Redacted: Published January ’18

Anonymous

Abstract

In this work we present a psychometric, visual search-based study analyzing the perceptual appearance uniformity of 3D printed materials. A 3D printer’s quality, precision, and capacity to produce smooth surfaces directly affects the perceived uniformity of its outputs. This work represents the first steps towards building a perceptual model of uniformity for 3D printing. Such a model will greatly assist in advancing the quality of 3D printers, especially as they become capable of creating complex, spatially-varying appearances.

We demonstrate the effectiveness of applying visual search to appearance perception problems by analyzing 288 appearance variations formed from the combination of 18 printed surfaces, 8 virtual transformations of those surfaces, and two illumination conditions. The virtual transformations allowed us to explore the impact of bumpiness, glossiness, and spatially-varying color on perceived uniformity. Significant effects were found to be caused by several of these dimensions. Additionally, the measured psychophysical data is a valuable contribution to the general study of the perception of spatially-varying appearances.

Introduction

Appearance uniformity is a perceptual measure of a material’s apparent spatial homogeneity, and it can depend on how the material is illuminated and from where it is viewed. Often in product design the goal is to achieve uniformity for individual components, such as a plastic panel or a casing. Uniformity can intentionally be reduced by adding patterns and texture, but it can also be diminished naturally by scratches, scrapes, and other forms of wear-and-tear. It is possible to define material uniformity in terms of physical attributes, like the measured height profile or color distribution of the surface. However, in this work we approach uniformity as a perceptual quantity, acknowledging that certain conditions can alleviate or expose non-uniformities in a material. Given the limits of our visual system, some non-uniformities captured by a physical measurement may not be visibly noticeable or significant. Focusing on perceived appearance uniformity has advantages, because, compared to physical measurements, it represents a less stringent threshold for quality control and evaluation.

Uniformity is of particular interest to us because of how it plays a role in developing quality 3D printing devices. However, we contend that uniformity is a useful benchmark for any printing task, whether it is in two or three dimensions. Outputting a smooth surface with homogeneous color requires the device to have high precision and consistency. The printing of colored patterns or surface texture can in fact mask issues concerning the quality of the output.

While most consumer color paper printers are able to achieve satisfiable levels of uniformity across a page, 3D printing is in the unique position of controlling the third dimension. Control over both macroscopic surface curves and mesoscale bumps and texture present new challenges in measuring and quantifying effects on appearance uniformity. Current 3D printers have not yet reached a comparable quality level to their 2D counterparts, and frequently introduce unintentional non-uniformities. This work seeks to quantify the perceptual dissimilarities in color, surface, and other appearance non-uniformities that can occur in 3D printing. While there are tools available to measure specific physical quantities related to printed appearance, there is not a perceptual metric that combines the effects of spatially-varying color, texture, and gloss.

Figure 1. *Comparison between a printed appearance (left)–highlighting its non-uniformities–and the computed ideal uniform (right).*

Thus, we sought to collect psychometric evaluations of a representative sampling of printed materials. We define appearance uniformity as the perceptually linear dissimilarity between a material’s appearance and that of a constructed, ideally uniform appearance. A constructed ideal would match the material’s intended color, glossiness, and other appearance behaviors but lack the spatial variations across these attributes that produced visible non-uniformities. An example of a printed appearance and its computed ideal are shown in Figure 1. As the dissimilarity tends towards zero, the original material approaches full uniformity. Defined in this manner, appearance uniformity becomes a specific instance of the more general appearance similarity problem.

Even though a computational model for the general similarity problem does not exist yet, quantifying the perceived appearance uniformity has multiple immediate applications to the field of 3D printing. Uniformity measurements can provide guidance to non-appearance experts who are otherwise advancing 3D print technology. Comparing measurements across multiple printed units can be used for quality control and quantifying the reproducibility of a 3D printer. Currently, it is common for 3D printed objects to feature a single color, or disparate regions of solid colors. Much as color science helped ensure the best color match between a photo on a monitor and its printed result, appearance uniformity can be used to match a digital design to the solid, colored output.

