https://doi.org/10.2352/CIC.2023.31.1.9

Improve image white balance by facial skin color

Liqing Wang[1], Yuechen Zhu[1], Xiaoxuan Liu[1], Ming Ronnier Luo[1,2,*]
[1]State Key Laboratory of Extreme Photonics and Instrumentation,Zhejiang University,CHINA.
[2]School of Design, Leeds University, Leeds, UK.

Abstract

In area of white balance, the process of large colored background images seems to be a problem. Regarding this issue, a white balance algorithm based on facial skin color was proposed. A neural network-based object detection algorithm and an adaptive threshold segmentation algorithm were combined to achieve the accurate segmentation of skin color pixels. Then a 3-dimension color gamut mapping method in CIELAB color space was used to do the illumination estimation. Last, CAT16 model was applied to rendering the images to standard lighting condition. Besides, an ill white balanced images dataset taken against large colored backgrounds were prepared to test the present algorithm and others' performance. The results show the proposed algorithm performs better on the dataset.

1. Introduction and related work

Color constancy refers to the ability of our visual system to perceive the colors of objects consistently, regardless of changes in lighting conditions. It allows us to perceive objects with relatively stable colors despite variations in the illumination. To digital imaging devices, color constancy is achieved by an approach called white balance. The typical white balance methods generally consist of two steps, first step is to estimate the illumination, and the second step is to adjust the color balance accordingly to produce a visually pleasing and color-accurate image. About white balance, a lot of previous algorithms are proposed, such as gray-world, white patch, max RGB, edged based white balance[1], gamut mapping[2], machine learning based white balance, and etc[3-6]. While numerous white balance algorithms have been proposed, there is no single algorithm that performs optimally in all situations. Each algorithm has its own set of conditions in which it performs well.

Typically, traditional auto white balance algorithms operate by detecting the dominant color in the scene that should be perceived as neutral (usually white or gray) and then adjusting the color channels of the image to remove the color cast and achieve a more accurate representation of colors. However, when there is not enough neural pixels in an image, some existing method may not perform well. To solve this problem, in this study, a new white balance algorithm based on memory skin color was proposed.

Memory colors refer to the perceived or remembered colors associated with specific objects or elements in our environment. These colors are deeply ingrained in our visual memory and are often strongly associated with certain objects, such as the blue sky, green grass, red apples, or the skin tone of a person. In the context of white balance, memory colors can be used as reference points or targets for color correction. By using memory colors as a basis for white balance, the algorithm can prioritize the accurate representation of these familiar and memorable colors. This can help in producing more visually pleasing and natural-looking images, as our perception is often influenced by our memory and expectations of how certain objects should appear in terms of color[7].

Although there are currently algorithms that perform white balance using skin tones[8]. Such algorithm that used hue angle for skin segmentation in the HSV color space may have limited universality. When used across different devices, or the lighting condition changes a lot, the hue angle may vary, and maybe it's very hard to segment the skin color pixel.

Compared with the existing white balance method using faces, the method proposed in this research has made improvements mainly in three aspects: skin color segmentation, light source estimation, and image rendering.

First, an computational fast object detection algorithm yolov5[9] was applied to detect human face in the images, then an adaptive threshold skin color ellipsoid in CIELAB color space instead of HSV color space was used to get accurate skin color pixels segmentation. Second, a three-dimension gamut mapping in CIELAB color space was applied to estimate the illumination. And the third, CAT16 model[10] was used to render the image to target illumination environment.

Besides, a dataset including images captured under five different color temperatures, featuring three different ethnicities, five colored backgrounds were prepared to test our algorithm and others to evaluate their performance.

