https://doi.org/10.2352/ISSN.2470-1173.2017.17.COIMG-424

Performance of the 14 skin-colored patches in accurately estimating human skin color

Hayan Choi, Kyungah Choi, Hyeon-Jeong Suk; Department of Industrial Design, Korea Advanced Institute of Science and Technology; Daejeon, Korea

Abstract

This research examined the performance of skin colored-patches for accurately estimating human skin color. More than 300 facial images of Korean females were taken with a digital single-lens reflex camera (Canon 550D) while each was holding the X-Rite Digital ColorChecker® semi-gloss target. The color checker consisted of 140 color patches, including the 14 skin-colored ones. As the ground truth, the CIE 1976 $L^*a^*b^*$ values of seven spots in each face were measured with a spectrophotometer. For an examination, three sets of calibration targets were compared, and each set consisted of the whole 140 patches, 24 standard color patches and 14 skin-colored patches. Consequently, three sets of estimated skin colors were obtained, and the errors from the ground truth were calculated through the square root of the sum of squared differences (ΔE). The results show that the error of color correction using the 14 skin-colored patches was significantly smaller (average $\Delta E = 8.58$, $SD = 3.89$) than errors of correction using the other two sets of color patches. The study provides evidence that the skin-colored patches support more accurate estimations of skin colors. It is expected that the skin-colored patches will perform as a new standard calibration target for skin-related image calibration.

Introduction

Research on skin appearance has received great attention from a variety of fields such as cosmetology, medicine, computer vision, and computer graphics [1]. In particular, the extraction of accurate skin color is in high demand for the purpose of shade matching or skin diagnosis in the fields of cosmetology and dermatology. There are two ways to extract skin color: the first way is to measure the skin with devices, and the other way is to photograph the skin. The devices for measuring skin color provide more accurate results; however, they are expensive and hard for end users to use [2]. The importance of end users is increasing due to the development of technology as well as the need for such applications in everyday life.

In order to improve the convenience of these measurements, some studies tried to use digital images for skin measurement without using any additional devices [3, 4]. However, the applications are meaningful only when the skin color is accurately captured. Unfortunately, digital images cannot capture the accurate color of an object due to their imperfect illuminant compensation as well as camera characteristics. Therefore, color correction and the standardization of the images are regarded as mandatory procedures in order to extract accurate skin color [2, 5, 6]. These procedures facilitated the color constancy theory in order to correct the color in the images to make it seem as if the images were taken under the standard illuminant.

To measure the performance of the color constancy algorithm, the color correction using a color checker is usually considered the ground truth. In other words, the performance of a color constancy algorithm is measured by comparing the result of using the algorithm with the result of using the color checker. However, Harville, et al. [2] showed that the primary colors perform worse than skin-colored patches in the skin color range. Because people's perceptions of the skin color are more sensitive than that of other colors, a higher level of performance is required for the skin color range. Thus, Harville et al. [2] proposed using additional calibration targets that focus on a representative range of human skin colors. However, few have tried to verify whether the skin-colored patches actually capture the skin colors more accurately than the 24 patches of the standard calibration targets.

Hence, this study investigated the performance of color correction using different color patches in color checkers. Three sets of color patches were compared, and each set consisted of the extended 140 patches, 24 standard patches and 14 skin-colored patches. The extracted skin colors from each calibrated image were compared with the measured skin color values from more than 300 females.

Related Works

The assessment of skin color involves using colorimetric devices, imaging sensors and digital images. These devices and sensors make it possible to get more accurate and precise values. However, these devices and sensors are limited due to the lack of accessibility to them, the relatively small interrogation area and the need for contact [3]. End users now need the applications of skin measurements in everyday life, but the devices involved are expensive and hard for end users to use [2].

On the other hand, digital images are more accessible and cost less. Since the image captured by a camera is influenced by the illuminant and camera specifications, color correction should be used to eliminate these influences.