Our work is a first step towards addressing this problem, so while we do not develop a predictive model for uniformity in this paper, we demonstrate how to efficiently measure uniformity and we are able to report observed qualitative effects on printed uniformity. We focus instead on collecting a large set of perceptual judgments to learn more about this subset of the general appearance space.

The visual search paradigm was employed to efficiently and directly measure perceived dissimilarities across a wide range of 3D printed appearances. The task presents a target appearance amongst a field of distracting appearances, and the time to find the target indicates the similarity between the target and distractors. This avoids the need for describing perceptual scales to the subjects or performing forced-choice comparisons for all combinations of appearances. This efficiency enabled us to collect perceived dissimilarities for 288 appearances that varied across color, pattern, glossiness, bumpy texture, and lighting environment. We made use of X-Rite's TAC7 material scanner[1] to digitize accurate representations of printed surfaces. We then constructed appropriate ideals and made additional variations with controlled appearance transformations. To the best of our knowledge, this represents the first application of the visual search task to the perception of appearance, and it demonstrates an exciting combination of psychophysics, 3D printing, and material scanning.

In the next section we present background literature across the range of topics that intersect this work. Following that, we describe both the physical prints and virtual appearance stimuli used in our psychophysical study, and we provide the details of our visual search study. We conclude with analysis of the collected measurements, demonstrating that a consistent perceptual response was provided by all subjects, and we discuss some of the more obvious variable effects and what this implies for 3D printing appearance in the future.

Background

This work lies at the intersection of a number of fields. While it is motivated by additive manufacturing and 3D printing, the focus on appearance requires color science, computer graphics (the study of material modeling and simulation), and psychophysics (understanding the visual system and human subjects research). All of these topics are introduced in the following sections.

3D printing

A range of additive manufacturing techniques have been developed over the last few decades, such as fused deposition modeling (*FDM*), selective laser sintering (*SLS*), and most recently multi-jet fusion (*MJF*). Additional approaches have been developed that work with metals and ceramics, broadening the range of materials usable with 3D printing. 3D printing has already revolutionized industrial part manufacturing and prototyping, where the emphasis is on function, durability, and fast turnaround. It has also become prominent with hobbyists and artists, and will eventually be useful for consumer products. In these domains, the appearance of the 3D printed object becomes paramount.

To that end, researchers have been advancing the appearance capabilities of 3D printers. A relatively simple extension to FDM printers is to make it two-toned [12, 24], allowing a printed object to have separately colored regions or even intermediate colors via half-toning. Full color 3D printing, where each voxel element can independently be colored, further improves upon the realism of the printed object [4], although this is currently limited to diffuse color only.

As 3D printing is used for more professional, consumer or appearance-oriented products, it will be necessary to improve print quality. Combining knowledge of human perception with additive manufacturing has led to noticeable improvements in visual quality. Stair-stepping, or the visible layering of material planes, in the object is a common objectionable defect. Wang's work [29] uses a perceptual saliency model to control the resolution of the printer to minimize the stair steps in complex areas while still minimizing the overall print time. It is also possible to separate an object into multiple parts and orient them so that each parts' local structure better aligns with the printer's coordinate system [28], which can improve surface quality and smoothness.

Eventually printers will have the capability to print materials that exhibit a broad range of reflectance profiles. The future of 3D printing represents a way to print and manufacture full spatially-varying materials on arbitrary and curved surfaces. Reaching that point requires a thorough understanding of how to model physical materials that artists, designers, and engineers would like to print.

Material modeling

An opaque material can reasonably be described with a spatially-varying bidirectional reflectance distribution function (*SVBRDF*) [16]. An SVBRDF is a two dimensional map of parameters to a homogeneous BRDF, of which there are numerous definitions that seek to describe how light is reflected by a material [2, 5, 30]. At a high level, an SVBRDF describes the bumpiness, color, and roughness over the surface of an object. However, these parameters exist in an arbitrary or physically-aligned space, which makes it difficult for artists to pick values that simulate realistic materials.

