Gloss-Aware Color Correction for 3D Printing

Jorge Condor
jorge.condor@usi.ch
Università della Svizzera italiana
Lugano, Ticino, Switzerland

Michal Piovarci
michael.piovarci@ist.ac.at
ISTA
Klosterneuburg, Austria

Bernd Bickel
bernd.bickel@gmail.com
ISTA
Klosterneuburg, Austria

Piotr Didyk
piotr.didyk@usi.ch
Università della Svizzera italiana
Lugano, Ticino, Switzerland

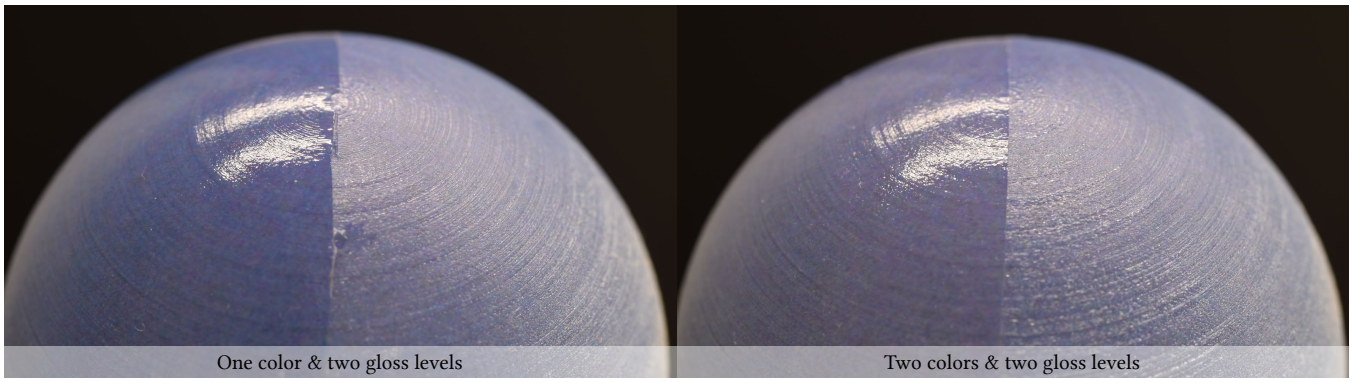

Figure 1: An example application of our method, where we account for any color shift stemming from the usage of different varnishes and varying gloss, to enable uniform and predictable color even in spatially-varying gloss conditions. Left is a sphere 3D printed with a single color, half of it varnished glossy, while right uses our method, changing slightly the base color of the varnished side. Pictures taken under complex illumination and identical conditions.

ABSTRACT

Color and gloss are fundamental aspects of surface appearance. State-of-the-art fabrication techniques can manipulate both properties of the printed 3D objects. However, in the context of appearance reproduction, perceptual aspects of color and gloss are usually handled separately, even though previous perceptual studies suggest their interaction. Our work is motivated by previous studies demonstrating a perceived color shift due to a change in the object's gloss, i.e., two samples with the same color but different surface gloss appear as they have different colors. In this paper, we conduct new experiments which support this observation and provide insights into the magnitude and direction of the perceived color change. We use the observations as guidance to design a new method that estimates and corrects the color shift enabling the fabrication of objects with the same perceived color but different surface gloss. We formulate the problem as an optimization procedure solved using differentiable rendering. We evaluate the effectiveness of our method in perceptual experiments with 3D objects fabricated using a multi-material 3D printer and demonstrate potential applications.

SIGGRAPH '23 Conference Proceedings, August 06–10, 2023, Los Angeles, CA, USA

ACM ISBN 979-8-4007-0159-7/23/08.
https://doi.org/10.1145/3588432.3591546

CCS CONCEPTS

• **Computing methodologies** → **Reflectance modeling**.

KEYWORDS

color, gloss, perception, color compensation, color management

ACM Reference Format:
Jorge Condor, Michal Piovarci, Bernd Bickel, and Piotr Didyk. 2023. Gloss-Aware Color Correction for 3D Printing. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Proceedings (SIGGRAPH '23 Conference Proceedings), August 06–10, 2023, Los Angeles, CA, USA.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3588432.3591546

1 INTRODUCTION

Accurate color reproduction is fundamental for many applications ranging from display technology to fabrication, and art. A key problem in color reproduction is ensuring the final visible color remains consistent across different media. To this end, extensive color management systems were developed [X–Rite 2005]. Typically, the systems contain a small set of well-behaved colors that achieve consistent reproduction on both digital and physical devices.

Unfortunately, accurate color reproduction still remains a challenging endeavour due to the complex nature of color perception. To predict a visible color one has to account for both the surface properties of the observed object as well as the ambient illumination. The effects of illumination are well known and understood with ready-made models available [Changjun et al. 2017; Moroney et al. 2002]. The interaction between visible color and material is significantly more complicated. Traditionally the color was assumed to be governed by sub-surface scattering. Yet practitioners observed that by applying translucent varnishes of different glossiness the same color can appear more or less saturated [Jakob 2010]. While this effect is sufficiently strong to motivate manufacturers to provide two sets of color references[1], there is no systematic study on how exactly different glossy finishes affect the perceived color of a surface.

In this work, we investigate the coupling between color and gloss perception. We start by investigating how the surface finish affects the perception of color. To this end, we propose a novel experimental design that improves on prior studies. Based on the experimental results, we propose a model that predicts the visible color difference. We further provide an end-to-end differentiable rendering pipeline that can automatically compensate for the perceived color shifts. To validate the pipeline, we run several studies investigating its performance on varying colors, geometries, and environmental conditions. Finally, we showcase an application of our novel color management tool towards 3D printing.

2 RELATED WORK

The human visual system (HVS) is one of the most studied senses [Hutmacher 2019]. In this section, we provide a brief overview of color and gloss perception work and delegate the reader to more exhaustive surveys. We then focus more extensively on coupled color-gloss perception studies and the state-of-the-art in color-gloss compensation for fabrication.

Color Perception. Color is one of the key visual attributes of a surface. As such many works have attempted to characterize the perceived color differences. The seminal work of Hunter and colleagues [1958] proposes a perceptually uniform color space, i.e, a color space where the Euclidean distance corresponds with the perceived difference. This gave rise to the very first color difference formula the so-called $\Delta E76$. Since then many improvements have been proposed [Changjun et al. 2017; Moroney et al. 2002; Sharma et al. 2005]. In this work we rely on the most widely used version of color difference, the $\Delta E00$ [CIE 2001]. For a more detailed overview of color perception we refer the reader to the classical book of [Wyszecki and Stiles 1982].

Gloss Perception. Next to color, gloss is perhaps the second most important visual attribute of a surface [Anderson 2011; Fleming 2014]. Proper modeling of perceived gloss is a long-standing problem [Marlow et al. 2012]. Over time it was shown that gloss depends on the ambient illumination [Zhang et al. 2020], overall object geometry [Serrano et al. 2021], bumpiness of the surface [Ho et al. 2008], and surface curvature [Kim et al. 2012]. By contrast, angle of illumination does not impact gloss perception in flat surfaces when illuminated by point light sources [Obein et al. 2004]. We take inspiration from prior work in the design of our stimuli. Our correction model relies on natural illuminations that were shown to perform best for visual gloss estimation [Fleming et al. 2003]. With regards to geometries, we base our correction on a smooth object with overall low curvature commonly used to preview car lacquers in the automotive industry. In validation studies, we explore how our model generalizes to different illumination conditions [Pereira and Rusinkiewicz 2012], and more complex geometries [Havran et al. 2016]. For a more detailed overview of gloss perception, we refer the reader to the excellent survey of [Chadwick and Kentridge 2015].