Color correction can be divided into two types. The first type aims to correct an image based on assumptions [7-9], and the other type corrects an image based on a calibration target, which uses the known color values. As an example of color correction using assumptions, Bianco and Schettini [7] assumed that the skin color is narrowly distributed in the color space. They detected the face and transformed the facial skin color into the predefined skin color range so that the other parts of the image will be corrected together. However, this method only transforms the skin pixels into the predefined skin color range, so it is not appropriate for capturing the accurate color values of the skin. Do, et al. [8] assumed that the sclera of the eye is white under standard illuminant, which means that the illuminant by the sclera region can be estimated. Störring, et al. [9] assumed that the highlights of the skin can be used to estimate the illuminant, which is a similar concept to that in the study of Do, et al. [8].

On the other hand, several studies used various calibration targets to extract skin color [2, 10]. Color targets are used to measure the physical properties such as dynamic range, the noise levels, the optical resolution and the skin color calibration. Some examples of color checkers can be seen in Figure 1. Among the calibration targets, the standard calibration target, which is composed of 24 color patches, is regarded as the most precise color correction, so it is considered the ground truth. However, human vision is more sensitive to skin colors than any other colored objects [11]. Harville, et al. [2] found that color correction using the primary colored patches may skew the pixels in the skin range. Since a little difference in skin color may induce a large difference in the skin related applications, another standard target for skin colors is needed. For this purpose, we aimed to determine the performance of skin-colored patches in order to use them as the standard target for skin colors.

To confirm the performance of color correction using skin-colored patches, we selected the Digital ColorChecker® semi-gloss target produced by X-Rite Inc., which includes 14 skin-colored patches. However, there has been little evidence of whether the 14 skin-colored patches support the color constancy in order to capture the color of human skin more accurately. This research attempted to examine the performance of the skin-colored patches compared with the existing color constancy algorithms. For a comparison, the actual skin colors were measured using a spectrophotometer, and the CIE 1976 $L^*a^*b^*$ values were considered as the ground truth.

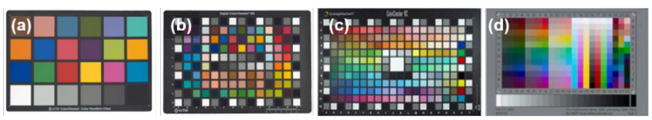

Figure 1. Examples of standard calibration targets: (a) X-Rite ColorChecker Classic; (b) X-Rite Digital ColorChecker SG; (c) Macbeth ColorChecker DC and (d) Agfa IT8.

Objectives

This research intended to examine the performance of the skin-colored patches in estimating human skin color more accurately than the conventional whole sets of color checkers. In the skin color range, human vision is highly sensitive, and the image calibration based on the conventional sets are questionable. However, accurate skin color values are needed for the skin-color-based applications. If the skin-colored patches have a higher accuracy, they will perform as a new standard calibration target in the skin color range.

For this purpose, the color value of skin was measured with the spectrophotometer, and facial images were collected while the females in the images held a color calibration target, which consisted of 140 color patches. Based on this, three types of color calibrations were performed: a calibration using all 140 patches, a calibration using 24 patches and a calibration using 14 skin-colored patches. The performance of the calibrations using each color patch group was calculated by the difference between the CIE 1976 $L^*a^*b^*$ values of the measured skin color and the calibrated result.

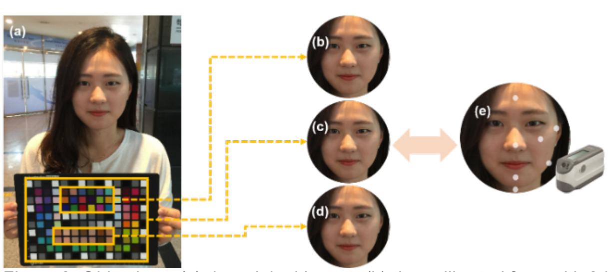

Figure 2. Objectives: (a) the original image; (b) the calibrated face with 24 patches; (c) the calibrated face with 140 patches; (d) the calibrated face with skin-colored patches and (e) the skin color dataset that was measured using a spectrophotometer.

