Countering Racial Bias in Computer Graphics Research

Theodore Kim
Holly Rushmeier
Julie Dorsey
Yale University

Derek Nowrouzezahrai
McGill University

Raqi Syed
Victoria University of Wellington

Wojciech Jarosz
Dartmouth College

A.M. Darke
University of California, Santa Cruz

ABSTRACT

Current computer graphics research practices contain racial biases that have resulted in investigations into "skin" and "hair" that focus on the hegemonic visual features of Europeans and East Asians. To broaden our research horizons to encompass all of humanity, we propose a variety of improvements to quantitative measures and qualitative practices, and pose novel, open research problems.

ACM Reference Format:
Theodore Kim, Holly Rushmeier, Julie Dorsey, Derek Nowrouzezahrai, Raqi Syed, Wojciech Jarosz, and A.M. Darke. 2021. Countering Racial Bias in Computer Graphics Research. In *Proceedings of Under Review*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Over the last year, the murder of George Floyd and the worldwide protests that erupted in its wake have foregrounded the pervasive nature of systemic racism. Computer graphics research is no exception, as racial homogeneities in the historical composition of our community have contributed to racially biased research practices. The pale skin and straight hair targeted by our algorithms for virtual humans are a direct reflection of the European and East Asian researchers that designed them [11].

To build a better future, and realize ACM SIGGRAPH's Vision Statement of *Enabling Everyone to Tell Their Stories*, we must expand our palette of research problems to encompass the full spectrum of humanity. In the following, we will detail how racial bias pervades the technical language and numerical measures we use in research. To push against this historical inertia, we propose a new *complex, quarter circle* numerical measure, propose qualitative improvements to current research practices, and pose several historically-neglected research questions. Implementing these practices and investigating these questions are a first step towards a more comprehensive approach to computer graphics research.

1.1 *Acts*, Not *People*

For the current discussion, we define *acts*, and not *people*, as racist. In the absence of constant vigilance, we are all capable of committing *racist acts* that perpetuate existing systemic inequalities [9]. This perspective allows us to examine how seemingly neutral practices in computer graphics research have resulted, independent of any individual intent, in measurably biased outcomes. Our supplement provides further details and a bibliography.

Translucency and the corresponding physical mechanism of subsurface scattering has become synonymous with "human skin" in rendering. However, translucency is only the dominant visual feature of young, white Europeans and fair-skinned East Asians. We found 19 graphics publications, including the seminal works on the topic, that solely present renderings of white humans as evidence that subsurface scattering algorithms can faithfully depict "skin", "human skin" and "human faces." In at least 4 instances, this bias is then reflected in commercial software. Several other publications that include darker skin present them as deviations from the white baseline, further reinforcing the supremacy of whiteness. Researchers performing the seemingly neutral act of capturing their own appearances have instead perpetuated existing inequalities.

Similarly, "hair" has become synonymous with straight or wavy hair, and simulation and rendering papers cluster around this type. However, over a billion humans in Africa and its attendant diaspora have "afro-textured" or "kinky" hair. We only found *two* works in the graphics literature that attempt to capture the visual phenomena associated with these billion people. In contrast, 41 graphics publications, again including seminal works on the topic, solely present images of straight or wavy hair as evidence that the algorithms can faithfully depict "human hair". If we do not actively guard against our own biases, we will reproduce existing inequalities.

1.2 Existing Quantitative Measures

One potential solution is to use medical and cosmetics scales to precisely quantify which human characteristics that specific graphics algorithms are intended to depict. Multiple dermatological systems measure the darkness of skin, such as the Fitzpatrick scale [7] which classifies white European skin as Type I, and darker skin using progressively higher numbers up to Type VI, or the von Luschan and Taylor Hyperpigmentation scales, which assign Type 0 to white skin and respectively Types 36 and 10 to dark skin.

These measurement systems all share a problem clearly identified in Feminist and Critical Race Studies [6, 12]: European features are granted primary numerical status in the Fitzpatrick scale (literally the #1 ranking), or placed at the *center* or *origin* [2] in the cases of the von Luschan and Taylor scales. This grants white, European appearances centrality, while darker skin is placed at inconsistent locations (VI, 36, and 10) along an unbounded number line.

Existing hair typing systems share the same problem. The Walker Hair system widely used by stylists classifies straight hair as Type 1a. Progressively curlier hair is assigned higher numbers, up to Type 4c. The L'Oréal [4] system assigns Type I to straight hair, and

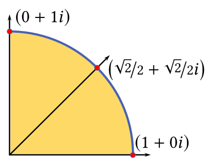
Under Review, Under Review, Online

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

Type VIII to kinky hair. Again, European features are #1 while Black features are inconsistently classed as either 4c or VIII.

This seemingly subtle ranking has visible impacts in the graphics literature on human appearance. Past works have included Fitzpatrick Types III and IV [5] in their measurements, but final renderings still only showcase white skin. Other works [10] have measured hair types with different elliptical cross sections (again, European hair scores closest to 1), but the final renders only depict straight or wavy blonde hair. The implicit ranking explicitly reproduces bias.

1.3 Proposed Anti-Racist Practices



To push against this historical inertia, we first propose a numerical measure that specifically targets the bias in existing systems: a *complex number* along the *unit quarter circle*. We apply this to hair by mapping Walker Type 1a to $(1+0i)$, and 4c to $(0+1i)$. A hair type in between then has a *Complex Walker* number of $(1/\sqrt{2}+1/\sqrt{2}i)$. This measure has several features.