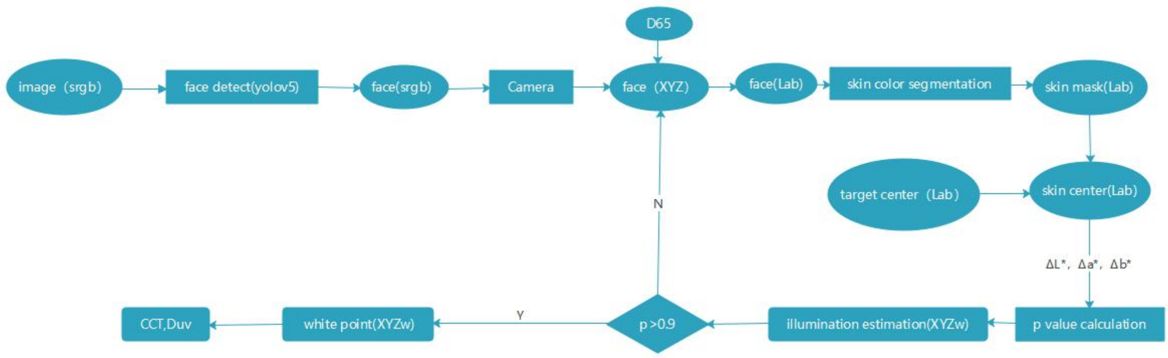


Figure 1 Workflow of the proposed illumination estimation method; sRGB image input, after object recognition process, divide face region from the full image. Using an adaptive threshold skin color ellipsoid to achieve pixel-level skin color segmentation. Then a 3-dimension gamut mapping method in CIELAB color space was applied to estimate the illumination. Check the result, and output the CCT and Duv of the illumination.

2. Memory color white balance

Figure 1 presents the proposed illumination estimation algorithm. The sRGB image input, after a series of processing steps, estimation of the light source is obtained.

2.1 Facial skin pixel segmentation

The first step of the process is to get the actual skin color pixels. A neural network-based object detection algorithm with an adaptive threshold segmentation algorithm were combined to achieve the accurate segmentation of skin color pixels.

The sRGB image input, then a face detect algorithm yolov5[9] was applied to find human face in the image. Afterwards, the facial region is segmented from the entire image to be used for subsequent skin color pixel segmentation. The reason for this step is because the segmented facial region typically contains a significant proportion of skin color pixels, which helps to eliminate interference from background pixels that may be close to skin color.

After facial detection, the second step is segmentation of skin color pixels. To achieve this, a skin color segmentation method based on adaptive threshold was proposed. The first step is to convert the image from RGB to XYZ color space by the standard sRGB to XYZ matrix. Then, using D65 as the reference white point, the images were converted from XYZ to CIELAB color space. From previous research, it is known that skin colors exhibit an ellipsoidal distribution in the CIELAB color space[11]. Moreover, under the D65 illumination condition, the parameters of the ellipsoid corresponding to a specific ethnicity in the CIELAB color space are well-defined[12].

$$P = \frac{1}{1+e^{\Delta E-\alpha}} \tag{1}$$

$$\Delta E = \sqrt{\begin{aligned} &k_1 * (L-L_s)^2 + k_2 * (a-a_s)^2 + \\ &k_3 * (b-b_s)^2 + k_4 * (a-a_s) * (b-b_s) \end{aligned}} \tag{2}$$

Equation 1 and 2 are skin color ellipsoid equation in PMCC (prefer memory color chart) [13-15].

k_1-k_4 , α , are ellipsoid parameters. L_s , a_s , b_s are prefer memory skin color ellipsoid center. When P=0.5, the equation represents a 50% acceptance skin ellipsoid in CIELAB color space.

So, the above 50% acceptance skin color ellipsoid from PMCC was used as a reference and each pixel in the image was checked to see whether it falls within the skin color ellipsoid. If a pixel falls within the ellipsoid, it is considered as skin color; otherwise, it is not. After one round of detection, the accumulated skin color pixels are used to compute a mean Lab value, which becomes the new center of the skin color ellipsoid, while keeping the other parameters unchanged. Then another round of skin color detection is performed, and the process is repeated until the Lab center value of the skin color ellipsoid no longer changes, which means a significant majority of the skin color pixels in the image are within the current skin color ellipsoid. At this point, the new skin color ellipsoid can represent the skin color region in the corresponding image, achieving the segmentation of skin color pixels.

The position of the ellipsoid found using this method will vary as the image changes. It allows for finding skin color pixels in different lighting conditions and for different ethnicities, aiming to capture as many skin color pixels as possible in the image. So it’s called an adaptive threshold skin color ellipsoid.

