Analysis and Synthesis of Facial Color Using Color Image Processing*

Tsutomu KURODA** and Tomio WATANABE**

The natural color of five representative areas of the human face, the lip, brow, nose, left cheek and right cheek, is analyzed using the HSV (hue, saturation, value) color space that is fitted with human color sense. From the relationships among the HSV parameters for these areas, it is found that the hue of the lip is not correlated with that of the other areas. This suggests that the hue of the lip in a virtual facial image can be treated independently from other facial colors. A virtual facial color image is synthesized using the average facial color in HSV color space determined from 64 males. On the basis of the difference in color of each pixel between the average facial image and the individual one, a method for emphasizing the individual characteristics in facial color is proposed. The effectiveness of the method is demonstrated through sensory evaluation by paired comparisons of enhanced virtual facial images.

Key Words: Color Image Processing, Human Interface, Facial Color, Virtual Facial Image, Characteristic Emphasis

1. Introduction

The facial image plays an important part in smooth communications between humans and computer systems functioning as virtual agents. Research aimed at the improvement of the human interface has been carried out using images of faces such as the analysis of synthesis of image coding system for the human face\(^{(1,2)}\). However, this research focuses mainly on facial expressions, and facial colors are taken care of only by texture mapping onto the surface. Though facial color is expected to play an important role in identifying individuals, making the virtual facial image effective and displaying one's emotion, so far no research has been carried out on facial color itself.

We reported a method for lip extraction from the face image by color image processing using HSV color space that is fitted with human color sense and found that this method could improve the extracting precision\(^{(9)}\). In this paper, we propose a method for the analysis and synthesis of facial color using color image processing, with the objective of constructing a facial color database and producing virtual face images. We first divide the face image into five areas including the lip, left cheek, right cheek, nose and brow, and then analyze the relationships among the HSV parameters of these areas using correlation analysis. By synthesizing a virtual facial color image with the average facial color in HSV color space, we propose a method for emphasizing individual characteristics in facial color on the basis of the difference in color of each pixel between the average face image and an individual one. Finally we demonstrate the effectiveness of the proposed method.

2. Quantitative Analysis of Facial Color

2.1 HSV color space

The HSV color space is used to express color by the three parameters of hue ($H$), saturation ($S$: colorfulness) and value ($V$: lightness), which are represented in a cylindrical coordinate system defining
2.2 Apparatus and measuring conditions

Figure 1 shows the set up of the experiment to take a face image. To fix the light source condition, the measurement is carried out in a dark room where outdoor daylight is excluded, and unwanted reflection of the light in the dark room is prevented by a black-out curtain behind the examinee. The color image processing unit (Mitani LMC-512V8) used in this research provides 16,770,000 colors (256 levels per each RGB) in each of 512x480 pixels.

For a light source, four daylight photo-reflector lamps (color temperature: 5500°K) were used. The distance between the lamps and the examinee’s face was adjusted to be about 100 cm, and the lamps were set to be 120 cm above the floor so that the lamps and the face might be at the same level. The distance between the camera and the examinee was set at 100 cm so that the face occupied most of a screen.

The distance between the examinee and the lens was adjusted to 100 cm. The lens aperture was adjusted for the condition that the value of the V-parameter of the white JIS standard color chip (N 9.0) was 80. At that time, the output ratio of the camera’s RGB signals was controlled so that the saturation of the image taken of the color chip might be achromatic, or S=0. By this adjustment, the influence of the color temperature was minimized.

The examinees were 64 male students ranging from 18 to 21 in age, and they closed their lips and revealed their foreheads when their images were taken.

2.3 Definition of analyzing area

We chose five areas of the face to be measured such as the lip (LP), left cheek (LC), right cheek (RC), nose (NS) and brow (BR) as shown in Fig. 2. The coordinates of the apex of each area were calculated according to the coordinates of the centers of both eyes (E1, E2), right and left nasal cavities of the nose (N1, N2), upper and lower lips (L1, L2) and the center of the mouth (Lc).

We define the coordinates of each measured point as follows. Incidentally, the absolute size of each area differs among individuals.

