REGULAR PAPER



Auto white balance method using a pigmentation separation technique for human skin color

Satomi Tanaka¹ · Akihiro Kakinuma² · Naohiro Kamijo² · Hiroshi Takahashi² · Norimichi Tsumura¹

Received: 17 May 2015/Accepted: 5 November 2016/Published online: 15 November 2016 © The Optical Society of Japan 2016

Abstract The human visual system maintains the perception of colors of an object across various light sources. Similarly, current digital cameras feature an auto white balance function, which estimates the illuminant color and corrects the color of a photograph as if the photograph was taken under a certain light source. The main subject in a photograph is often a person's face, which could be used to estimate the illuminant color. However, such estimation is adversely affected by differences in facial colors among individuals. The present paper proposes an auto white balance algorithm based on a pigmentation separation method that separates the human skin color image into the components of melanin, hemoglobin and shading. Pigment densities have a uniform property within the same race that can be calculated from the components of melanin and hemoglobin in the face. We, thus, propose a method that uses the subject's facial color in an image and is unaffected by individual differences in facial color among Japanese people.

Keywords Illuminant color estimation \cdot Color collection \cdot Auto white balance \cdot Human skin color model

Part of this paper was presented at the SPIE'2015 conference held in San Francisco USA.

 Satomi Tanaka satomi.st.tanaka@nts.ricoh.co.jp
 Norimichi Tsumura tsumura@faculty.chiba-u.ac.jp

¹ Graduate School of Advanced Integration Science, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

² Ricoh Company, Ltd., 16-1 Shinei-cho, Tsuzuki-ku, Yokohama, Kanagawa 224-0035, Japan

1 Introduction

The human visual system has a feature called color constancy that maintains a person's perceived color of the same object under various light sources, yet the object's physical color strongly depends on the illuminant color [1]. Current digital cameras have this feature of the human visual system in the form of an auto white balance (AWB) function, where the illuminant color is estimated and the output color is corrected as if the photograph was taken under a specific light source.

Many studies have proposed methods of estimating the illuminant color from an image [2]; e.g., methods using chromatic scene statistics [3, 4] and specular highlights [5]. To compare the estimated results and the ground truth, these illuminant color estimation methods use a key object whose color (RGB or spectroscopic information) is known, such as a gray board. The color of the key object is known in advance, and the illuminant color is then calculated from the change in the key object's color in the photograph. If a photographer takes a photograph with this key object, the illuminant color can always be obtained. However, using the gray board is not practical in nonprofessional use. Furthermore, the main subject in a photograph is often a person, and there is a strong desire to appropriately color correct the photograph in this case [6]. Here, color correction might be achieved using the person's face as the key object.

However, even among Japanese people, the color of facial skin varies according to the individual. Table 1 shows that the maximum chromaticity difference is 4.07 for mean values of the chromaticity for the areas of the cheek, chin, nose and mouth extracted from seven Japanese facial images. Here, the chromaticity difference is calculated using Δa^* and Δb^* in the CIELAB color space. Furthermore, the skin color is affected by pigment irregularities, and cannot be represented as a single value. Table 1Chromaticitydifference between the meanRGB values extracted fromseven subjects

| | Subject's number | | | | | | | |
|------------------|------------------|------|------|------|------|------|------|------|
| Subject's number | | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | 1 | 0.00 | 1.54 | 2.57 | 1.46 | 4.07 | 1.88 | 2.73 |
| | 2 | 1.54 | 0.00 | 1.30 | 1.34 | 2.53 | 0.40 | 1.67 |
| | 3 | 2.57 | 1.30 | 0.00 | 2.63 | 2.01 | 1.33 | 0.50 |
| | 4 | 1.46 | 1.34 | 2.63 | 0.00 | 3.38 | 1.36 | 3.01 |
| | 5 | 4.07 | 2.53 | 2.01 | 3.38 | 0.00 | 2.22 | 2.37 |
| | 6 | 1.88 | 0.40 | 1.33 | 1.36 | 2.22 | 0.00 | 1.78 |
| | 7 | 2.73 | 1.67 | 0.50 | 3.01 | 2.37 | 1.78 | 0.00 |