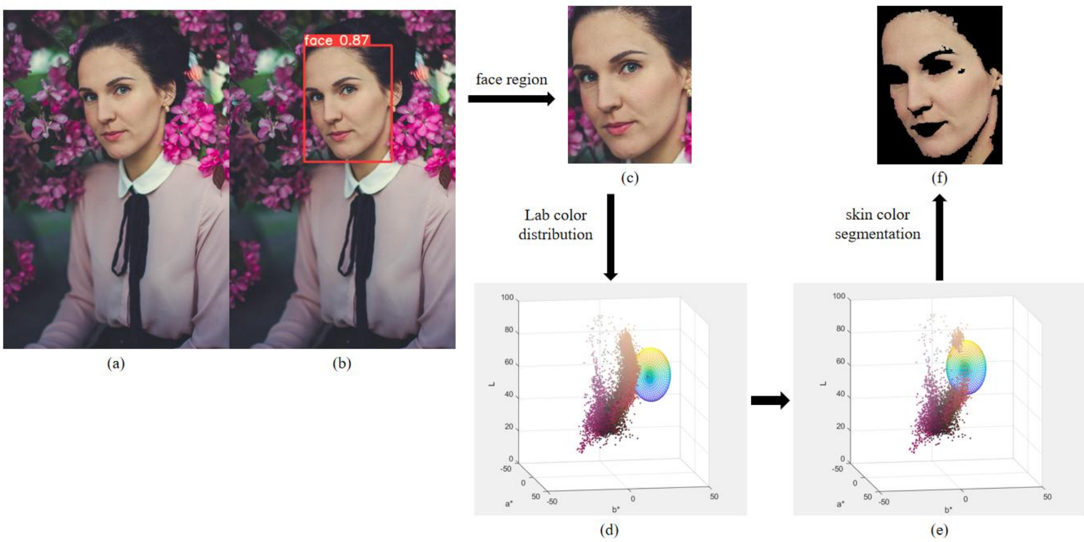


Figure 2. An example of skin color pixel segmentation process ; (a) the example image. (b) the image processed after face detection, the location of face region was determined. (c) the face region was divided from the full image, eliminate interference from background pixels that may be close to skin color. (d) The color distribution of the face region image pixels in the CIELAB color space, and 50% acceptance skin color ellipsoid mentioned above. (e) After adjustment, an appropriate center for the skin color ellipsoid is found, ensuring that the majority of the skin color pixels in the image fall within the new ellipsoid. Using this new ellipse, we’ll be able to create a skin color mask. (f) Apply morphological operations, such as erosion and dilation, to refine the skin color mask and remove noise or small artifacts. Eventually get the accurate skin color mask.

2.2 Illumination estimation

CIELAB is widely used in color-related applications, such as color calibration, color matching, and color difference calculations. It provides a perceptually uniform color space, meaning that equal distances in the space represent approximately equal perceived color differences to the human eye. This makes color comparison and adjustments more intuitive and accurate. Besides, CIELAB color space decomposes colors into two components of lightness (L*) and chroma (a* and b*). This allows for more precise control and adjustment of colors without being influenced by the lightness component. This makes it easier to maintain color consistency across different devices. What's more, CIELAB color space is a device independent color representation, unaffected by specific display or output devices. And it's computational efficient and cross-cultural.

Based on all these advantages, a convenient illumination estimation method in CIELAB color space was proposed.

As is mentioned above, it's known that for a specific ethnicity, the position and parameters of the skin color ellipsoid in the CIELAB color space is determined[12], and this skin color ellipsoid (50% acceptance ellipsoid in PMCC)[14] can be considered as a standard three-dimensional skin color gamut. Which means when capturing an image containing skin color, regardless of the lighting conditions during the photo capture, as long as the correct reference white point was chosen during the conversion process from XYZ to LAB, the result Lab value of skin color should fall within this standard skin color gamut.

However, in the example of figure 2, in the process of skin color segmentation, most of the skin color pixels obtained in the image were distributed in the CIELAB color space as a new ellipsoid instead of 50% acceptance ellipse in PMCC. The reason why the skin ellipsoid of the image is different from ellipsoid in PMCC is because D65 was chosen as the reference white point during the XYZ to LAB conversion, while the actual lighting conditions during the photo capture were not D65. So, based on these theories, the difference between the skin color ellipsoid of the image and the standard 3-d skin color ellipsoid gamut were used to predict the illumination environment.