Method

To identify the effects of the color patches on image calibration and skin color estimation, we collected facial images of females holding the color checker, and we measured the colorimetric values of facial skin. Based on the dataset, we calibrated each face using the color patches in the color checker, including the skin-colored patches, the conventional 24 color patches and the whole 140 patches. After the calibration, the skin colors from the facial images were extracted and averaged into a single color. The calculated skin colors of each color patch group were compared with the measured colorimetric values.

Dataset Collection

The 348 Korean females were recruited, and their ages ranged from 17 to 56 years old (mean 21.56 years, standard deviation 5.56 years). Each female was a paid volunteer. A facial photo of each subject was taken with a digital single-lens reflex camera (Canon 550D), and the illuminant condition, including the illuminance and correlated color temperature (CCT), of each participant was also measured by a chromameter. The average illuminance was 20,016 lux (standard deviation = 22,972 lux), and the average CCT was 5,184 K (standard deviation = 1,890 K). All of the images were recorded in the RAW mode to avoid automatic calibration. When the photo was taken, the females were holding the Digital ColorChecker® semi-gloss target (X-Rite Inc.), which includes 140 color patches. That calibration target is extended from the 24 standard colored patches by including natural objects, a 17-step gray scale and the 14 skin-colored patches.

On the other hand, the ground truth for comparing the performance of each calibration was measured by the CIE 1976 $L^*a^*b^*$ values of the faces by using a spectrophotometer (Konica Minolta, CM 2600d). The spots for measurement were decided as suggested in the related previous studies [12-14]. The spots were the forehead, tip of the nose, chin, neck, cheek, cheekbone and jaw. These seven colors were averaged into a representative color of each face and considered the ground truth. The average values of the measured skin colors were [61.06, 9.65, 17.36], and the standard deviation of that was [2.18, 1.21, 1.67], $L^*a^*b^*$ respectively.

Image Calibration

The calibration aimed to minimize the errors between each patch and the target value of the patches by using the root mean square. The three types of calibration were performed using (1) the whole 140 color patches; (2) the 24 standard color patches and (3) the 14 skin-colored patches. Accordingly, each calculation resulted in an image so that three images for each input image were generated.

After the image calibration, the skin-colored region was extracted through the decision boundary that is defined $0 \leq H \leq 50$ for hue, $0.2 \leq S \leq 0.68$ for saturation and $0.35 \leq V \leq 1$ for value [15]. This skin boundary was commonly used in the existing research for comparing the results of algorithms such as those regarding color constancy [8].

We then averaged the colors of the skin region for each of the three calibrations. To avoid ambiguity, we considered the whole pixel values from the faces for the comparison. The skin color values were transformed into the CIE 1976 $L^*a^*b^*$ values, which is the same color space of the skin measurement. The measured values of each participant were averaged in order to derive a representative skin color for each participant as the ground truth. Accordingly, we calculated the differences, ΔE, between the calibrated colors and the ground truth for each subject.

Results and Analysis

The performance of each algorithm was verified as the difference between the ground truth skin color and the estimated skin color from each calibration. The difference was calculated as the ΔE, which is commonly used for calculating the difference between colors. Among the three calibrations, the calibration using the 14 skin-colored patches was the most accurate, having the smallest difference between the measured skin color and the calibrated skin color for each subject. The average ΔE was 8.58, and the standard deviation was 3.89. The calibration using the whole 140 patches performed better than the calibration using the 24 standard color patches. The average ΔE of the whole 140 patches was 8.88 (standard deviation = 3.92), and the average ΔE of the 24 standard color patches was 9.22 (standard deviation = 4.07). The experiment's results are shown in Figure 3.

