Collection and Analysis of a Representative Dataset of Human Skin Colors

Bolanle Dahunsi & Lucy Dunne, University of Minnesota

Introduction The fashion industry struggles with diversity of representation, and white female models tend to be overrepresented (Best & Choi, 2018; Butkowski et al., 2022). The ongoing dissatisfaction with this trend has led to efforts to include more people of color in fashion (The Fashion Spot, 2017). As representation improves, designers are beginning to consider a broader range of potential customers. However, the breadth of human skin tones is not well-characterized in a way that can be used by designers. This study therefore sought to clearly define the spectrum within which human skin colors fall using the RGB and LAB spaces, and understand which system provides a more clearly defined range for human skin colors.

Literature Review Many advanced techniques that could be used to evaluate skin color representation in the industry or allow designers to render concept garments on avatars with a representative breadth of skin tones rely on quantification of skin colors. This requires a clearly defined range of values within which human skin could fall. Color can be considered using multiple color spaces. The Basic RGB color space is a device dependent space that is most used in digital devices, and divides colors into red, green and blue components on a 0-255 intensity scale (Vezhnevets et al., 2003). The LAB space that is a perceptually uniform color space that represents color based on how two colors vary in appearance to a human observer (Ganesan et al., 2010). Here color is represented as three coordinates where L represents the lightness of the color (0 to 100), A represents its position between red/magenta and green and B represents its position between yellow and blue with A and B taking values between -128 and 127.

Sobottka & Pitas (1996) defined skin color ranges using the HSV color space as colors falling within the saturation values of Smin = 0.23, Smax = 0.68, and hue values of Hmin = 0°, Hmax = 50°. Yang et al. (1997) also conducted a study analyzing 48 human faces. They defined skin color ranges in the RGB space as having a mean of m_R = 188.9069, m_G = 142.9157, m_B =115.1863 and variance of σR = 58.3542, σ_G = 45.3306, σ_B = 43.397. The main limitation of using these studies is that they were conducted with the aim of being used in algorithms for human facial recognition in images or identifying pornographic images and used small convenience sample sizes. Given the differences in representation of skin colors for certain populations and the weakness of prior studies for defining skin representation for fashion industry purposes, we sought to understand the actual range of values within which human skin color could fall using two spaces that are widely used in the Apparel industry, basic color space using the RGB color space and an example of the perceptually uniform color space using LAB.

Methodology The sample consisted of color images of members of parliament from 32 countries. First, a list of the five countries in every continent with the highest percentage of female representation in the parliament was obtained from IPU Parline (2021) to ensure adequate representation of genders. Pictures of the members of parliament were then downloaded from the country's official webpage. Two other countries were then added to this list as they were found

to have a wider spectrum of the darker ends of skin tones (South Sudan) or a large variation in skin tones (India). Figure 1 shows the distribution of countries from which images were collected. The final set of images consisted of n=3566 unique images of which 40.8% were female. To extract color values, each image was imported into Photoshop and the average RGB and LAB numerical values for a visible section of skin with no shadows was extracted. The full dataset is accessible at https://conservancy.umn.edu/handle/11299/229549.

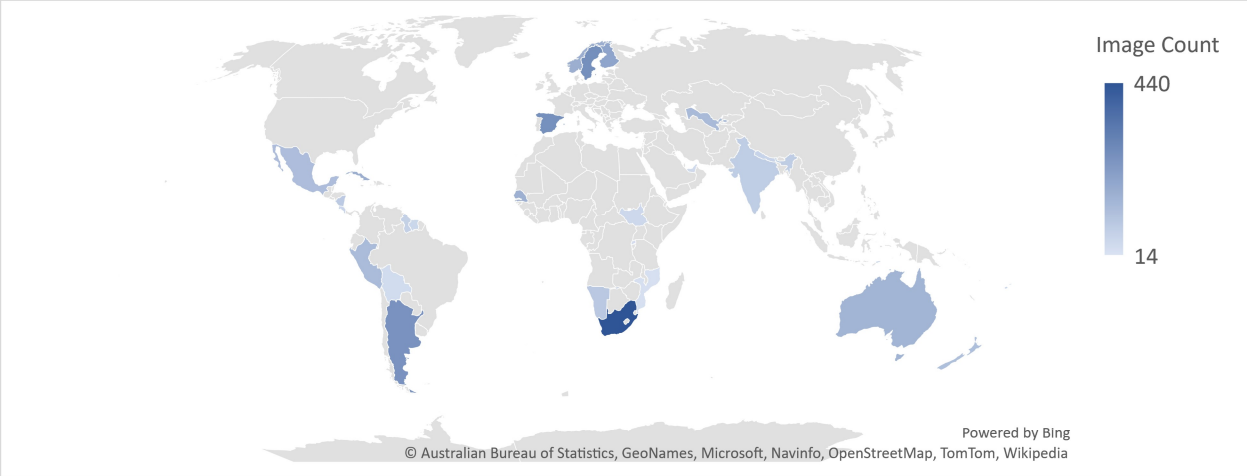

Figure 1: Countries sampled

Results Tables 1 and 2 provide the range, mean, variance, standard deviation, and the percent of color space occupied by skin tones for each color space. The distribution of numerical values in 3D of skin distribution in both color spaces are shown in Figures 2a and b.