The cells of the first row and third column represent the chromaticity differences between subjects 1 and 3 $\,$

In this paper, we propose a white balance algorithm that uses the human face in a photograph as the key object, and considers the individual variation in the skin color based on a human skin color model [7-9]. Following [7], in our approach, the individual variation among people of the same race (such as the Japanese race) corresponds to the difference in the density of melanin and hemoglobin, which are the main pigments found in human skin. Assuming that the minimum densities of the two pigments are uniform among individuals of the same race, we extract the bias component that represents the illuminant color from the skin color of the person in the photograph. Furthermore, we correct the color of six photographs taken under six light sources first using our algorithm based on the human skin color model and then using a method that assumes that the human skin color does not have individual variation. It is found that the color correction using our approach based on the human skin color model outperforms the method without individual variation.

The remainder of the paper is organized as follows. Section 2 introduces the imaging model used in related works on illuminant color estimation and color correction [10–13] and the human facial skin color analysis as a pigmentation separation method [7–9]. Section 3 introduces the AWB method using the pigmentation separation method. Section 4 describes an experiment on color correction using our AWB method. Section 5 presents a discussion and suggests directions of future research.

2 Related work

2.1 Imaging model of illuminant color estimation and color correction

This section introduces the imaging model of the illumination color estimation and the color correction introduced in [2]. The AWB is achieved by estimating the illuminant color from an image and transforming an input image to an output image that is as taken under a certain light source. We discuss first the object color model of an image and then the transformation resulting from a change in illuminant color.

Under the Lambertian assumption [14, 15], image values f_i depend on an object's surface reflectance $S(x, \lambda)$, the light source $I(\lambda)$ and the camera sensitivity function $C_i(\lambda)$, where $i = \{R, G, B\}$, λ is the wavelength of the light and x is the spatial coordinate. The sensitivity function of a regular camera can be approximated by delta functions that have peaks near the wavelengths of red, green and blue light. The dependence of image values f_i on these factors is expressed as

$$f_i(x) = m(x) \int_{\omega} S(x,\lambda) I(\lambda) C_i(\lambda) d\lambda, \qquad (1)$$

where ω is the visible spectrum and m(x) is a scale factor that models the Lambertian shading that contributes to the overall light reflected at location x. Assuming that the scene is illuminated by a single light source and that the illuminant color e_i depends on the color of the light source Ias well as the camera sensitivity function C_{i} , e_i is given as

$$e_i = \int_{\omega} I(\lambda) \cdot C_i(\lambda) \, d\lambda. \tag{2}$$

All *S*, *I* and *C* are unknown, and estimating the illuminant color from an image is an under-constrained problem without assumptions. Assuming that the object color S_i can be expressed as

$$S_i(x) = \int_{\omega} S(x,\lambda) \cdot C_i(\lambda), \qquad (3)$$

the image value is

$$f_i(x) = S_i(x) \ e_i. \tag{4}$$

Knowing the object color S_i in advance allows the estimation of the illuminant color e_i using Eq. (4).

Color correction is often modeled using a linear transformation, which can be simplified as a diagonal transformation [10, 11, 13]. In this paper, we use the diagonal transform or von Kries model [12], expressed as

$$f_t = D_{u,t} f_u, \tag{5}$$

where $f = (f_R, f_G, f_B)^T$, f_u denotes the pixel values of an image taken under an unknown light source, and f_t denotes the resultant pixel values of an image transformed under the arbitrary light source. **D**_{*u*,*t*} is a diagonal matrix that maps colors obtained under an unknown light source to their corresponding colors obtained under the arbitrary illuminant *t*. Using Eq. (4), the diagonal matrix **D**_{*u*,*t*} is rewritten as

$$D_{u,t} = \begin{pmatrix} \frac{e_R^t}{e_R^u} & 0 & 0\\ 0 & \frac{e_G^t}{e_G^u} & 0\\ 0 & 0 & \frac{e_B^t}{e_B^u} \end{pmatrix},$$
(6)

where e_i^u and e_i^t are the illuminant colors of the unknown and arbitrary light sources, respectively. Using this imaging model, if we know the illuminant color e_i^u from an image, it is easy for us to correct the color using Eqs. (5), (6).