The equations are listed below:

$$\begin{aligned} XYZ_{W1} &= \text{Lab2xyz}\,(100+\Delta L, \Delta a, \Delta b, XYZ_W) \\ \Delta L &= L_i - L_s \\ \Delta a &= a_i - a_s \\ \Delta b &= b_i - b_s \end{aligned} \tag{3}$$

Lab2xyz() represents a function that can convert Lab to XYZ value. L_s , a_s , b_s represent the center value of standard skin color ellipsoid. L_i , a_i , b_i represent the center value of skin color ellipsoid in the image. XYZ_W is white point used before, XYZ_{W1} is the new white point.

After converting the image to the LAB color space using D65 as the reference white point, a new white point can be calculated based on the equation above. Then, using this calculated new white point as the reference, another conversion from XYZ to LAB color space was performed, then equation (1) and (2) was used to calculate the p value of this image's average skin color. Repeated these above procedures until the p value of the image's average skin color is big enough to meet our requirement. In this research, 0.9 was set as a standard.

Once the p value meets the standard, the actual skin color ellipsoid will have a significant overlap with the standard ellipsoid. Then the white point now can be regarded as the actual lighting condition. The illumination estimation was achieved.

Figure 3 visualizes the process of illumination estimation. The standard skin color ellipsoid is fixed. A new white point was found, using which as a reference white to convert image to LAB color space, the skin color in image will mainly fall within the standard ellipsoid.

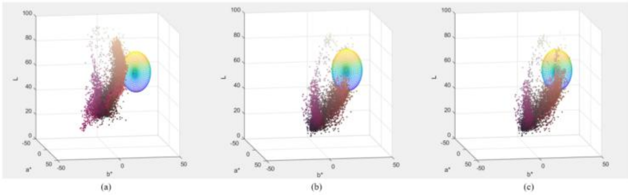


Figure 3. *An example of illumination estimation process : using different white point to convert image to Lab color space, and plot their pixel in Lab space together with the standard skin color ellipsoid(50% acceptance prefer skin color ellipsoid). (a) Converting the image to the LAB color space using D65 as the reference white. p<0.5 in this situation. (b) Converting the image to the LAB color space using predicted white point XYZ_W as the reference white. p=0.5 in this situation. (c) Converting the image to the LAB color space using predicted white point XYZ_W1 as the reference white. p=1 in this situation. XYZ_W1 is the estimated result.*

Figure 4 shows some results of illumination estimation, and the corresponding data was listed on the image. The estimated CCT(correlated color temperature) is very close to the actual value, and the Duv value is very low too. All images' illumination estimation results were gathered together and averaged. And the results in Table 1 shows that the relative error of the CCT of estimated illumination is lower than 0.2, and it decreases as the CCT becomes lower.

Figure 4 . *The results of several illumination estimation : (a) (b) (c) (d) (e) are respectively images captured in lighting condition D80, D65, D50, D40, D30. The different row represents pictures taken against different color background. The estimated illumination were listed on the image.*

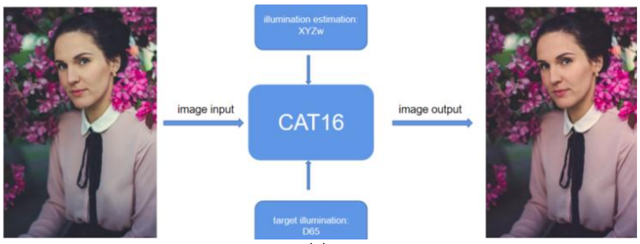
Table 1 Illumination estimation results for images captured under five different lighting condition.

Actual CCT	8000K	6500K	5000K	4000K	3000K
Average estimation	7692K	5997K	4456K	3548K	2644K
$\Delta T/T$	0.04	0.08	0.11	0.11	0.12

2.3 Image rendering

Through the above two step, the skin color segmentation and illumination estimation were done. With knowledge of the lighting conditions during photography, the CAT16 model[10] can be used to render the images to appear as if it were captured under any desired lighting environment. When the target output light source was set D65, the output image was white balanced.

Input the sRGB image, the illumination environment, and the target lighting condition, CAT16 model will calculate the transform matrix, and apply the matrix to rendering the image to any lighting condition. Figure 5 shows an example of image rendering.