Table 1

	Range	Mean	Var	SD	% skin
L	$8 \leq L \leq 89$	63.67	170.49	13.06	81.00
A	$-2 \leq G \leq 42$	17.14	28.98	5.38	17.25
B	$-3 \leq B \leq 50$	19.76	47.20	6.87	20.78

Table 2

	Range	Mean	Var	SD	% skin
R	$30 \leq R \leq 253$	192.10	1350.70	36.76	87.45
G	$10 \leq G \leq 216$	143.12	1137.79	33.74	80.78
B	$5 \leq B \leq 206$	120.61	1203.43	34.70	78.82

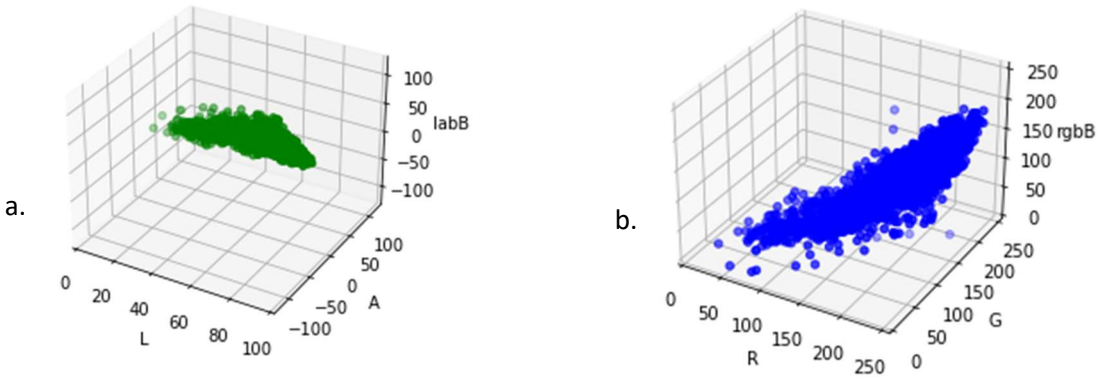

Figure 2a. 3D distribution of LAB values, b. 3D distribution of RGB values

Discussion Human skin color was found to fall within a clearly defined range in both the RGB and LAB spaces. This space was found to be more compact in the LAB space with values occupying only a small section of the A and B space while the lightness component was more widely dispersed. The percentage of the skin distribution in the RGB space was much larger as skin could fall anywhere within 78 to 87 percent of the color space. The LAB space could therefore prove to be a more useful space for identifying skin colors and determining the level of representation of people of all colors in the apparel industry.

References

Best, K. R., & Choi, E. (2018). Korean Fashion Media, Beauty Ideals, and Colorism: Examining the Prominence of Whiteness Between 2013 and 2017 in Ceci Magazine. *International Textile and Apparel Association Annual Conference Proceedings*, *75*(1), Article 1. https://www.iastatedigitalpress.com/itaa/article/id/1194/

Butkowski, C., Humphreys, L., & Mall, U. (2022). Computing colorism: Skin tone in online retail imagery. *Visual Communication*, 14703572221077444. https://doi.org/10.1177/14703572221077444

Ganesan, P., Rajini, V., & Rajkumar, R. I. (2010). Segmentation and edge detection of color images using CIELAB color space and edge detectors. *INTERACT-2010*, 393–397. https://doi.org/10.1109/INTERACT.2010.5706186

IPU Parline. (2021, June). *Monthly ranking of women in national parliaments*. Parline: The IPU's Open Data Platform. https://data.ipu.org/women-ranking?month=4&year=2021

Kakumanu, P., Makrogiannis, S., & Bourbakis, N. (2007). A survey of skin-color modeling and detection methods. *Pattern Recognition*, *40*(3), 1106–1122. https://doi.org/10.1016/j.patcog.2006.06.010

Sobottka, K., & Pitas, I. (1996). Extraction of facial regions and features using color and shape information. *Proceedings of 13th International Conference on Pattern Recognition*, *3*, 421–425 vol.3. https://doi.org/10.1109/ICPR.1996.546982

The Fashion Spot. (2017). *Diversity Report: Every Runway at New York Fashion Week Featured at Least 2 Models of Color for Spring 2018*. Https:// Www.Thefashionspot.Com/Runway-News/765783-Diversityreport-Every-New-York-Fashion-Week-Spring-2018/. https:// www.thefashionspot.com/runway-news/765783-diversityreport-every-new-york-fashion-week-spring-2018/

Vezhnevets, V., Sazonov, V., & Andreeva, A. (2003). A Survey on Pixel-Based Skin Color Detection Techniques. *IN PROC. GRAPHICON-2003*, 85–92.

Yang, J., Lu, W., & Waibel, A. (1997). Skin-color modeling and adaptation. In R. Chin & T.-C. Pong (Eds.), *Computer Vision—ACCV'98* (pp. 687–694). Springer. https://doi.org/10.1007/3-540-63931-4_278