Material scanning helps address this difficulty, and in many ways is the inverse of 3D printing: it creates a digital model of a physical object. Scanning is directly relevant to the future of 3D printing. Scanned materials will form libraries from which designers can make selections. Printed objects can subsequently be scanned to evaluate how accurate the print is compared to the original digital model. An object manufactured by some other process can be scanned so that multiple copies can later be printed.

Measured BRDFs have been acquired for homogeneous materials [11, 15], where it is assumed that the entire surface can be represented by a single function. Spatially-varying materials can be scanned with a myriad of techniques that combine multiple calibrated lights and cameras to solve for BRDF parameters. Examples of this include spatial gonioreflectometry [16], linear light scanning [10], and spherical harmonic scanning [26]. More recent work has focused on simplifying the scanning process by using a handheld mobile device [7, 25]. X-Rite, has built the TAC7 Scanner (Total Appearance Capture) that integrates much of the research into an easy-to-use black box device. We have made use of the TAC7 to help prepare the stimuli used in our experiments.

Appearance perception

Material scanning captures and records a physical description of a material and object, while 3D printing exports a phys-

[1] https://www.xrite.com/categories/appearance/tac7

ical description. There is a perceptual disconnect between the scanned model, printed object, and what a person perceives in both cases. The field of color science has studied a similar problem for diffuse color on two dimensional surfaces, mapping from the physical spectrum to perceptual dimensions [34]. Appearance and its perception involves much more than just an object's color, combining attributes such as glossiness [21, 31], lighting environment [9], and shape [27]. For the most part, these have been studied as independent effects but other research has looked into how they interact: gloss and color [6]; gloss and texture [14, 20]; and brightness and texture [13].

Appearance perception can also be studied in terms of a distance or similarity metric between two materials. While there are metrics designed for homogeneous BRDFs [22], there is currently no such general perceptual model for the range of appearances representable by 3D printing. Once such a metric is developed, it can be used in quality control for 3D printing, for design and prediction in the printing pipeline, and as a tool to be used by engineers developing printers who themselves may not be appearance experts. This work does not seek to provide such a model, and exploring appearance uniformity is a simpler subset of the problem of general appearance similarity. Our work does represent initial steps towards a solution and presents results that are immediately valuable for evaluating current 3D print quality.

Visual search

Without an appearance model that could be applied directly to the problem of measuring appearance uniformity for 3D printing, we sought to directly measure human judgments of perceptual dissimilarity between a non-uniform appearance and its corresponding idealized, uniform representation. The visual search [8, 32] study paradigm is an effective and rapid way of measuring perceptual dissimilarity. The study task presents a target and a field of distractors to the subject and measures their reaction time to find the target. The target and the distractors can vary along a single perceived dimension, or across multiple dimensions [17], such as size, contrast, or frequency. The premise is that the more similar the target is to the distractors, the longer it takes for the subject to identify it. Arun [1] showed that the reciprocal of response time is actually a better value to use as a measure of dissimilarity. While visual search has been widely used in the psychophysics community to learn more about the visual system using comparably primitive stimuli, this work is to our knowledge the first that applies visual search to appearance similarity.

Experiment

We define the appearance uniformity of a printed material as inversely related to the perceptual dissimilarity to a corresponding idealized, uniform material. Thus, appearance uniformity is mapped to a perceptual similarity problem. As the perceived distance between the material's appearance and its ideal approaches zero, it becomes more and more uniform. The only caveat is that for a given material, its uniform, ideal variant must be defined. Often this will be the desired, digital model that was printed but it can also be defined after printing by using a material scanner and filtering or smoothing the reflectance model parameter maps produced by the device. This is useful when characterizing a 3D printer or if a printer is not capable of producing an ideal material sample.