Color-Gloss Coupling. To investigate the effect of surface color on perceived gloss researchers conducted trials in both virtual [Pellacini et al. 2000] and real environments [Samadzadegan et al. 2014]. The results suggest that the perceived gloss is relatively constant under hue and saturation changes. On the other hand, lightness change directly correlates with perceived gloss, with darker materials appearing glossier.

The interplay between gloss and perceived surface color has received little scientific interest in the past and its conclusions are conflicting. While some studies claim that there is in fact a small perceptual color shift due to the presence of varying levels of gloss in objects [Giesel and Gegenfurtner 2010; Xiao and Brainard 2008], others attribute it to just a measurable change in contrast and lightness, and affirm it can be controlled via measuring the variation between the specular and diffuse components [Dalal and Natale–Hoffman 1999; Ma et al. 2009]. The studies are based on the assumption that gloss is a physical property of objects which can be defined by a single objective measure, in this case, the specularity of a surface. However, it has been shown that such description is insufficient and gloss is primarily a perceived quality [Chadwick and Kentridge 2015]. In contrast, our psychophysical experiments investigate the color-gloss interplay as a perceptual problem that depends on cues such as geometry and illumination.

Color-Gloss Fabrication. In computer graphics, accurately reproducing the target appearance of digital assets is one of the longest standing challenges. Recent works investigate individual aspects such as color [Babaei et al. 2017; Elek et al. 2017; Sumin et al. 2019], translucency [Brunton et al. 2018; Urban et al. 2019], or gloss [Elkhuizen et al. 2019; Piovarči et al. 2020]. In the more traditional 2D printing context the visible difference in color reproduction based on glossiness of the substrate is a well known problem [X–Rite 2016]. State-of-the-art research and commercial solutions propose a correction of the perceived color based on colorimetric measurements [Baar et al. 2014; Datacolor 2022]. The key idea is that the specular reflections will reflect more light towards the measurement sensor and the final color will appear brighter and less saturated. However, colorimetric data alone is not well suited to estimate the shift in perceived color. Measurements only account for light scattering in the varnish surface which reduces the problem to adding/removing light. However, the added/removed light would be entirely dependent on the object's illumination, and would not consider any potential perceptual interaction. In contrast, our model takes into account the full picture: objects' color, gloss, illumination, and geometry. We demonstrate that our model can

[1]Pantone Formula Guide | Coated & Uncoated

correctly estimate a significant perceived difference in color due to varying levels of gloss even for cases where colorimeter devices would not measure any.

3 GLOSS-COLOR INTERACTION EXPERIMENT

To assess how humans perceive the colors of objects with different gloss levels and derive a perception-informed method for correcting potential color shifts, we first conducted a perceptual experiment.

3.1 Methods

Our methodology is inspired by previous literature on color perception, and more specifically, color constancy [Arend and Reeves 1986]. We perform a color-matching experiment where subjects match color patches with a color of an animated geometry rendered under complex illumination. We vary both color and surface finish which enables us to assess perceived color shifts as a function of gloss level.

Stimuli. A single stimulus consisted of a video of a rotating *car* geometry rendered under complex illumination. Each video was rendered using a physically-based path tracer [Nimier-David et al. 2019] in 1080×1080 resolution and was shown in the center of the screen using its original size (Figure 2, top).

We considered 16 different base colors and 5 different gloss levels (Figure 2, bottom). For the gloss levels, the BRDF used for rendering was obtained by measuring samples of different varnishes [Piovarči et al. 2020] and fitting them to the Cook-Torrance model [Cook and Torrance 1982] and a GGX normal distribution [Trowbridge and Reitz 1975]. Another two were chosen manually as very matte and glossy materials. We used all color-gloss combinations leading to 80 different stimuli. For further information on the stimuli and hardware, we refer to the supplemental material.

Task. In each trial, the participants were shown a continuous video of one rendered geometry. On the right-hand side of the video, a 3×3 grid of similar colors was shown, which was part of a bigger 21×21 grid through which the subjects could navigate lightness and chroma axis from the CIECAM16 color appearance model [Changjun et al. 2017] at a fixed distance of 1.35 ΔE00 between row-adjacent pairs. The task was to navigate this table using the arrow keys to find the color they perceived as closest to the color of the *car*. All participants performed the task for 80 stimuli without repetitions, at a randomized order and initial grid positions. The experiment took between 20 and 40 minutes, and participants were instructed to take as much time as needed. They were also allowed to take breaks in case of fatigue. We got 11 participants, which we tested for color acuity [Ishihara 1917] [Farnsworth 1943]. For more details on participants, we refer to the supplemental material.

Discussion. Our experiment addresses some of the limitations of previous works [Dalal and Natale–Hoffman 1999; Ma et al. 2009; Xiao and Brainard 2008]. We use rotating references, which prevent participants from performing pixel-to-pixel comparisons and provide additional visual cues for disambiguation of view-dependent effects, such as reflections and highlights, from the pure surface color. Forced motion on the screen also better simulates natural viewing conditions, where the subjects are expected to move, and improves gloss perception [Scheller Lichtenauer et al. 2013].

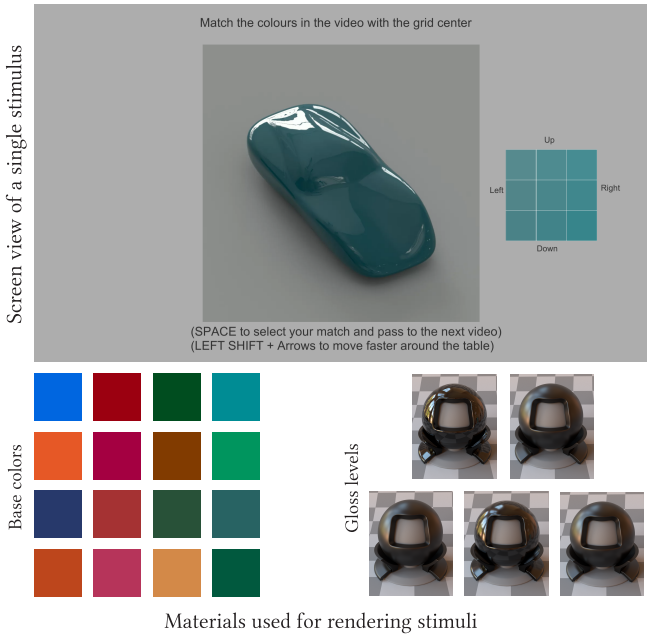

Figure 2: Overview of the stimuli used in our experiment. Top: screenshot of the rotating *car* video in the middle and the 3×3 sliding window. Bottom: all colors and materials used for rendering the *car* geometry.

Illumination is a significant factor in material perception [Zhang et al. 2019]. Our stimuli were rendered using environment maps representing natural, real-world illumination to improve the material depiction [Fleming et al. 2003] and simulate natural viewing conditions.

Crucially, we use color patches to match the color of the rendered geometry. An alternative approach is to use rendering for both test and reference. Since on-the-fly rendering is prohibitively expensive (if high physical accuracy is desired), such an approach would need to rely on image blending, which does not account for the non-linearity between the base color of a sample and on-surface color [Xiao and Brainard 2008]. Our approach avoids this problem by analyzing color shifts between matched patches.