(a)

(b)

Figure 5. *An example of image rendering process : (a) Original image, estimate illumination, CCT = 7904K, Duv = 0.0175, XYZ_W= [86.8,100,114.5], target illumination D65 input, the CAT16 model rendering the image and output, finishing the process of white balance. (b) Four different target illumination (D65, D50, D40, D30) rendering result.*

2.4 Dataset

To test the performance of our algorithm, a dataset was prepared, which contains 150 incorrectly white balanced images, including 2 genders (male and female), 3 ethnicities (Caucasian, Pakistani, Chinese), 5 different colored backgrounds (gray, red, green, yellow, blue), and 5 lighting source color temperatures.

The color temperatures include D80, D65, D50, D40, D30, all pictures were taken under the camera's WB setting D65, so the D65 lighting source group images can be regarded as ground-truth. The white balance algorithm proposed can be used to correct the ill white balanced images. Then, compared the corrected images with ground truth and calculate the color difference to evaluate the performance of the proposed algorithm and others.

Besides, all the images were taken against a large uniform colored background. White balance in images with a single-color background is a challenging problem in the field of white balance. Currently, there is not a specific white balance algorithm designed for this type of image. Therefore, we are creating this database to test the ability of our algorithm to process such images.

3. Result

For certain gender, ethnicity, colored background, the corresponding D65 illumination image were considered as the ground truth, and different white balance methods were used to correct the images taken under other illumination. Then calculated the color difference between the corrected images and the ground truth, and evaluated the performances of different algorithms on our database.

For better performance evaluation of the algorithms, two different color difference calculation methods were applied. One is the full image pixel difference calculation, the other is the skin color pixel color difference calculation.

The color difference calculate equation is based on CIELAB color space, and according to existing research, multiplying the L-index by a factor of 0.5 in color difference calculation may fit our visual perception better.

$$\Delta E = \sqrt{(\frac{\Delta L}{2})^2 + \Delta a^2 + \Delta b^2} \tag{4}$$

The compared white balanced methods were traditional gray-world, max RGB, and recently proposed WB_correction[6], and Deep WB[5]. The result of several images was shown in Figure 6.

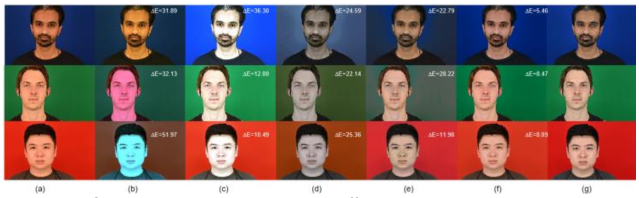


Figure 6. *Comparision between several different white balance algorithms on our database : (a) is the ill white balanced image. (b), (c), (d), (e), (f) representing respectively the gray world, max RGB, WB_correction[6], Deep WB[5], and method proposed in this research. (g) is the ground*

The first part is the skin pixel color difference, considering that skin color restoration is an important aspect in evaluating image quality, skin color pixels in the images were segmented and compared with the skin color in ground truth image. The result color differences are listed in Table 2.

Table 2 Skin pixels' color difference between corrected white balanced image and ground truth.The best performance data were marked red below.

Skin color	summary	
	mean	max
Gray world	45.58	77.82
Max RGB	34.37	78.18
WB_correction	14.21	31.19
Deep WB	9.58	20.90
ours	8.46	16.18

The second part is the full image color difference, the data were listed in Table 3. According to different background colors, the color difference data is collected and divided into six tables, including a summary table and five tables for different background colors.

As is listed in the Table 2 and 3. In the summary table, both full image and skin color, our method got lower mean value and max value, representing that in large colored background images' white balance correction, the algorithm proposed in this research achieved better result.

The images corrected by this method are closer to the ground truth. Besides, the reproduction was based on PMCC(prefer

Table 3 Full images color difference between corrected white balanced image and ground truth. The data were divided into six parts including a summary table and five tables for different colored background. The best performance data were marked red below.

Full image	summary		gray		red	
	mean	max	mean	max	mean	max
Gray world	36.01	59.62	7.60	13.84	54.07	59.62
Max RGB	26.39	66.19	25.53	39.18	16.56	25.73
WB_correction	19.67	45.30	6.66	13.45	21.89	42.35
Deep WB	22.05	46.53	8.13	14.66	14.77	23.06
ours	10.82	32.99	8.62	18.39	8.98	22.77

Full image	green		yellow		blue	
	mean	max	mean	max	mean	max
Gray world	32.85	38.80	51.10	55.96	34.44	49.74
Max RGB	23.77	38.62	34.15	66.19	31.95	59.62
WB_correction	23.97	38.84	25.52	42.90	20.30	45.30
Deep WB	25.83	33.59	37.66	46.53	23.88	38.93
ours	9.70	18.02	15.88	31.68	10.90	32.99

memory color chart), the result images shows more visual preference.