2.2 Imaging model of the pigmentation separation method

Section 2.1 presented the imaging model of the general object. This section introduces the imaging model of human skin used in the literature to separate pigmentation from the human skin color [7-9].

Igarashi et al. assumed a schematic model of the human skin as shown in Fig. 1 [16, 17]. The human skin has layers that contain various pigments such as melanin, hemoglobin, and bilirubin. Melanin and hemoglobin are mostly contained in the epidermis and dermal layers, respectively. Tsumura et al. [7–9] assumed that (i) spatial variations in skin color are caused by spatial variations in the quantities



Fig. 1 Schematic model of human skin with layers

of the pigments melanin and hemoglobin and (ii) the pigment quantities are mutually independent spatially. We further assume that (iii) light reflected from human skin follows the modified Lambert–Beer law [18] and (iv) the spectral distribution of skin does not abruptly change in the imaging system. The third and fourth assumptions ensure linearity in the optical density domain [19]. The linearity and assumptions (i) and (ii) allow us to calculate the image values f_i of human skin as

$$f_{i}(x) = k \int_{\omega} \exp\{-\rho_{m}(x) \sigma_{m}(\lambda) l_{m}(\lambda) -\rho_{h}(x) \sigma_{h}(\lambda) l_{h}(\lambda)\} I(x,\lambda) C_{i}(\lambda) d\lambda,$$
(7)

where $i = \{R, G, B\}$, k is a constant determined from the gain of the camera, ω is the visible spectrum, x is the positional variable and λ is the wavelength of the light. $\rho_m(x)$, $\rho_h(x)$, $\sigma_m(\lambda)$, and $\sigma_h(\lambda)$ are the pigment densities and the spectral cross sections of melanin and hemoglobin. $l_m(\lambda)$ and $l_m(\lambda)$ are the mean path lengths of photons in the epidermis and dermis layers, respectively. $I(x,\lambda)$ and $f_i(x)$ are the incident spectral radiance and reflected spectral radiance, respectively, at the position x. C_i (λ) is the camera sensitivity function. In addition, it is assumed that the lighting environment is distant and that its spectrum does not vary with direction. The illuminant color can be written as $I(x, \lambda) = p(x) \bar{I}(\lambda)$, where p(x) encodes shapeinduced shading variation and $\overline{I}(\lambda)$ is the position-independent incident spectral radiance. Furthermore, approximating $C_i(\lambda)$ by the delta function $\delta(\lambda - \lambda_i)$ [20] gives

$$f_i(x) = k \exp\{-\rho_m(x) \sigma_m(\lambda_i) l_m(\lambda_i) - \rho_h(x) \sigma_h(\lambda_i) l_h(\lambda_i)\} p(x) \bar{I}(\lambda_i).$$
(8)

By taking the logarithm of Eq. (8), the image values of human skin color are described as

$$f^{\log}(x) = -\rho_m(x)\sigma_m - \rho_h(x)\sigma_h + p^{\log}(x)1 + e^{\log}.$$
 (9)

We exclude the term k from Eq. (8) because this term is a constant. Where,

$$f^{\log} = [\log(f_R(x)), \log(f_G(x)), \log(f_B(x))]^T$$

$$\sigma_m = [\sigma_m(\lambda_R) l_m(\lambda_R), \sigma_m(\lambda_G) l_m(\lambda_G), \sigma_m(\lambda_B) l_m(\lambda_B)]^T,$$

$$\sigma_h = [\sigma_h(\lambda_R) l_h(\lambda_R), \sigma_h(\lambda_G) l_h(\lambda_G), \sigma_h(\lambda_B) l_h(\lambda_B)]^T,$$

$$1 = [1, 1, 1]^T,$$

$$e^{\log} = [\log(\overline{I}(\lambda_R)), \log(\overline{I}(\lambda_G)), \log(\overline{I}(\lambda_B))]^T,$$

$$p^{\log}(x) = \log(p(x)) + \log(k) .$$

(10)

Here, *1* is the light intensity vector. The skin color vector f^{\log} can thus be represented by the weighted linear combination of the three vectors σ_m , σ_h , and *1* with the bias vector e^{\log} . Equation (9) is presented in Fig. 2. If the



The optical density domain

Fig. 2 Skin color model in the optical density domain

pure pigment color vectors σ_m and σ_h are known, a facial color image can be separated into the component images of melanin, hemoglobin and shading (as shown in Fig. 3) by projecting the skin color vector f^{\log} to the three vectors σ_m , σ_h , and I in the density domain. In previous studies [7–9], the bias vector e^{\log} that represents the illuminant color was ignored during the projection of a skin color.