A significant difference from other studies is the color dimensions we used in our experiments. While many previous works[Giesel and Gegenfurtner 2010; Granzier et al. 2014; Xiao and Brainard 2008] rely on standard color spaces, such as CIELAB, HSV, and HSL, we rely on color attributes of a recent color appearance model, CIECAM16. The advantage of such an approach is that the model already accounts for many perceptual effects related to color perception (i.e. the Abney effect). The use of CIECAM16 for the construction of our reference grids also allowed us to maintain a constant perceived hue across the whole grid.

Previous studies [Giesel and Gegenfurtner 2010; Granzier et al. 2014] and our own early experiments indicated that people do not perceive a change in hue among varying degrees of gloss, other than the expected shift from altering other color parameters such

as saturation or lightness (i.e. Abney effect [Abney 1909]). By limiting the search space this way, we make our experiment easier to navigate and analyze.

3.2 Data Filtering

Even though our experiment considers only two color attributes (lightness and chroma), the experiment proved to be challenging for participants (see Figure 3 for data spread). Apart from normal experimental variability, some subjects occasionally skipped color pairs accidentally, which partially explains some clear outliers. Therefore, we apply outlier removal. However, the problem with visualizing data in lightness/chroma axis is that the color space of our grid is not linear, as a constant perceived difference between co-linear colors row-wise is ensured but it can drastically vary between colors diagonally and column-wise. A better way of visualizing this data is to compute a confusion matrix with the pairwise distance of every data sample to the rest (Figure 3). We filtered data by creating a distribution with the cumulative sum of ΔE00 differences of each sample to the rest of them for every given color-material pair. We performed the Shapiro-Wilk normality test on these distributions, which showed that most (77.5%, $p > 0.05$) followed normal behavior. We then rejected samples above $\mu + \sigma$ (84.1% of the samples are kept). The result is akin to fitting a 2D Gaussian to reject samples outside of the main cumulus or "blobs", rejecting data points far away from the main consensus.

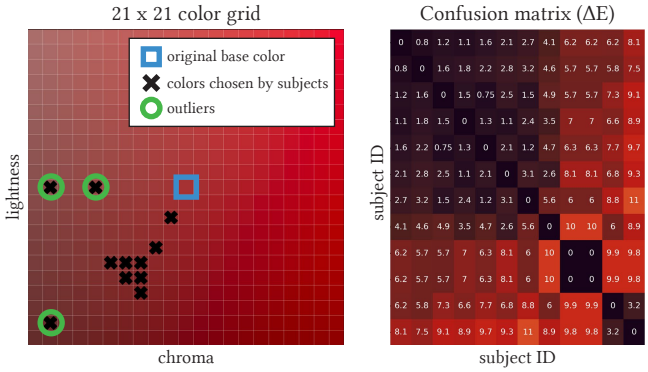

Figure 3: The figure presents the outcome of the experiment for one color-gloss combination. Left: 21×21 color grid with subjects' matches, the original color used for rendering, and the removed outliers. Right: confusion matrix containing pairwise Δ*E*00 differences between subject's choices.

3.3 Results and discussion

Gloss-color correlation analysis. First, we look for correlations between gloss levels and any other color parameter. Previous research has shown some degree of correlation between perceived gloss and fitted model material roughness [Pellacini et al. 2000], hence we will use it as our measure of gloss to facilitate the analysis. The results can be seen in Figure 4. We analyze picked colors according to the CIECAM16 color appearance model, of which we show the two most significant axis, lightness and saturation. For each color, we plot the average value of the different color components across all test subjects against the roughness of the sample. A single reference is created with the roughest or most matte sample by averaging all answers for that given color-material. The Pearson correlation test reveals a very strong negative linear correlation between gloss and perceived saturation shift ($p = -0.9249$).

In order to answer whether this shift suggests a perceptual effect or not, we looked at how subjects interacted with the experiment. For the given color matching task, there are several approaches to assess a sample's global color. One could average across various pixels, fixate on specific areas or obtain a global assessment that does not closely match any particular area of the object. We analyze this by computing the ΔE00 of the average picked color to every pixel of the reference videos for every color-material pair (Figure 5). We found that test subjects tended to match their colors with a specific area in our objects, one directly illuminated by the most prominent light source in the scene and with no self-shadowing, reflections or high frequency illumination of any kind, an area of "constant" color. At the same time, their choice tended to be substantially more saturated and darker the glossier the sample was, and likewise is more saturated than the global average of the whole surface, discounting just specular highlights. This is in agreement with previous color research on the topic of global color assessment [Sunaga and Yamashita 2007] and suggest the perceptual effect of gloss over color stems from the different ways of matching and interacting with the range of colors displayed on a given surface, which changes with gloss.

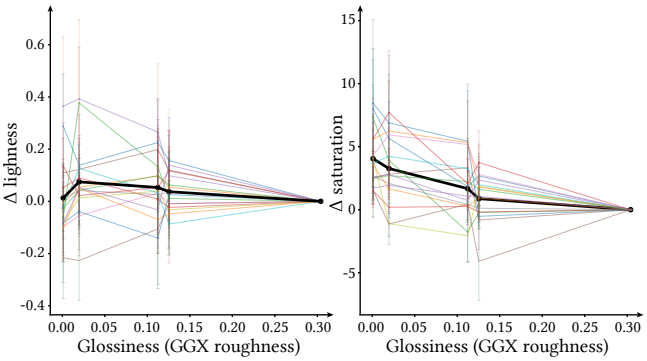

Figure 4: Average picked colors for every color-material pair, with error bars showing 1.0 standard deviation (faded), broken down into lightness and saturation using the CIECAM16 color appearance model. In black, average saturation and lightness across colors for every material, which showcases a strong correlation between perceived color saturation shift and gloss levels, while discarding any consistent lightness and gloss interaction.

4 GLOSS-AWARE COLOR COMPENSATION FOR 3D PRINTING

Our psychophysical experiment revealed an interaction between surface gloss and the perceived color. We now seek to formulate a computational method for correcting the perceived color shift. The input to our method is a color and two glossy finishes, an

original and a novel one. The output is a new color that when viewed under the novel gloss appears the same as the original color under the original gloss. Unfortunately, deriving the correction directly from experimental data is challenging as it depends on the complex interplay between objects geometry, color, gloss, and the environment illumination. Instead, we assume that the perceived color shift for different gloss levels can be largely explained by physical changes in the surface appearance and how people judge the color of the complex surface under complex illumination. In this section we describe how we correct for the perceived color shift of surfaces with varying gloss.

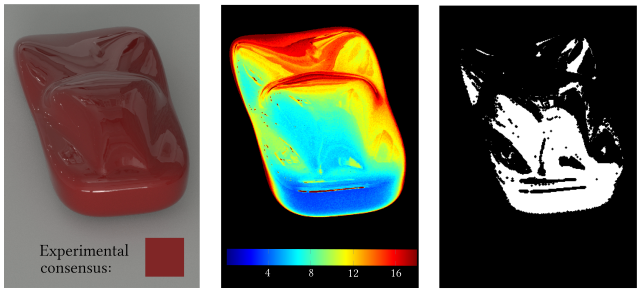

Figure 5: #a53233 (high gloss) video frame shown during the experiment as reference (left); CIE2000 ΔE heatmap between the experimental consensus color and the frame (middle); mask obtained through our method (right). We can see how subjects disregarded illumination-specific effects such as reflections, sheens and specular highlights, and focused on color-stable regions in the video (back of the car). The observation was confirmed in informal discussions with subjects after the experiment.