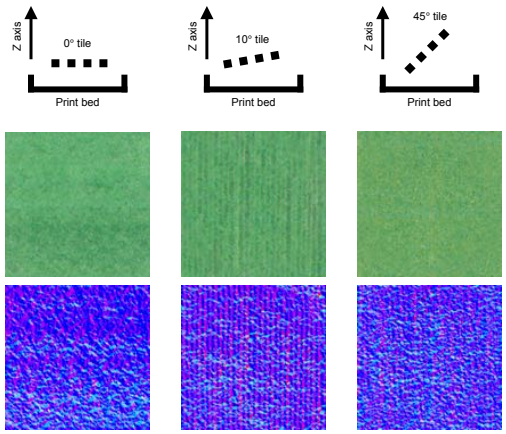

Figure 2. *Tile orientation and example SVBRDF maps for diffuse color and surface normal (in pseudo color).*

For the purposes of this experiment we consider three broad appearance attributes and how they can independently deviate from a uniform ideal:

- *bumpiness* – the just-visible, mesoscale height profile of the surface, which deviates from a perfectly smooth surface.
- *color* – spatial changes in the diffuse scattering of the material, such as hue or lightness, deviating from a solid color.
- *glossiness* – changes to the microscopic roughness that affects specular scattering of the material.

These attributes represent the primary sources of appearance variation with the 3D printer technology and material substrate available to us. We will explore how modifications to the spatial distributions of these attributes in printed materials impact appearance uniformity judgments. The creation of the appearance images used for stimuli in our experiments are described in the next section. The psychometric task to collect the appearance dissimilarity judgments of the 3D printed materials is specified in the subsequent section.

Stimuli design

To take advantage of computer-driven study protocols, the stimuli images used in this experiment are simulated images. The materials are based on BRDF scans of actual 3D printed surfaces, along with digital modifications to create a controlled range of parameter changes. To get realistic base material definitions, we printed 1cm $\times$ 1cm tiles arranged at different orientations with respect to the Z-axis of the printer. While more orientations were tested, tiles oriented at 0°, 10°, and 45° were found to span the range of patterns and non-uniformities formed by the 3D printer that was used. Figure 2 shows the print configuration for these three angles and example SVBRDF maps corresponding to the print results.

Tiles were printed in several colors–cyan, green, and purple–to explore how the intended color affects perceived uniformity and to include non-uniformities caused by the inking process. Additional physical variants were created by post-processing the printed tiles to smooth and coat them. The post-processing is one example of finishing processes that might be applied to a