3 Proposed approach

3.1 Extracting the bias vector

In Sect. 2.2, the skin color analysis [7–9] ignored the bias vector when obtaining the components of melanin, hemoglobin and shading from a facial image. To estimate the illuminant color employing such skin color analysis, we propose a method of calculating the bias vector from a facial image. This section describes our method of

Fig. 3 Pigmentation separation employing skin color analysis

extracting the bias vector and correcting the color of an image containing a person as the main subject using the extracted bias vector.

According to the analysis of human skin color, the individual variation in facial skin color corresponds to the difference in the densities of melanin and hemoglobin; so that the skin color can be plotted on a skin color plane composed of pure pigment color vectors of melanin and hemoglobin in the density domain. Figure 4 shows the distribution of skin color on the skin color plane. There are differences in the density of pigments between the two Japanese subjects in the figure (where the left and right graphs, respectively, show the pigment distributions for subjects 1 and 5 in Table 1). It is seen that subject 1 has a larger variance than subject 5. However, as shown by the red and orange lines in Fig. 4, the minimum densities of melanin and hemoglobin are similar for the two subjects. Therefore, in this paper, we assume that the minimum densities of the two pigments are constant within the same race under the same light source, and we use these minimum values for the bias vector expressed in Eq. (11). Figure 5 shows the steps of our algorithm used to extract the bias vector from the skin color vector and pigment color vector. On the bottom left graph, the pink dashed lines are lines running parallel to the pigment's color vector based on the minimum vector of melanin and hemoglobin. The shading component represents only the brightness level of an image, and we thus exclude the shading density in calculating the bias vector. The bias vector is thus expressed as

$$-e^{\log} = \min\{\rho_m(x)\}\,\sigma_m + \min\{\rho_h(x)\}\,\sigma_h.$$
(11)

The bias vectors extracted using our method from seven Japanese facial images taken under the same light source and for the same camera parameters have a few individual



(a) Input image



(b) Melanin image



(c)Hemoglobin image



(d) Shading image



The optical density domain

Fig. 5 Processing flow of the proposed method of extracting the bias vector

variations. In fact, the variance of the bias vectors in the CIELAB space is 0.58 whereas the variance of the facial color in the CIELAB space is 2.05. This result shows that the bias vector is more uniform among Japanese people than the mean facial color.

3.2 AWB algorithm using the bias vector

This section introduces the steps of our AWB algorithm using the bias vector. The processing flow is shown in Fig. 6. First, standard facial images are taken under a



Fig. 6 Processing flow of the proposed method of AWB for the facial image using the bias vector

standard light source, and using our method introduced in Sect. 3.1, we calculate the bias vectors from patch images extracted from standard facial images. By calculating the mean vector among these bias vectors, we can decide the standard bias vector e_s that gives the standard illuminant color. Next, from the input facial image taken under an arbitrary light source, we extract the observed bias vector e_o that gives the arbitrary illuminant color. We then color correct the input facial image as if the image were taken under the standard light source by exchanging the bias



The optical density domain

Fig. 7 Color correcting by exchanging the bias vector of the facial image $% \left({{{\mathbf{F}}_{{\mathbf{F}}}}_{{\mathbf{F}}}} \right)$

from e_s to e_o in the density domain. Figure 7 shows the color correction method that uses the bias vector and replaces the observed bias with the standard bias.

4 Experiments

4.1 Experiment method

We used a digital camera (Nikon D5100, 2464×1632 pixels) and five light sources, namely a xenon light source (6000 K), halogen light source (3100 K), and light-emitting diode (LED) light source (5600 K) with white, orange, and green filters. We obtained each facial image in a darkroom with a single light source, where the distance between the camera and subject was 90 cm. Each subject was photographed with a Macbeth color checker positioned close to his/her face. The camera parameters (F number, shutter speed, and ISO speed) were the same for all subjects under the same light source.