The mean difference was analyzed by a repeated measure one-way ANOVA. The result yielded statistical significance at an alpha level of 0.05 [$F(2, 1804) = 34.02$, $P < 0.05$]. From the post hoc test, the pairwise mean differences were also statistically significant, indicating that the type of calibration target influences the color correction. In particular, the mean differences provide evidence that it is more accurate to perform a color calibration only using the 14 skin-colored patches rather than using the whole set of calibration targets.

Since the calibration is computed to minimize the errors between the captured color patches and the target color patches, the computational load for calibration depends on the number of color patches used. The computational load is lower and the speed is faster when fewer patches are considered. Several applications put a bigger emphasis on the computational load and speed, such as mobile applications and real–time applications. This implies that, even though the performance of the calibration using the whole 140 patches was better than the 24 patches, those applications might use the 24 patches due to the faster speed. Regardless, this study verified that the calibration using the 14 skin-colored patches has the lowest computational load as well as the highest performance in the skin range.

Previously, in the study of Harville, et al. [2], the skin-colored patches performed best for distinguishing between 16 ethnic groups based on skin color. In line with their assertion, we can conclude that the skin-colored patches support a more accurate and distinguishable calibration. This result implies the potential of introducing a specialized calibration target. In the case of calibrations which need high accuracy such as skin estimation, it would be more accurate and efficient to calibrate using only the related colors rather than using the whole set of colors.

Result: The mean ΔE between each calibration

Calibration Type	Mean ΔE (SD)
Whole 140 patches	8.88 (3.92)
Standard 24 patches	9.22 (4.07)
Skin-colored 14 patches	8.58 (3.06)

Discussion

From the results, the skin-colored patches were verified as having the most accurate performance of skin color estimation among the three calibrations. However, the color correction focusing on skin color may distort the other range of colors. We have not yet tested whether the non-skin areas were also properly calibrated. As the human eyes are more highly sensitive to changes of skin tone than other colors [11], the calibration of skin color might affect viewers' satisfaction more than calibration of other colors. In addition, this implies the possibility of using a specialized calibration target, which is composed of strategically selected color patches. This specialized calibration target would not be limited to the skin color range but would also be extended to any color range that focuses on accurate color estimation.

This examination was limited to Korean females whose skin color was concentrated in a narrow region in a color space. Since the 14 skin-colored patches on the Digital ColorChecker® semi-gloss target include the range of various ethnic groups, these results are expected to be valid for the different ethnic groups. Further studies with different ethnic groups will be conducted to validate the empirical findings of this study.

Nevertheless, further investigation is necessary to successfully utilize the skin-colored patches for practical applications. For example, estimation of skin color is a primary step in cosmetic applications such as matching foundations. As presented by Harville, et al. [2], a new portable calibration target can be proposed by using skin-colored patches.

Finally, the calibration based on the skin-colored patches is most accurate, but errors still remain. This result may be due to the illuminant and the specifications of the devices as well as the distance between an object and camera, the textures and the three–dimensional structure of the object. From the perspective of usability, the color checker has a great advantage of convenience. If possible, the colorimetric values will provide more accurate values to be used as the ground truth.

Conclusion

Calibration targets are widely used in imaging applications as ground truths because of their great advantage of usability. This study verified that the skin-colored patches provide more accurate skin color estimations than the 24 colored patches and 140 colored patches through the skin color dataset of 300 females. Although previous research proposed additional color calibration targets focusing on a representative range of human skins, few have tried to prove whether the patches actually capture the skin colors more accurately. This study provides evidence that the skin-colored patches support more accurate estimations of skin colors measured with a spectrophotometer. Since people perceive skin colors more sensitively than the colors of other things, the approach toward reducing the color difference is much more meaningful.

References

[1] T. Igarashi, K. Nishino, and S. K. Nayar, "The appearance of human skin: A survey," Foundations and Trends® in Computer Graphics and Vision, vol. 3, pp. 1-95, 2007.