4.1 Rendering

To model the complex interplay between surface geometry, illumination, and the material properties we rely on physically-based rendering [Jakob et al. 2022]. The materials we aim to model are highly translucent plastics potentially covered with a thin dielectric varnish, which we use to generate different surface finishes in our applications. We model our materials from captured data using the Cook-Torrance model with GGX distribution [Trowbridge and Reitz 1975]. For rendering, as in our perceptual experiment, we use these captured BRDFs with Mitsuba's rough-dielectric plugin, which we apply on a generic geometry. For a more in-depth look as to why this is a good approximation of physical behavior, please see the supplemental material. Finally, we illuminate the scene with high-resolution photometrically captured gray-scale environment maps representing natural and complex illumination conditions.[2]

4.2 Computing Reference Color

To drive the optimization of the color correction, our method requires a strategy to aggregate the rendered spatially-varying surface appearance into one single color for comparison. This strategy should ideally resemble the way the HVS assesses the color of complex objects. Following previous research [Sunaga and Yamashita 2007], and based on the findings of our perceptual experiment, we assume that subjects discard regions that are largely affected by the view-dependent illumination effects such as reflections, specular highlights, and sheens, and use a simple strategy that computes a map that excludes these regions from the mean color computation based on the used varnishes, while biasing towards higher saturation colors when presented with a same-hue, spatially varying color appearance. We formulate this as follows:

$$\mathcal{M}(c, f_g, f_m) = \bigcap_{b \in B} (\mathcal{M}_s \circ \mathcal{M}_v)_{-b}, \tag{1}$$

$$\text{where} \quad \mathcal{M}_s = \left[\mathcal{R}^{sat}(c, f_g) - \mathcal{R}^{sat}(c, f_m)\right] > 0,$$

$$\mathcal{M}_v = \left[\mathcal{R}^{L}(c, f_g) - \mathcal{R}^{L}(c, f_m)\right] > \mu,$$

$$\text{st.} \quad f_g < f_m,$$

where $\mathcal{M}$ is the pooling mask, c is the color of the geometry, f_g and f_m are roughness coefficients for a glossy and matte varnish, respectively. $\mathcal{R}^{sat}$ is the saturation of a rendered image, $\mathcal{M}_s$ behaves similarly to the participants of our experiments and filters out colors where the saturation is not significantly affected by view-dependent effects, $\mathcal{R}^{L}$ is the lightness of a rendered image, with $\mathcal{M}_v$ acting as a mask that removes sheens and bright specular reflections based on the threshold $\mu = -0.05$, which enables small color variations but fundamentally excludes these effects. Both masks are combined through a Hadamard product ($\circ$). Finally, we perform a binary erosion, where B is the binary erosion structuring element and $(\mathcal{M}_s \circ \mathcal{M}_v)_{-b}$ denotes the translation of our mask by $-b$, the morphological structuring element at a given pixel position. This reduces the impact of rendering artifacts (i.e., fireflies). The binary mask $\mathcal{M}$ needs two different gloss levels to estimate the view-dependent effects. In practice we found out the mask shows little variation with respect to the difference in gloss between the two varnishes and across different colors. However, for maximum accuracy during optimization we pre-compute it at the beginning for every color and material pairs. Our resulting masks closely follow the heatmaps (Figure 5), indicating that our perceptually-motivated heuristic models the color-gloss interaction accurately.

4.3 Optimization

To minimize the perceived difference between reference colors of two objects with varying glossy finish we rely on differentiable rendering [Jakob et al. 2022]. Our optimization seeks to match the perceived color of a sample with a specific gloss appearance while preserving the hue of the desired color. More formally it attempts to minimize our loss function $\mathcal{L}$ as defined below:

$$\min_{c} \quad \mathcal{L} = \mathcal{L}_1 + w_{reg}\mathcal{L}_2, \tag{2}$$

$$\text{where} \quad \mathcal{L}_1 = \frac{1}{\mathcal{P}} \sum_{p \in \mathcal{P}} || \left[\mathcal{M} \circ \mathcal{R}(c_r, f_r)\right](p) - \left[\mathcal{M} \circ \mathcal{R}(c, f_t)\right](p) ||_2^2,$$

$$\mathcal{L}_2 = ||c_r^{hue} - c^{hue}||_2^2,$$

$$\mathcal{M} = \mathcal{M}(c, \min(f_g, f_m), \max(f_g, f_m)),$$

[2] https://polyhaven.com/

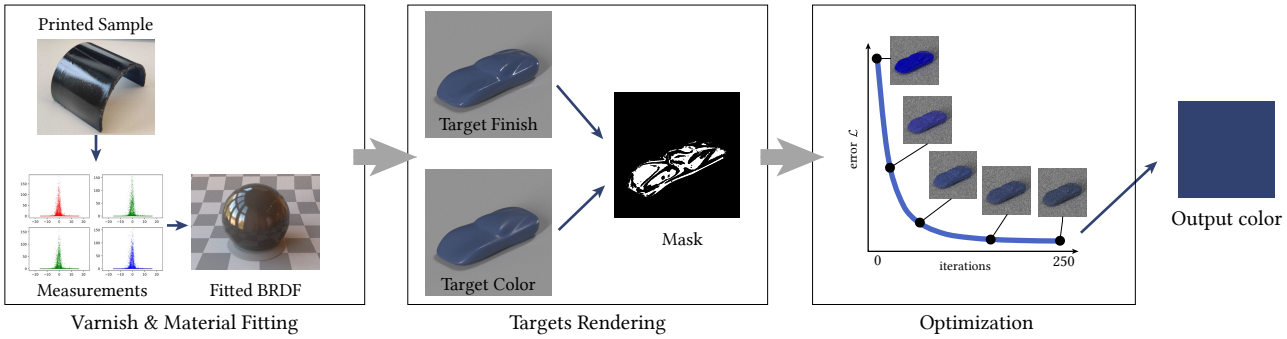

Figure 6: Method overview. First, we capture and fit the BRDF of the materials. We then render a geometry with both target and reference finishes, with the desired color.A mask is computed from them, selecting the pixels that will be entering our optimization stage. We use differentiable rendering and our custom color loss to obtain a surface color for our target gloss that will be perceptually equivalent to the desired one in the reference gloss.

where we optimize for a new color c for a target gloss f_t that matches the perceived color of reference color c_r with reference gloss f_r. The loss function $\mathcal{L}$ is composed of two different terms. $\mathcal{L}_1$ is defined as the per-pixel (p from set of image pixels $\mathcal{P}$) mean squared difference between the Hadamard product of our perceptually inspired pooling strategy $\mathcal{M}$ and both the target appearance's render $\mathcal{R}(c_r, f_r)$ and currently proposed compensation's render $\mathcal{R}(c, f_t)$, the second term is strictly a hue loss computed in CIECAM16 space, between the base rendering color of the target appearance, and $w_{reg} = 0.05$ is a small regularization weight. Neither our psychosphysical experiment nor previous works[Giesel and Gegenfurtner 2010; Xiao and Brainard 2008] have found any correlation between gloss and perceived hue, which is why it is crucial to get the exact same hue across both reference and target. This loss term ensures that any remaining error does not manifest in hue, as slight shifts in chroma or lightness are more acceptable. The optimization takes around 5 minutes for a given target color-gloss. For more details about the parameters and settings please see the supplemental material.