Right eye (E1) : (x_{e1}, y_{e1}), Left eye (E2) : (x_{e2}, y_{e2})
Right nasal cavity (N1) : (x_{n1}, y_{n1}), Left nasal cavity (N2) : (x_{n2}, y_{n2})

Upper lip (L1) : (x_{l1}, y_{l1}), Lower lip (L2) : (x_{l2}, y_{l2})
Center of a mouth (Lc) : (x_{lc}, y_{lc})
Midpoint between the eyes (Ec) : (x_{ec}, y_{ec})
Provided that : x_{ec}=(x_{e1}+x_{e2})/2, y_{ec}=(y_{e1}+y_{e2})/2

Midpoint between both nasal cavities (Nc) : (x_{nc}, y_{nc})
Provided that : x_{nc}=(x_{n1}+x_{n2})/2, y_{nc}=(y_{n1}+y_{n2})/2

On the basis of the coordinates of the measured points, the position of each area and the equations to find the coordinates for each apex are as follows;

(1) Brow (BR)
In the direction of the x-axis:
Middle of the trisect between the both eyes.
In the direction of the y-axis:
Lower side: Midpoint between the eyes.
Upper side: Half the distance between the eyes and the nasal cavities above the lower side.

\[
\begin{align*}
  x_{br1} &= x_{ec} + |x_{e1} - x_{e2}|/3 \\
  x_{br2} &= x_{ec} + 2|x_{e1} - x_{e2}|/3 \\
  y_{br1} &= y_{ec} \\
  y_{br2} &= y_{ec} - |y_{n1} - y_{n2}|/2
\end{align*}
\]  \hspace{1cm} (1)

(2) Right cheek (RC)
\[ x_{rc} \] axis:
At the same distance as the one between the nasal cavities rightwards from the right eye.
\[ y_{rc} \]
Upper side: Middle of the trisect between the right eye and the right nasal cavity.

Lower side: At the position of the right nasal cavity.

\[
x_{sc1} = x_{el} \\
x_{sc2} = x_{sc1} + \frac{|x_{el} - x_{sc2}|}{2} \\
y_{sc1} = y_{el} + \frac{|y_{el} - y_{ns}|}{3} \\
y_{sc2} = y_{el} \tag{2}
\]

(3) Left cheek (LC)
The coordinates symmetrical with those of the right cheek are used.

\[
x_{lc1} = x_{ea} \\
x_{lc2} = x_{lc1} + \frac{|x_{ea} - x_{lc2}|}{2} \\
y_{lc1} = y_{ea} + \frac{|y_{ea} - y_{ns}|}{3} \\
y_{lc2} = y_{ea} \tag{3}
\]

(4) Nose (NS)
In the direction of the \( x \)-axis:
The value between the nasal cavities.
In the direction of the \( y \)-axis:
The middle of the trisect between the eyes and the nasal cavities.

\[
x_{ns1} = x_{el} \\
x_{ns2} = x_{ns} \\
y_{ns1} = \frac{y_{nc} + y_{ns}}{2} \\
y_{ns2} = 2 \times \frac{y_{nc} + y_{ns}}{3} \tag{4}
\]

(5) Lip (LP)
In the direction of the \( x \)-axis:
The value between the nasal cavities.
In the direction of the \( y \)-axis:
The value entered from the screen.

\[
x_{lp1} = x_{na} \\
x_{lp2} = x_{lp1} \\
y_{lp1} = y_{l} \\
y_{lp2} = y_{l} \tag{5}
\]

The apex of each area is obtained from combinations of these values of the coordinates.

2.4 Range of the areas in HSV color space
The value of hue of the images of the human face ranges from 340 to 359 and 0 to 40 (from reddish purple to red or orange), while that from 90 to 300 (from green to blue) does not exist. The range of saturation is over 20, and that of value is over 30. In each area, the average of each HSV parameter in all pixels is taken as a representative HSV value calculated to one decimal place.

3. Correlation of Facial Color among Areas

3.1 Distribution of hue and saturation
Figure 3 shows the distribution on the H-S plane for each area. According to Fig. 3, the distribution for the lip on the H-S plane is different from that of the other areas. It was also found that the saturation of the nose was higher than that of the cheek and brow.

---

**Table 1 Correlations of hue**

<table>
<thead>
<tr>
<th>Area</th>
<th>NS</th>
<th>BR</th>
<th>RC</th>
<th>LC</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP</td>
<td>-0.001</td>
<td>0.031</td>
<td>-0.073</td>
<td>-0.137</td>
</tr>
<tr>
<td>LC</td>
<td>0.516</td>
<td>0.531</td>
<td>0.937</td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>0.519</td>
<td>0.570</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR</td>
<td>0.586</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3 Distribution in the \( H \)-S plane for each area

Fig. 4 Correlation of the hue of lip to that of the other areas

Fig. 5 Average and standard deviation of hue for each area
the hue in each area. The correlation coefficients between the lip (LP) and the other areas were almost 0, and the highest coefficient was \(-0.137\) at the left cheek (LC). On the assumption that there is no correlation \((H_0: r=0)\) of hue between the lip and the other areas, the threshold of \(r\) was 0.177 at the significance level of 5% for the degree of freedom 62 (the number of data minus 2) using a two-sided test in \(t\)-distribution, and the assumption was not rejected. Therefore, we can assume that there is no correlation of the hue between the lip and the other areas. This result indicates that the lip is independent of other areas with respect to hue, and it is, therefore, enough to focus on the lip area for extracting the lip. Accordingly, this suggests that the effectiveness of our lip extraction method based on hue is guaranteed\(^{39}\).