And in colored background groups, apart from the neutral gray background, our algorithm performs better, the mean value, max value are both lower.

v

4. Conclusion

Based on the PMCC, object detection algorithm, adaptive threshold skin color segmentation, illumination estimation in CIELAB color space, and image rendering model CAT16, these four parts were combined together, achieving a white balance algorithm which performs good on a large colored background database.

Considering the important effect of skin color on our visual perception, prefer skin color was used as a basis for white balance. The algorithm proposed can prioritize the accurate and preferred representation of skin colors. This makes the producing images more visually pleasing and natural-looking.

Furthermore, by changing the target skin color center, the skin color of the images can be rendered as personal preference, like healthier, more attractive,or etc. Using the accurate skin mask got ahead, the skin color in images can be reproduced more preferable while other pixels remaining unchanged. That’s to say, the white balance method proposed in the study can be regarded as a skin enhancement algorithm as well.

At last, the method mainly focus on images in the uniform lighting environment, in the future, we will improve our method, and try to apply this method in solving the problem of mix-illumination images processing.

5. References

[1] J. Van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *Ieee Transactions on Image Processing,* vol. 16, pp. 2207-2214, Sep 2007.

[2] D. A. Forsyth, "A NOVEL ALGORITHM FOR COLOR CONSTANCY," *International Journal of Computer Vision,* vol. 5, pp. 5-36, Aug 1990.

[3] J. T. Barron, Y. T. Tsai, and Ieee, "Fast Fourier Color Constancy," in *30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 6950-6958.

[4] Y. M. Hu, B. Y. Wang, S. Lin, and Ieee, "FC4: Fully Convolutional Color Constancy with Confidence-weighted Pooling," in *30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 330-339.

[5] M. Afifi, M. S. Brown, and Ieee, "Deep White-Balance Editing," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Electr Network, 2020, pp. 1394-1403.

[6] M. Afifi, B. Price, S. Cohen, and M. S. Brown, "When Color Constancy Goes Wrong: Correcting Improperly White-Balanced Images," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[7] J. J. M. Granzier and K. R. Gegenfurtner, "Effects of memory colour on colour constancy for unknown coloured objects," *I-Perception,* vol. 3, pp. 190-215, 2012.

[8] S. Bianco and R. Schettini, "Adaptive Color Constancy Using Faces," *IEEE Trans Pattern Anal Mach Intell,* vol. 36, pp. 1505-18, Aug 2014.

[9] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, and Ieee, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, 2016, pp. 779-788.

[10] C. J. Li, Z. Q. Li, Z. F. Wang, Y. Xu, M. R. Luo, G. H. Cui, *et al.*, "Comprehensive color solutions: CAM16, CAT16, and CAM16-UCS," *Color Research and Application,* vol. 42, pp. 703-718, Dec 2017.

[11] H. Z. Zeng and R. Luo, "Preferred Skin Color Enhancement for Photographic Color Reproduction," in *Conference on Color Imaging XVI - Displaying, Processing, Hardcopy, and Applications*, San Francisco, CA, 2011.

[12] H. Z. Zeng, R. Luo, and Is, "Preferred Skin Colours of Africans, Caucasians, and Orientals," in *19th Color and Imaging Conference - Color Science and Engineering Systems, Technologies, and Applications*, San Jose, CA, 2011, pp. 211-216.

[13] R. Peng, M. R. Luo, Y. C. Zhu, X. X. Liu, and M. Pointer, "Preferred skin reproduction of different skin groups," *Vision Research,* vol. 207, p. 15, Jun 2023.

[14] M. K. Cao and M. R. Luo, "Preferred skin tones on mobile displays under illuminants having different correlated colour temperatures," *Vision Research,* vol. 198, p. 12, Sep 2022.

[15] M. K. Cao, M. R. Luo, Y. Lu, and K. Xaio, "A study of preferred and natural memory colours across different ethnic groups," *Color Research and Application,* vol. 48, pp. 178-200, Mar 2023.