5 RESULTS AND APPLICATIONS

In this section, we start by validating our correction on a wide range of colors. Next, we compare our correction method with state-of-the-art color-gloss compensation. To further validate our method, we conduct several ablation studies. Finally, we demonstrate an application of our correction to 3D printing.

Method. For the evaluation we consider a forced choice experiment. We present the participants with a design manufactured with a reference color and finish (i.e., matte). We then show them two possible reproductions using a different finish (i.e., glossy) and either the reference color or our compensation. We then ask the participants which design is preferred. To physically realize our stimuli we use the J750. To validate that the printed colors lie inside the gamut of the device, we rely on a color validation tool provided by the manufacturer. After printing, the objects are cleaned and varnished.

Participants. We recruited 18 participants (14 male, 4 female) aged between 20 – 31 with normal or corrected to normal vision and perfect color acuity [Farnsworth 1943; Ishihara 1917].

5.1 Validation Experiment

For the validation experiment, we manufactured 48 reference stimuli (24 color pairs) and 24 proposed compensations (one for each pair). We further divided our stimuli into 3 different experiments. 8 samples were varnished *Edding Clear Matte*, while their pairs were varnished *Edding Clear Gloss*. Our aim was to compute a sample that, when varnished with *Edding Clear Gloss*, would be perceived as having the same color as the one varnished in matte. In other 9 samples, the target was changed from *Edding Clear Matte* to *Schmincke Matte*, with the same purpose, and varnish-specific compensations were proposed and printed for these pairs. Finally, the last 7 samples attempted the opposite target (matching a reference in *Edding Clear Gloss* while being varnished in *Edding Clear Matte*). A small subset of these color triads, as presented to test subjects, can be seen in Figure 9, while the full results of our study are shown in Figure 10, left.

We can observe that our correction achieved significant improvement over the reference color. On average, around 90 percent of the participants preferred the corrected colors ($p < 0.001$) for both *Schmincke* and *Edding Clear* varnishes, even with large corrections of up to 8.737 $\Delta E00$. In the cases where our performance was below average, it never dropped under 50% suggesting our correction never worsened the perceived mismatch between the colors based on varying glossy finish. However, results do suggest worse performance for the reference in gloss, target in matte compensation (3rd column 10, left), while still being statistically significant ($p < 0.001$). We attribute the loss in performance to the direction of the shift towards more saturated colors that can fall outside of the printing gamut.

5.2 Comparison with State-of-the-Art

State-of-the-art research and commercial methods rely on computing corrections based on colorimetric measurements [Baar et al. 2014; Datacolor 2022]. To compare with them, we prepared a set of

8 samples with reference *Edding Clear Glossy* finish compensated towards *Edding Clear Matte*. We experimentally verified with an xRite i1 Pro 3 spectrometer featuring a 45/0 measurement geometry that the glossy and matte colors have the same measured value (measured $\Delta E00 < 1$ calculated in CIE L*a*b* using a D65 illuminant and the 2° observer). As such, state-of-the-art methods would propose no correction. In contrast, our system for the same colors proposes a correction of up to 8 $\Delta E00$. The results of the experiment are shown in Figure 10, left.

We can see that in each investigated case, participants preferred our correction ($p < 0.001$), which demonstrates a shift between perceived colors was indeed present despite the lack of significant measured colorimetric differences. This demonstrates that our method has a closer agreement with human perception of color and gloss than prior work. We attribute our improvement to the more complex estimation of reference color that includes the effects of geometry, illumination, and perception, while also addressing the physical source of these shifts through physically-based rendering, as opposed to attempting to assess gloss through a single measurement regardless of its physical origin.

5.3 Ablation Study on Geometry

Our color correction method has been validated on a single geometry that manifests mostly smooth features. Since the effect of geometry on perceived gloss is well known [Serrano et al. 2021], we investigate the robustness of our method to changes in geometry. Based on prior work, we employ two geometries, *bunny* geometry with higher frequency features [Ho et al. 2008], and *ghost* designed for material perception studies [Havran et al. 2016], Figure 8. The results of the experiment can be appreciated in Figure 10, right.

For the *bunny* geometry, we see no significant degradation of performance ($p < 0.001$). In contrast, the performance for the *ghost* is lower while still being positive and statistically significant ($p < 0.001$). We attribute this to the challenging nature of the geometry, which features many inter-reflections and self-shadowed regions. A small decrease of performance was also expected given the slight influence of surface texture on the perception of glossiness [Baar et al. 2016]. Even in such a challenging scenario, our correction was mostly preferred and, at worse, performed on par with respect to the reference. However, we expect the performance to improve on any geometry when the method uses the target geometry during optimization, as we have shown in our *cars* validation tests (Figure 10, right).

5.4 Ablation Study on Illumination

The original illumination used during our optimization followed guidelines that suggested using a scene with complex natural illumination to mimic normal viewing conditions. To stress-test our method, we evaluate using a complimentary environment map that captures a tunnel with very strong, focused, and sparse area lights (Figure 8). We proceeded to replicate these in our lab conditions by lowering the blinds and turning on our fluorescent lamps. We then tested both proposed compensations (which were computed under these two different environment maps) under loosely matching (i.e., *museum* with a mix of natural and point-light artificial lights) and non-matching environmental lighting conditions (i.e., *museum* with solely artificial fluorescent tubes). The results of our method are shown in Figure 10.

We can observe a mild degradation of performance with respect to our validation study. The correction still achieved significant improvement in color and once again never performed worse than the baseline ($p < 0.001$ for both matching and non-matching conditions). We believe that our correction is quite robust and should handle various natural environments well while naturally being slightly illumination-dependent.

5.5 Application to 3D Printing

Lastly, we present an application of our correction in the context of 3D printing. 3D printers require supporting structures to produce overhangs. These support structures often affect the finish of the print. As such, a simple sphere printed in a glossy finish will be manufactured with a half-glossy and half-matte finish. While such a variance in gloss might be acceptable for rapid prototyping, the resulting variance in perceived color would require manual post-processing samples for a consistent finish. We demonstrate that our correction can help to maintain color consistency. We manufactured 5 spheres with half matte and half glossy finish. For the reference spheres, we used the same color on both halves. For the corrected spheres, we applied our correction on the glossy side. We then asked the participants which of the spheres was perceived as more uniform.

The participants in each scenario considered the color-compensated spheres as more uniform (Figure 7, $p<0.001$). We showcase an example of this application in Figure 1 where our compensated half has an actual color difference of $4.047\Delta E00$ with respect to the original half. In the same vein, if spatially-varying gloss is desired [Piovarči et al. 2020], we can ensure that color will be consistent with the digital design, regardless of the surface finish.

6 LIMITATIONS AND FUTURE WORK

The method presented here computes a reference perceived color based on several assumptions. While we show that the final method works well and is reasonably robust, a more in-depth evaluation of the individual factors merits further study.

Our method estimates the mean perceived color from a single view. We did not evaluate how viewpoint selection impacts the correction. However, there are parallels between viewpoint selection and illumination changes as both change the distribution of colors on the surface of the reference geometry. Therefore, the performance of our correction from different view locations should loosely follow our evaluation using different environment maps.