We can find the high correlations between the areas except the lip. In particular, the correlation coefficient between the two cheeks (LC, RC) is 0.937.

### 3.3 Correlations of saturation

The correlation coefficients among the areas are shown in Table 2. We also find a high correlation between areas except the lip. The correlation coefficient between the two cheeks is especially good at 0.945.

Figure 6 shows the average and standard deviation of saturation in each area, and the areas except the lip area and brow have almost the same saturation. After the lip extraction, we can extract the area of the nose using saturation.

### 3.4 Correlations of value

The correlation coefficients between the areas are shown in Table 3. There are close correlations among the areas. It is suggested that individual characteristics appear as the value for a whole face.

---

**Table 2 Correlations of saturation**

<table>
<thead>
<tr>
<th>Area</th>
<th>NS</th>
<th>BR</th>
<th>RC</th>
<th>LC</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP</td>
<td>0.358</td>
<td>0.404</td>
<td>0.258</td>
<td>0.264</td>
</tr>
<tr>
<td>LC</td>
<td>0.791</td>
<td>0.772</td>
<td>0.945</td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>0.791</td>
<td>0.772</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR</td>
<td>0.772</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 3 Correlations of value**

<table>
<thead>
<tr>
<th>Area</th>
<th>NS</th>
<th>BR</th>
<th>RC</th>
<th>LC</th>
</tr>
</thead>
<tbody>
<tr>
<td>LP</td>
<td>0.579</td>
<td>0.519</td>
<td>0.573</td>
<td>0.511</td>
</tr>
<tr>
<td>LC</td>
<td>0.818</td>
<td>0.748</td>
<td>0.939</td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td>0.866</td>
<td>0.725</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR</td>
<td>0.690</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Figure 7 shows the average and standard deviation of value in each area, and the difference between the lip and the other areas is not as distinct as seen for hue and saturation.

In summary, we clarified that the hue of the lip is independent of the other areas. We also found the color symmetrical on the face from the high correlation of hue, saturation and value between the right and left cheeks.

---


### 4.1 Standardization of facial image size

As the absolute size of a facial image depends on the size of an individual face and the distance between the subject and the camera, the size of a face was standardized as follows: In the direction of the \(x\)-axis, we define the width of a face, denoted as \(W\), as the distance between two eyes. In the direction of the \(y\)-axis, we define the height of the face, denoted \(T\), as the distance between the eyes and the mouth.

The definitions of the width and the height of a facial image using the notations introduced in 2.3 are:

\[
W = |x_{\text{left}} - x_{\text{right}}| \quad \text{and} \quad T = |y_{\text{top}} - y_{\text{bottom}}|
\]

(6)

The definitions of the average width \(\bar{W}\) and the height \(\bar{T}\) are:

\[
\bar{W} = \frac{\sum_{i=1}^{n} W_i}{n}
\]

(7)

\[
\bar{T} = \frac{\sum_{i=1}^{n} T_i}{n}
\]

The individual sizes of face were enlarged or reduced both in the direction of the \(x\)-axis and the \(y\)-axis for standardization to make \(W\) and \(T\) of each examinee equal to \(\bar{W}\) and \(\bar{T}\). The standardized facial
images were synthesized by enlarging, reducing or shifting the examinee's facial images.

4.2 Synthesis of the average facial color image

The average facial color image was synthesized by calculating the average of each HSV parameter in every corresponding pixel of the each examinee's standardized facial image. Figure 8 shows the average facial color image obtained.

4.3 Method for emphasizing the individual characteristics using facial color

The position of a pixel on the standardized facial image of an examinee is expressed by the notation $i (0 \leq i \leq 511)$ in the direction of the $x$-axis and the notation $j (0 \leq j \leq 479)$ in the direction of the $y$-axis. The HSV parameters at the positions $(i, j)$ were expressed as $H_i$, $S_i$, and $V_i$. In the same way, we denote the HSV parameters for the average facial color image by $H_o$, $S_o$, and $V_o$. The HSV parameters of the emphasized facial color image by $h_o$, $s_o$, and $v_o$. The HSV parameters of an emphasized facial color image can be obtained from Eq. (8), in which the color distance between the average facial color and an individual's facial color was multiplied by a, an emphasis coefficient, and added to the average color.