[2] M. Harville, H. Baker, N. Bhatti, and S. Susstrunk, "Consistent image-based measurement and classification of skin color," in IEEE International Conference on Image Processing 2005, 2005, pp. II-374-7.

[3] C. S. Kim, M. Ki Kim, B. Jung, B. Choi, W. Verkruysse, M. Y. Jeong, et al., "Determination of an optimized conversion matrix for device independent skin color image analysis," Lasers in surgery and medicine, vol. 37, pp. 138-143, 2005.

[4] H. Wannous, Y. Lucas, S. Treuillet, A. Mansouri, and Y. Voisin, "Improving color correction across camera and illumination changes by contextual sample selection," Journal of Electronic Imaging, vol. 21, pp. 023015-1-023015-14, 2012.

[5] C.-C. Chang, Y.-C. Chuang, C.-T. Wu, B. I. Choi, K. Lee, S. Woo, et al., "A low-cost mobile device for skin tone measurement using filter array spectrum sensor," in IEEE SENSORS 2014 Proceedings, 2014, pp. 499-502.

[6] J. Van De Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," IEEE Transactions on image processing, vol. 16, pp. 2207-2214, 2007.

[7] S. Bianco and R. Schettini, "Color constancy using faces," in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, 2012, pp. 65-72.

[8] H.-C. Do, J.-Y. You, and S.-I. Chien, "Skin color detection through estimation and conversion of illuminant color using sclera region of eye under varying illumination," in 18th International Conference on Pattern Recognition (ICPR'06), 2006, pp. 327-330.

[9] M. Störring, E. Granum, and H. J. Andersen, "Estimation of the illuminant colour using highlights from human skin," in First International Conference on Color in Graphics and Image Processing, 2000, pp. 45-50.

[10] S. Bianco and C. Cusano, "Color target localization under varying illumination conditions," in International Workshop on Computational Color Imaging, 2011, pp. 245-255.

[11] P. Bodrogi and T. Khan, Illumination, color and imaging: evaluation and optimization of visual displays: John Wiley & Sons, 2012.

[12] J. Tajima, M. Tsukada, Y. Miyake, H. Haneishi, N. Tsumura, M. Nakajima, et al., "Development and standardization of a spectral characteristics database for evaluating color reproduction in image input devices," in SYBEN-Broadband European Networks and Electronic Image Capture and Publishing, 1998, pp. 42-50.

[13] M. Wang, K. Xiao, S. Wuerger, V. Cheung, and M. R. Luo, "Measuring human skin colour," in Color and Imaging Conference, 2015, pp. 230-234.

[14] K. Xiao, N. Liao, F. Zardawi, H. Liu, R. Van Noort, Z. Yang, et al., "Investigation of Chinese skin colour and appearance for skin colour reproduction," Chinese Optics Letters, vol. 10, p. 083301, 2012.

[15] Y. Wang and B. Yuan, "A novel approach for human face detection from color images under complex background," Pattern Recognition, vol. 34, pp. 1983-1992, 2001.

Author Biography

Hayan Choi received her B.S. in industrial design(ID) and minored computer science. received M.S. in ID from the Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea. Currently, she is a Ph.D. candidate of KAIST in a color laboratory (color.kaist.ac.kr). Her research interests include computational imaging and skin color.

Kyungah Choi received her B.S. and M.S. in industrial design from the Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea. Currently, she is a Ph.D. candidate of KAIST in a color laboratory (color.kaist.ac.kr). Her research interests include psychophysics, color perception, and lighting technology.

Hyeon-Jeong Suk received her BS and MS degrees in industrial design from KAIST, Republic of Korea, and her PhD degree in psychology from the University of Mannheim, Germany. Currently, she is an associate professor of KAIST leading a color laboratory. Her research interests include color psychology and emotional design.