For the estimation of the reference color, we use a neutral *car* geometry commonly used in the automotive industry. Our ablation studies suggest the method generalizes well to other geometries. However, a notable exception to this would be geometries with strong view-dependent effects. The simplest example is a flat plane where the visibility of the specular reflection strongly depends on the viewing direction. An interesting direction of future work would be to investigate if observers can compensate for these strong view-dependent effects.

Our method currently needs to be re-computed for specific color-gloss pairs and geometries. However, our ablation studies suggest

the method generalizes well to other geometries, potentially enabling a more widespread application of our method by precomputing large numbers of color-gloss corrections and deploying them as LUTs during print time.

During our optimization, we disregard the gamut of the output device when computing the correction. This can lead to generating corrections that lie outside of the gamut of the output device. In our studies, it is most visible when compensating towards a glossier finish which generally increases saturation. An interesting direction of future work would be to integrate the gamut of the output device as a constraint in the optimization.

Lastly, we use grayscale environment maps to illuminate our geometry. We chose a grayscale illuminant based on color constancy [Arend and Reeves 1986]. Our results suggest that the effect of different illuminations is small. However, further investigation whether the coupled color-gloss perception follows the same color constancy rules would be needed.

7 CONCLUSIONS

We present a perceptually-driven, differentiable rendering-based color correction approach that maintains a consistent perceived color on a surface regardless of the gloss level. To this end, we conduct a novel psychophysical experiment , demonstrating that the HVS estimates perceived color for a complex object under complex illumination by integrating specific locations that are more saturated and stable under view-dependent effects. Based on this insight we design a novel color correction method that models how subjects aggregate color from surfaces given varying gloss levels, estimating and correcting the perceived color shift. We show that our correction outperforms the state-of-the-art, and demonstrate the first practical application of such correction to color 3D printing.

ACKNOWLEDGMENTS

We thank Matthew S Zurawski for the 3D model of the *car* speed shape, Creative Tools[3] and Stanford Computer Graphics Laboratory for the geometry of the *bunny*, and Havran et al. [2016] for the geometry of the *ghost*. This research has been supported by the Swiss National Science Foundation (SNSF, Grant 200502), the FWF Lise Meitner (Grant M 3319), and an academic gift from Meta.

REFERENCES

William De Wiveleslie Abney. 1909. On the change in hue of spectrum colours by dilution with white light. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 83, 560 (1909), 120–127. https://doi.org/10.1098/rspa.1909.0085 arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1909.0085

Barton L Anderson. 2011. Visual perception of materials and surfaces. *Current biology* 21, 24 (2011), R978–R983.

Lawrence Arend and Adam Reeves. 1986. Simultaneous color constancy. *J. Opt. Soc. Am. A* 3, 10 (Oct 1986), 1743–1751. https://doi.org/10.1364/JOSAA.3.001743

Teun Baar, Maria V Ortiz Segovia, and Hans Brettel. 2014. Colour management of prints with variant gloss. In *Color and Imaging Conference*, Vol. 2014. Society for Imaging Science and Technology, 48–52.

Teun Baar, Sepideh Samadzadegan, Philipp Urban, and Maria V. Ortiz Segovia. 2016. Interrelation between gloss and texture perception of 2.5D-printed surfaces. *electronic imaging* 2016 (2016), 1–6.

Vahid Babaei, Kiril Vidimče, Michael Foshey, Alexandre Kaspar, Piotr Didyk, and Wojciech Matusik. 2017. Color Contoning for 3D Printing. *ACM Trans. Graph.* 36, 4, Article 124 (jul 2017), 15 pages. https://doi.org/10.1145/3072959.3073605

[3]https://www.thingiverse.com/creativetools/designs

Alan Brunton, Can Ates Arikan, Tejas Madan Tanksale, and Philipp Urban. 2018. 3D printing spatially varying color and translucency. *ACM Transactions on Graphics (TOG)* 37, 4 (2018), 1–13.

Alice C Chadwick and RW Kentridge. 2015. The perception of gloss: A review. *Vision research* 109 (2015), 221–235.

li Changjun, Zhiqiang Li, Zhifeng Wang, Yang Xu, Ming Luo, Guihua Cui, Manuel Melgosa, Michael Brill, and M. Pointer. 2017. Comprehensive color solutions: CAM16, CAT16, and CAM16-UCS. *Color Research & Application* 42 (06 2017). https://doi.org/10.1002/col.22131

CIE. 2001. *Improvement to industrial color difference evaluation.* Technical Report. Vienna: CIE Publication No. 142, Central Bureau of the CIE.

R. L. Cook and K. E. Torrance. 1982. A Reflectance Model for Computer Graphics. *ACM Trans. Graph.* 1, 1 (jan 1982), 7–24. https://doi.org/10.1145/357290.357293

Edul N. Dalal and Kristen M. Natale–Hoffman. 1999. The effect of gloss on color. *Color Research & Application* 24, 5 (1999), 369–376.

Datacolor. 2022. Color Theory - Visual and Instrumental Evaluation. https://www.datacolor.com/wp-content/uploads/2022/05/Datacolor_ColorTheory_Webianr_Part5.pdf. (2022). Accessed: 2023-01-22.

Oskar Elek, Denis Sumin, Ran Zhang, Tim Weyrich, Karol Myszkowski, Bernd Bickel, Alexander Wilkie, and Jaroslav Křivánek. 2017. Scattering-Aware Texture Reproduction for 3D Printing. *ACM Trans. Graph.* 36, 6, Article 241 (nov 2017), 15 pages. https://doi.org/10.1145/3130800.3130890

Willemijn Elkhuizen, Tessa Essers, Yu Song, Jo Geraedts, Clemens Weijkamp, Joris Dik, and Sylvia Pont. 2019. Gloss, color, and topography scanning for reproducing a Painting's appearance using 3D printing. *Journal on Computing and Cultural Heritage (JOCCH)* 12, 4 (2019), 1–22.

Dean Farnsworth. 1943. The Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color Vision*. *J. Opt. Soc. Am.* 33, 10 (Oct 1943), 568–578. https://doi.org/10.1364/JOSA.33.000568

Roland W Fleming. 2014. Visual perception of materials and their properties. *Vision research* 94 (2014), 62–75.

Roland W Fleming, Ron O Dror, and Edward H Adelson. 2003. Real-world illumination and the perception of surface reflectance properties. *Journal of vision* 3, 5 (2003), 3–3.

Martin Giesel and Karl R. Gegenfurtner. 2010. Color appearance of real objects varying in material, hue, and shape. *Journal of Vision* 10, 9 (09 2010), 10–10. https://doi.org/10.1167/10.9.10 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/933478/jov-10-9-10.pdf

Jeroen J. M. Granzier, Romain Vergne, and Karl R. Gegenfurtner. 2014. The effects of surface gloss and roughness on color constancy for real 3-D objects. *Journal of Vision* 14, 2 (02 2014), 16–16. https://doi.org/10.1167/14.2.16 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/932814/i1534-7362-14-2-16.pdf

Vlastimil Havran, Jiri Filip, and Karol Myszkowski. 2016. Perceptually motivated BRDF comparison using single image. In *Computer graphics forum*, Vol. 35. Wiley Online Library, 1–12.

Yun-Xian Ho, Michael S Landy, and Laurence T Maloney. 2008. Conjoint measurement of gloss and surface texture. *Psychological science* 19, 2 (2008), 196–204.

Richard S. Hunter. 1958. Photoelectric Color Difference Meter. *J. Opt. Soc. Am.* 48, 12 (Dec 1958), 985–995. https://doi.org/10.1364/JOSA.48.000985

Fabian Hutmacher. 2019. Why is there so much more research on vision than on any other sensory modality? *Frontiers in psychology* 10 (2019), 2246.