- The origin $(0+0i)$ is excluded, so no value claims "the center."
- The appearance of a 1 signals the *endpoint* of a scale, but neither $(1+0i)$ nor $(0+1i)$ can claim the ranking of #1.
- All points on the scale have a numerical norm of 1. No value has a greater or lesser magnitude than any other.
- As additional parameters are needed, new complex axes can be added using quaternions or octonions.
- The dueling nomenclature of *complex* vs. *imaginary* reflects the larger social, political and biological discourse. Some will view this mapping as addressing an imaginary problem. We assert that the problems are both real and complex.

In other cases, switching to a quarter circle numerical measure will not be sufficient, because while adopting medical scales can seem appealingly objective, medical practices are equally susceptible to systemic bias [13]. Moreover, borrowing an existing scale can backfire if it misaligns with the appearance space of interest. Good-faith attempts to explore skin appearance across multiple Fitzpatrick values [14] then become beholden to a scale that was designed to classify skin according to its damage susceptibility under UV light and various treatments. We instead advocate multi-dimensional scales designed for visual appearance, such as the Pantone-based scale inspired by photographer Angélica Dass [3].

One-dimensional scales are appealing because they are straightforward to explore, but reproducing the range of human appearance will almost certainly require a higher-dimensional space. Rather than using the Fitzpatrick scale as a proxy for melanin concentration, we advocate direct reporting of concentrations, while underscoring that other factors like spatial and directional variation will introduce further non-linearities. Similarly, L'Oréal [4] uses PCA to reduce the wavelength, curvature and helicity of hairs into a single scale, but a comprehensive visual approach will need to disentangle these factors. Flattenings are a useful way to understand *outputs*, but can be misleading when used to simplify *inputs*.

Measuring and exploring the applicability of existing methods in higher-dimensional parameter spaces is an exacting and necessary process. In order to make progress, we as researchers must commit to encouraging and favorably reviewing work that both evaluates current methodologies, and develops novel techniques for measuring the applicable ranges of new and existing methods.

1.4 Future Research Directions

The preceding discussion suggests a variety of directions.

- What is a complete rendering model for Black skin? Where do the blue tones come from, and could it benefit from a custom, multi-layer BSDF model?
- What is an efficient simulation model for kinky, $(0+1i)$ hair? Collisions behave differently from the $(1+0i)$ case.
- What is an efficient rendering model for kinky, $(0+1i)$ hair? Can the near-isotropic strand distributions be leveraged? Do wave-based silhouette phenomena gain importance?
- What is the *gamut* of human appearance? How many coordinates are needed, and is it computationally tractable?

When investigating these questions, it is imperative to actively recruit and engage stakeholders that both inhabit the forms of humanity being modeled, and are the intended users of the technology [8]. This is scientifically sound practice: nobody is better positioned to formulate algorithms that capture the subtle qualities of Black skin and $(0+1i)$ hair than somebody who sees them in the mirror every morning. Such practice would increase the diversity of our community and reduce the risk of publishing well-intentioned but counter-productive "Blackface" skin or "Giant Afro" hair papers.

This work will require grappling with uncomfortable truths, and will not be immune from missteps and failure. However, these are necessary steps towards creating an environment that values honest self-reflection and open dialog. Progress is urgently needed, because computer graphics algorithms are increasingly being used to generate training sets for machine learning, so their potential to amplify existing inequalities has never been greater [1]. By learning from past missteps, we can instead move towards a future where everyone can tell their stories. We commit to rendering this future.

REFERENCES

[1] E. Bender, T. Gebru, A. McMillan-Major, and M. Mitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. In *Proc. of FAccT*.

[2] R. Benjamin. 2019. *Race After Technology*. Polity Press.

[3] A.E. Carrington, M. Vatanchi, and J. Jakus. 2020. Angélica Dass' Humanae, a spectrum of skin tones. *Int. J. Dermatol.* 59, 5 (2020), 640–642.

[4] R. De La Mettrie, D. Saint-Léger, G. Loussouarn, A. Garcel, C. Porter, and A. Langaney. 2007. Shape variability and classification of human hair: a worldwide approach. *Human biology* 79, 3 (2007), 265–281.

[5] C. Donner, T. Weyrich, E. d'Eon, R. Ramamoorthi, and S. Rusinkiewicz. 2008. A layered, heterogeneous reflectance model for acquiring and rendering human skin. *ACM Trans. Graph.* 27, 5 (2008), 1–12.

[6] R. Dyer. 1997. *White*. Routledge.

[7] T. Fitzpatrick. 1975. Soleil et peau. *J Med Esthet* 2 (1975), 33–34.

[8] B. Friedman, P. Kahn, and A. Borning. 2008. Value sensitive design and information systems. *The handbook of information and computer ethics* (2008), 69–101.

[9] I. Kendi. 2019. *How to be an antiracist*. One world.

[10] P. Khungurn and S. Marschner. 2017. Azimuthal Scattering from Elliptical Hair Fibers. *ACM Trans. Graph.* 36, 2, Article 13 (April 2017), 23 pages.

[11] T. Kim. 2020. The Racist Legacy of Computer-Generated Humans. *Scientific American* (2020). https://bit.ly/3thBeye

[12] A. Phillips. 2020. *Gamer Trouble*. NYU Press.

[13] M.W. Sjoding, R.P. Dickson, T.J. Iwashyna, S.E. Gay, and T.S. Valley. 2020. Racial Bias in Pulse Oximetry Measurement. *NEJM* 383, 25 (2020), 2477–2478.

[14] T. Weyrich, W. Matusik, H. Pfister, B. Bickel, C. Donner, C. Tu, J. McAndless, J. Lee, A. Ngan, H.W. Jensen, and M. Gross. 2006. Analysis of Human Faces Using a Measurement-Based Skin Reflectance Model. *ACM Trans. Graph.* 25, 3 (July 2006), 1013–1024.