For the calculation of the standard bias, four Japanese people were photographed under the standard light source. In our experiments, the xenon light source was chosen as the standard light source. Figure 8(a) shows a facial image of four images trimmed to about 1200×800 pixels. A patch image as shown in Fig. 8b, with 300×300 pixels is





(b) Skin color image

Fig. 8 Example images of a a facial image and b patch image extracted from the facial image, which are used in our experiment

extracted from the areas of the facial image as cheek, forehead, chin and so on. The extracted areas are shown as red squares in Fig. 8a. When employing this method, the black area of the skin color image is masked. The skin color vector of the facial pixel is projected to the pigment color vector, and the minimum vector of pigment density is calculated. We assumed that the pigment color vector is similar for subjects of the same race, and we thus used (0.3517, 0.4310, 0.8004) as the melanin vector and



(a) Halogen



(c) LED with orange filter





(b) LED with white filter



(d) LED with green filter

Fig. 9 Facial images taken under four light sources: a halogen light source and LED light source with white, orange and green filters



(a) Halogen



(**b**) LED with white filter



(c) LED with orange filter



(d) LED with green filter





(a) Halogen



(b) LED with white filter



(c) LED with orange filter



(d) LED with green filter

(0.2981, 0.7631, 0.5034) as the hemoglobin vector for the four subjects. These vectors are adequate for separating the melanin and hemoglobin components in the case of Japanese subjects. By adding the minimum vector of melanin and hemoglobin, bias vectors were extracted from the skin color images of the four subjects and the standard bias e_s was calculated as the mean of biases for the four subjects.

The four subjects were also photographed under the other four light sources. Figure 9 shows examples of facial images taken under the halogen light source and under the LED light source with white, orange, and green filters. Patch images were extracted as shown in Fig. 10, and the observed bias e_o of the facial image was extracted from each patch image. Then, by exchanging the observed bias vector with the standard bias vector, the facial color image was corrected as if the image was taken under the xenon light source.

4.2 Evaluation results

This section evaluates the effectiveness of our proposed method. We photographed subjects next to a Macbeth color checker. We cut off the area of the light skin from the color checker in the taken images. We calculated the mean pixel value for the area of light skin and then evaluated the differences in the mean pixel value between the xenon image and color-corrected images. The image taken under the xenon light source was the ground truth of the correction of each image using the AWB method.

In this paper, we evaluated the color shift from the ground truth to the corrected image using Δa^* and Δb^* in the CIELAB space. ΔL is not directly related to the illuminant color, and we thus excluded ΔL when comparing the color differences. We express ΔE as the Euclidean distance between the ground truth color and the corrected color in the two chromatic dimensions a^* and b^* .

Figure 11 shows the facial images corrected employing the proposed method. All images are clearly similar to the image taken under the xenon light source shown in Fig. 11a. Figure 12 shows the checker images extracted from the input images (Fig. 11) and the corrected images (Fig. 12). Figure 13 shows the color shift between the checker images taken under the xenon light source and those taken under the other four light sources. The red and gray bins represent the hue differences obtained using the proposed method and conventional method, respectively. The conventional method uses the mean RGB value of the Japanese facial color, considering that the Japanese face is



Fig. 12 Light-skin patch of the Macbeth color checker extracted from the input and the color-corrected images taken under five light sources: the xenon light source, halogen light source, and LED light source with *white*, *orange* and *green* filters (color figure online)



Fig. 13 Hue differences of the Macbeth color checker

the key object introduced in Sect. 2.1. The processing flow of this method is as follows. In advance, the mean RGB value of the Japanese facial color is calculated from facial images of four subjects taken under the xenon light source. In the color correction step, the input facial image is color corrected by matching the facial RGB values in the image taken under the xenon light source with facial RGB values in the input image. The performance of this method is thus affected by differences in individual skin color.

As shown in Fig. 13, considering all light sources, the proposed method realizes better color correction than the conventional method. The result shows that the performance of the proposed auto white balance algorithm is not affected by differences in individual skin color.