S. Ishihara. 1917. Tests for Colour-Blindness. *Handaya - Tokyo, Hongo Harukicho* (1917).

Wenzel Jakob. 2010. Mitsuba documentation. https://www.mitsuba-renderer.org/releases/current/documentation.pdf.

Wenzel Jakob, Sébastien Speierer, Nicolas Roussel, Merlin Nimier-David, Delio Vicini, Tizian Zeltner, Baptiste Nicolet, Miguel Crespo, Vincent Leroy, and Ziyi Zhang. 2022. *Mitsuba 3 renderer.* https://mitsuba-renderer.org.

Juno Kim, Phillip J Marlow, and Barton L Anderson. 2012. The dark side of gloss. *Nature neuroscience* 15, 11 (2012), 1590–1595.

Jian Ma, Haisong Xu, M. Ronnier Luo, and Guihua Cui. 2009. Color appearance and visual measurements for color samples with gloss effect. *Chin. Opt. Lett.* 7, 9 (Sep 2009), 869–872. https://opg.optica.org/col/abstract.cfm?URI=col-7-9-869

Phillip J Marlow, Juno Kim, and Barton L Anderson. 2012. The perception and misperception of specular surface reflectance. *Current Biology* 22, 20 (2012), 1909–1913.

Nathan Moroney, Mark Fairchild, Robert Hunt, li Changjun, Ming Luo, and Todd Newman. 2002. The CIECAM02 color appearance model. *IS & T/SID Color Imaging Conf* 10, 23–27.

Merlin Nimier-David, Delio Vicini, Tizian Zeltner, and Wenzel Jakob. 2019. Mitsuba 2: A Retargetable Forward and Inverse Renderer. *ACM Trans. Graph.* 38, 6, Article 203 (nov 2019), 17 pages. https://doi.org/10.1145/3355089.3356498

Gaë Obein, Kenneth Knoblauch, and Françise Viéot. 2004. Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision* 4, 9 (08 2004), 4–4. https://doi.org/10.1167/4.9.4 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/933508/jov-4-9-4.pdf

Fabio Pellacini, James A. Ferwerda, and Donald P. Greenberg. 2000. Toward a Psychophysically-Based Light Reflection Model for Image Synthesis. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '00)*. ACM Press/Addison-Wesley Publishing Co., USA, 55–64. https://doi.org/10.1145/344779.344812

Thiago Pereira and Szymon Rusinkiewicz. 2012. Gamut mapping spatially varying reflectance with an improved BRDF similarity metric. In *Computer graphics forum*, Vol. 31. Wiley Online Library, 1557–1566.

Michal Piovarči, Michael Foshey, Vahid Babaei, Szymon Rusinkiewicz, Wojciech Matusik, and Piotr Didyk. 2020. Towards Spatially Varying Gloss Reproduction for 3D Printing. *ACM Trans. Graph.* 39, 6, Article 206 (nov 2020), 13 pages. https://doi.org/10.1145/3414685.3417850

Sepideh Samadzadegan, Jana Blahová, and Philipp Urban. 2014. Color-Printed Gloss: relating measurements to perception. In *Color and Imaging Conference*, Vol. 2014. Society for Imaging Science and Technology, 207–211.

Matthias Scheller Lichtenauer, Philipp Schuetz, and Peter Zolliker. 2013. Interaction improves perception of gloss. *Journal of Vision* 13, 14 (12 2013), 14–14. https://doi.org/10.1167/13.14.14 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/933540/i1534-7362-13-14-14.pdf

Ana Serrano, Bin Chen, Chao Wang, Michal Piovarči, Hans-Peter Seidel, Piotr Didyk, and Karol Myszkowski. 2021. The effect of shape and illumination on material perception: model and applications. *ACM Transactions on Graphics (TOG)* 40, 4 (2021), 1–16.

Gaurav Sharma, Wencheng Wu, and Edul N Dalal. 2005. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur* 30, 1 (2005), 21–30.

Denis Sumin, Tobias Rittig, Vahid Babaei, Tim Weyrich, Thomas Nindel, Piotr Didyk, Bernd Bickel, Jaroslav Křivánek, Alexander Wilkie, and Karol Myszkowski. 2019. Geometry-Aware Scattering Compensation for 3D Printing. *ACM Trans. Graph. (Proc. SIGGRAPH)* (2019). https://doi.org/10.1145/3306346.3322992

Shoji Sunaga and Yukio Yamashita. 2007. Global color impressions of multicolored textured patterns with equal unique hue elements. *Color Research & Application* 32 (08 2007), 267 – 277. https://doi.org/10.1002/col.20330

TS Trowbridge and Karl P Reitz. 1975. Average irregularity representation of a rough surface for ray reflection. *JOSA* 65, 5 (1975), 531–536.

Philipp Urban, Tejas Madan Tanksale, Alan Brunton, Bui Minh Vu, and Shigeki Nakauchi. 2019. Redefining a in rgba: Towards a standard for graphical 3d printing. *ACM Transactions on Graphics (TOG)* 38, 3 (2019), 1–14.

Gunter Wyszecki and Walter Stanley Stiles. 1982. *Color science*. Vol. 8. Wiley New York.

Bei Xiao and David Brainard. 2008. Surface gloss and color perception of 3D objects. *Visual neuroscience* 25 (05 2008), 371–85. https://doi.org/10.1017/S0952523808080267

X–Rite. 2005. Complete Guide to Color Management. https://xritephoto.com/documents/literature/en/L11-176_Guide_to_CM_en.pdf. (2005). Accessed: 2023-01-22.

X–Rite. 2016. A Guide To Understanding Color. https://www.xrite.com/-/media/xrite/files/whitepaper_pdfs/l10-001_a_guide_to_understanding_color_communication/l10-001_understand_color_en.pdf. (2016). Accessed: 2023-01-22.

Fan Zhang, Huib de Ridder, Pascal Barla, and Sylvia Pont. 2019. A systematic approach to testing and predicting light-material interactions. *Journal of Vision* 19, 4 (2019), 11–11.

Fan Zhang, Huib de Ridder, Pascal Barla, and Sylvia Pont. 2020. Effects of light map orientation and shape on the visual perception of canonical materials. *Journal of vision* 20, 4 (2020), 13–13.

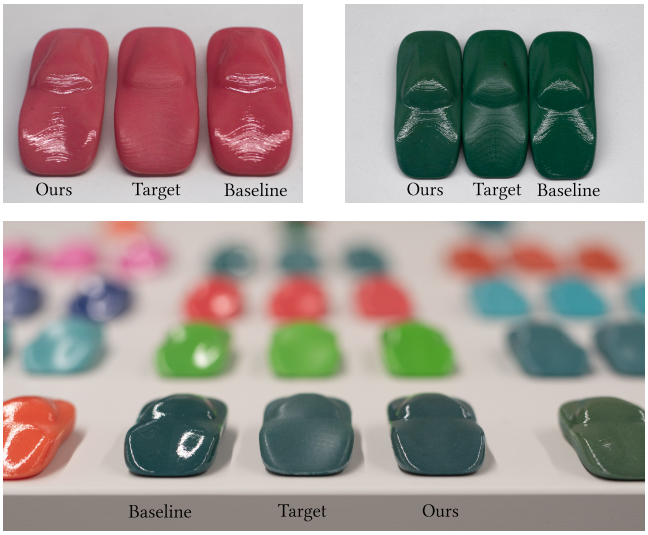

Figure 9: A small showcase of the color triads presented to test subjects during our experiment in different viewing angles and varnishes (top-left and bottom, *Schmincke Matte*, top-right, *Edding Clear Matte*). Target denotes the target color appearance, while baseline is the same base color on the target surface finish. Ours achieves a closer perceived color to the target, while maintaining the desired glossy surface appearance. Note that due to compression and SDR screen quality, color differences may look substantially smaller than in real life.

