection using HSV and YCbCr color spaces," in *Proc. International Conference on Intelligent Autonomous Agents, Networks and Systems*, 2014, pp. 58–61.

Keigo Kano received the B.S. degree from the Department of Computer Science and Engineering, Nagoya Institute of Technology, in 2017. He is currently pursuing the M.S. degree from the Department of Computer Science, Nagoya Institute of Technology, Japan. His research interests are signal and image processing and super-resolution processing.

Tomio Goto (M'01) received the B.S., M.S., and Ph.D. degrees from Nagoya Institute of Technology in 1997, 1999, and 2007, respectively. He is currently an associate professor with the Department of Computer Science, Nagoya Institute of Technology, Japan. His research interests include coding of images, non-linear filtering, noise reduction, and super-resolution, as well as most other applications of digital image processing. He is a senior member of the IEEE, IEICE, ITE, and IEEJ in Japan

Satoshi Hirano received the B.S., M.S., and Ph.D. degrees from Nagoya Institute of Technology in 1985, 1987, and 1991, respectively. He is currently an associate professor with the Department of Computer Science, Nagoya Institute of Technology, Japan. His research interests include analog to digital converter and filter circuit, and so on. He is a member of the IEEE, IEICE, and IEEJ.