\[
\begin{align*}
h_o & = a(H_o - H_i) + H_i \\
s_o & = a(S_o - S_i) + S_i \\
v_o & = a(V_o - V_i) + V_i \\
\end{align*}
\]

(8)

$a$: emphasis coefficient

However, in case where a one-to-one correspondence for each pixel was not found between the standardized facial image of an examinee and the average facial color image, data from the standardized facial image of the examinee were used preferentially as shown in Eq. (9).

\[
\begin{align*}
h_o & = H_o \\
s_o & = S_o \\
v_o & = V_o \\
\end{align*}
\]

(9)

The emphasis coefficient $a$ can be entered arbitrarily when the facial color image is synthesized.

4.4 Synthesis of the emphasized facial color image using hue

We synthesized an emphasized facial color image using the hue, which was expected to be the most effective parameter in color.

Figure 9 shows the emphasis on hue in the case of $a=1$. The notation ' shows the distribution of the left cheeks of all the examinees in the HS plane, while ● shows the average ($H=18.2$) of all the examinees. The notation ▲ shows the measured value ($H=23.5$) of the left cheek of the examinee who had the highest hue, and the notation ▼ shows the emphasized value ($H=28.8$) when $a=1$.

4.5 Evaluation of the emphasized facial color image

Two of the three kinds of emphasized facial images under conditions that $a=0$ (original facial color image), 0.5 and 1 were displayed side-by-side on a color monitor (SONY KV21-HVS) directly connected through RGB signals from the color image processor. Figure 10 shows an example of the displayed image in which the left is an emphasized facial color image ($a=1$) and the right is a standardized facial image ($a=0$). Considering the difference of position in left and right, six paired images were synthesized by switching right and left for the sensory evaluation by paired comparison. The judges were three students who knew the examinee, and chose the image which urged the examinee's facial color.

The emphasis coefficients $a$ were set at 0.5 and 1.0 in this experiment because the emphasized facial color images when $a=1.5$ and 2.0 were found unnatural in a preliminary experiment.

The results are shown in Table 4. To evaluate
Table 4 Paired comparison of enhanced hue

<table>
<thead>
<tr>
<th>i</th>
<th>0</th>
<th>0.5</th>
<th>1</th>
<th>Total</th>
<th>σ^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td></td>
<td>4.73</td>
</tr>
<tr>
<td>0.5</td>
<td>7</td>
<td>2</td>
<td>9</td>
<td></td>
<td>7.36</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td></td>
<td>17.91</td>
</tr>
</tbody>
</table>

the results quantitatively, we assumed the following Bradley-Terry model:

\[ P_{ij} = \pi_i / (\pi_i + \pi_j) \]

\[ \sum \pi_i = \text{const} \quad (\text{average}=10) \quad (\pi_i : \text{intensity of } i, \quad P_{ij} : \text{probability of judgment that } i \text{ is better than } j) \]

The model is able to determine the resemblance based on the paired comparison in Table 4. The value of \( \pi_i \) for the enhanced image when \( \alpha = 1 \) was higher than that for the original image when \( \alpha = 0 \). This indicates that the individual characteristics are emphasized by shifting the hue as much as the difference between the average and itself.

5. Conclusions

In this paper, to quantify human facial color by using HSV color space which is fitted with human color sense, we analyzed the correlations regarding color among five areas including the lip, brow, nose, left cheek and right cheek. We also synthesized an average facial color image and proposed a method for emphasizing the individual characteristics of facial color on the basis of the difference in color of corresponding pixels between the average facial color image and an individual facial color image, and demonstrated the effectiveness of the method. The summary of the results is as follows:

1. In the male examinees ranging from 18 to 21 years old, there is no relation between the hue of the lip and that of the other areas. Therefore, for synthesizing natural virtual facial images in the same generation, the hue of the lip can be treated independently from that of the other areas such as the cheeks, and to extract the lip from the facial image, it is sufficient to pay attention to the area around the lip.

2. Through sensory evaluation by paired comparisons of the enhanced facial color images based on the average facial color image, it was found that the enhancement of facial color by the proposed method increased individual characteristics. This indicates that facial color is the important attribute among the individual characteristics.

To generalize the result from the males ranging from 18 to 21 years old as subjects, we plan to analyze and evaluate facial color data for every generation. Moreover, the analysis and evaluation of facial color data covering a wider range and various conditions including the distinction of sex and seasons and the application to the coloration of virtual face images are subjects for a further study.

Acknowledgement

This research is supported in part by the Telecommunications Advancement Foundation.

References