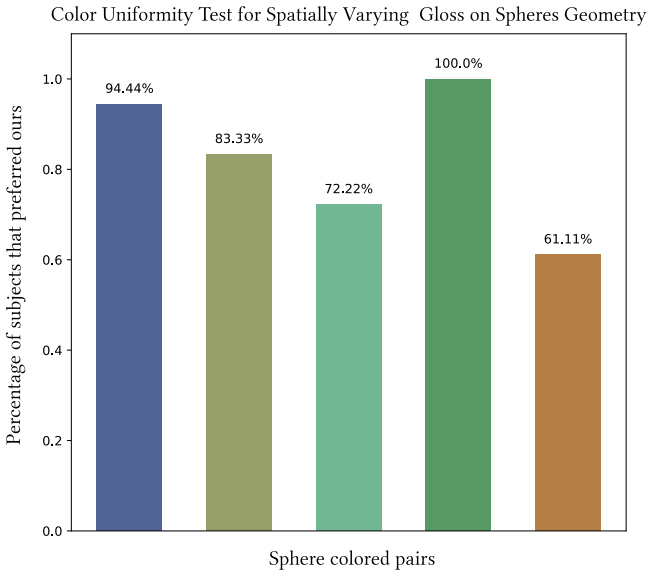

Figure 7: Results for our experiment with spatially varying gloss, broken down for individual color pairs. On average, we achieve an 82.22% of preference for our compensations, meaning that our compensated spheres had better color uniformity across different surface finishes.

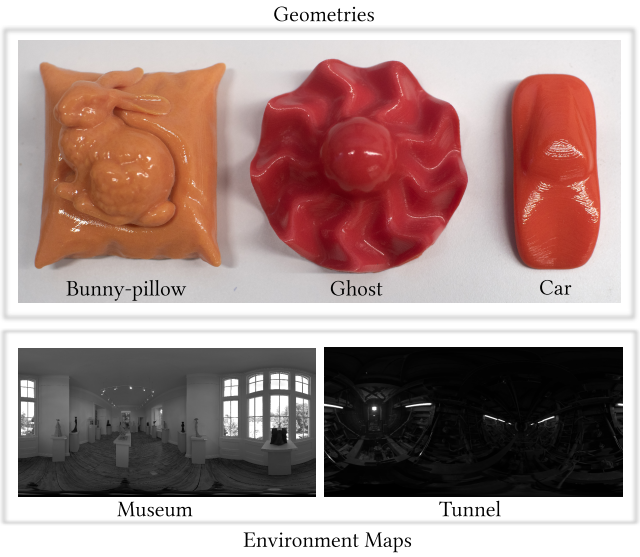

Figure 8: Showcase of the different geometries and real captured environment maps used throughout our ablation studies.

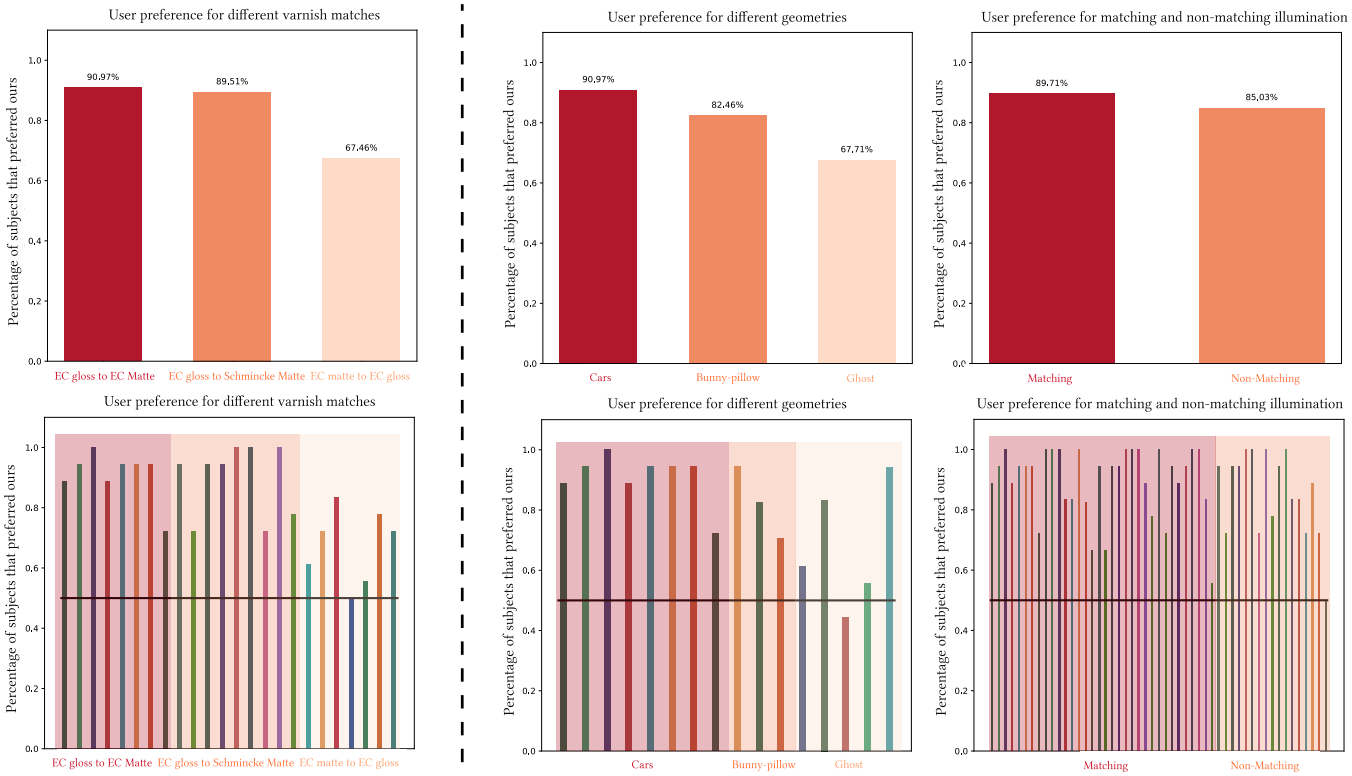

Figure 10: Results for our experiment with 3D printed samples. In the top row, we show averages across different colors for validation and ablation studies. In the bottom row, we show corresponding plots for its top counterparts, which break down averages into all individual tested colors, each the average response of the 18 test subjects. On the left, we have the validation study, while on the right we have our two ablations: on geometry (varying geometries while keeping compensation computation with *cars* geometry (centre)) and on environment illumination (using different environment maps during computation and testing the different proposed compensated colors within a loosely matching illumination or a largely different one (right)). A 50% preference (marked by the black line) means that while we don't improve over the baseline color, our proposed compensation does not worsen perceived color. As you can see, our method routinely achieves to reduce the perceived gap between different finishes.