5 Conclusion

This paper presented an AWB algorithm using a pigmentation separation method. The pigmentation separation method separates the human skin color into the components of melanin, hemoglobin and shading. We use the skin color property that the vector of minimum melanin and hemoglobin densities under the same light source is uniform among individuals of the same race. The bias vector calculated as the minimum vector of pigment components thus represents the color of the light source.

We performed experiments for the color correction of facial images and the evaluation of the effectiveness of the proposed method. We photographed four Japanese people under five light sources, and corrected the facial images so that they matched images taken under a xenon light source. The results of the experiments demonstrate the superiority of the proposed method over conventional color correction.

As shown in this paper, a white balance function can be realized when the faces of Japanese people are photographed. The subjects in our experiments were Japanese people. We expect that the proposed method can be adapted to Southeast Asians having facial pigmentation similar to that of Japanese people. However, Caucasians and African Americans were not accounted for in the present work. We will, therefore, endeavor to develop the proposed color correction method for people of various races as future work.

References

- 1. Foster, D.H.: Color constancy. Vision. Res. 51, 674-700 (2011)
- Gijsenij, A., Gevers, T., Van De Weijer, J.: Computational color constancy: survey and experiments. Image Process IEEE Trans 20, 2475–2489 (2011)
- Buchsbaum, G.: A spatial processor model for object colour perception. J. Franklin Inst. 310, 1–26 (1980)
- Land, E.H.: The retinex theory of color vision. Sci Am 237(6), 108–128 (1977)
- Lee, H.-C.: Method for computing the scene-illuminant chromaticity from specular highlights. JOSA A 3, 1694–1699 (1986)
- Tsukada M., Funayama C., Tajima J.: Automatic color preference correction for color reproduction. In: Photonics West 2001-Electronic Imaging, pp. 216–223 (2000)
- Tsumura, N., Haneishi, H., Miyake, Y.: Independent-component analysis of skin color image. JOSA A 16, 2169–2176 (1999)
- Tsumura N., Ojima N., Nakaguchi T., Miyake Y. (eds.): Skin Color Separation and Synthesis for E-Cosmetics. In: Signal Processing Techniques for Knowledge Extraction and Information Fusion, pp. 201–220. Springer, Berlin (2008)
- Tsumura, N., Ojima, N., Sato, K., Shiraishi, M., Shimizu, H., Nabeshima, H., et al.: Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. ACM Trans Gr (TOG) 22, 770–779 (2003)
- Finlayson, G.D., Drew, M.S., Funt, B.V.: Color constancy: generalized diagonal transforms suffice. JOSA A 11, 3011–3019 (1994)
- Funt, B.V., Lewis, B.C.: Diagonal versus affine transformations for color correction. JOSA A 17, 2108–2112 (2000)
- Von Kries J.: Sources of color science, Chromatic adaptation, pp. 109–119 (1970)

- West, G., Brill, M.H.: Necessary and sufficient conditions for von Kries chromatic adaptation to give color constancy. J. Math. Biol. 15, 249–258 (1982)
- Klinker, G.J., Shafer, S.A., Kanade, T.: A physical approach to color image understanding. Int. J. Comput. Vision 4, 7–38 (1990)
- Shafer, S.A.: Using color to separate reflection components. Color Res Appl 10, 210–218 (1985)
- 16. Igarashi T., Nishino K., Nayar S. K.: The appearance of human skin: a survey, In: Foundations and Trends[®] in Computer Graphics and Vision, vol. 3, pp. 1–95 (2007)
- Van Gemert, M., Jacques, S.L., Sterenborg, H., Star, W.: Skin optics. Biomed. Eng. IEEE Trans. 36, 1146–1154 (1989)
- Hiraoka, M., Firbank, M., Essenpreis, M., Cope, M., Arridge, S., Van Der Zee, P., et al.: A Monte Carlo investigation of optical pathlength in inhomogeneous tissue and its application to nearinfrared spectroscopy. Phys. Med. Biol. 38, 1859 (1993)
- Anderson, R.R., Parrish, J.A.: The optics of human skin. J. Investig Dermatol 77, 13–19 (1981)
- Drew M. S., Chen C., Hordley S. D., Finlayson G. D.: Sensor transforms for invariant image enhancement. In: Color and Imaging Conference, pp. 325–330 (2002)