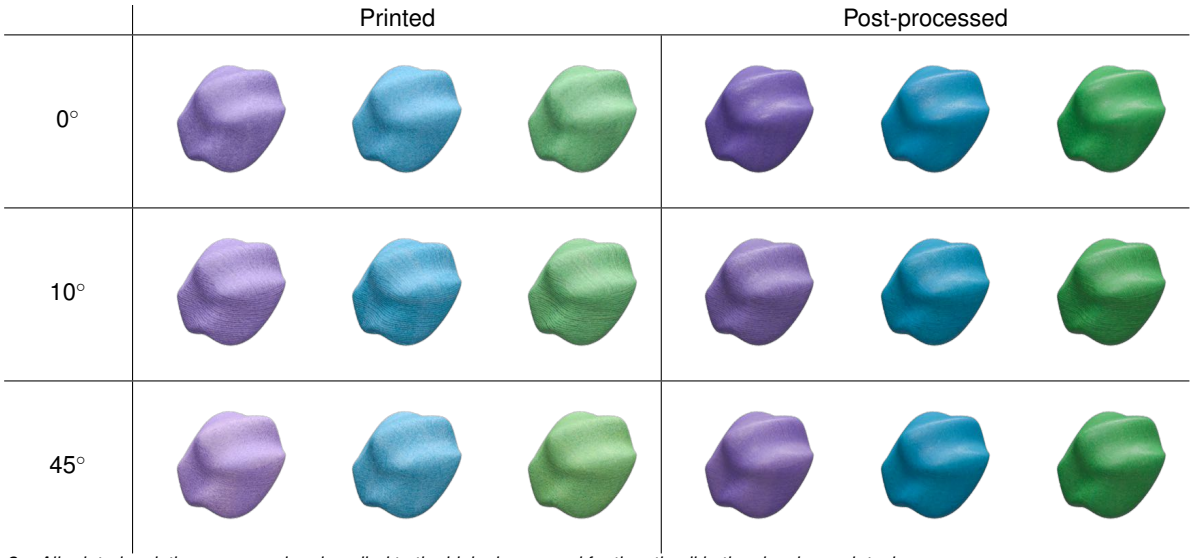

Figure 3. *All printed variations, scanned and applied to the blob shape used for the stimuli in the visual search task.*

3D print. This process can also cause substantial changes to the appearance of the object, such as reducing bumpiness, increasing glossiness, and reducing visible non-uniformities. Overall, 18 physical printed materials were made from the combination of three print orientations, three colors, and either including or omitting post-processing. Renderings from measurements of all of the printed variations are shown in Figure 3.

After printing the 18 tile variants, the samples were scanned using X-Rite's TAC7 device, producing SVBRDF parameter maps in the Ward [30] model for each. These SVBRDFs are what were used in the simulations shown in Figure 3, 1, and 4. From those images it is clear that non-uniformities, in both the diffuse color and the surface bumpiness, exist in the printed tiles.

Ideal uniform materials corresponding to the 18 tiles were created by filtering the diffuse color map and mesoscale height map. The ideal diffuse color map was calculated by reducing the Weber contrast by a factor of 10. Weber contrast is measured as the relative difference between a pixel's luminance and the background luminance, which we took as the median lightness of the diffuse color map. The ideal height profile was computed by scaling the scanned heights by 0.1. These transformations created SVBRDFs that are almost homogeneous, but still contain subtle non-uniformities that improve realism. The previously shown Figure 1 demonstrates a side-by-side comparison of a scanned material and its constructed ideal. It is clear that the ideal has improved uniformity but still exhibits an overall similar appearance to the original.

Having scanned the 18 materials with the TAC7 device, we used the SVBRDF measurements to create similar materials with known variations in the parameters. Using the same filters that were applied to create the ideal materials, seven additional materials were generated for each of the original scans. Non-uniformities were increased along the three appearance attributes of interest in different amounts and combinations: increasing diffuse contrast, bumpiness, and glossiness. Figure 4 describes the attribute modifications for the seven variations, with respect to an original, and shows an example for each. These variations preserve the stochastic and structural patterns that were produced by the 3D printer. It would not have been possible to use a current generation 3D printer to preserve such patterns while also accurately modifying the desired parameters. The varied materials for each original allows us to explore how the three appearance attributes interact and impact appearance uniformity.

While the original physical tiles were more-or-less two dimensional, all generated materials were rendered to a blob geometry that was originally designed to have a range of surface curvature, highlight the environment's lighting, and otherwise not resemble a real world object that might bias the subject [19]. This geometry is what's shown in Figures 1, 3, and 4. Although shape does influence appearance, we did not evaluate multiple geometries due to the already sizable parameter space being considered. However it has been shown in the literature that for realistic material visualizations blob like objects are preferable to simpler objects like spheres.

However, we did introduce an additional stimuli variable to consider the illumination conditions when viewing the material. For each of the materials, images were rendered in two different conditions. First, a diffuse, ambient lighting condition based on a D65 light booth intended to simulate the appearance of the object when placed under an actual booth (shown in Figure 5). When rendered, the D65 white point was adapted to be white. Second, a single spotlight source is used without any other ambient lighting, which represents a more extreme environment. The spotlight was also chromatically adapted to be the white point of the scene. All images were tonemapped to a low dynamic range. Figure 6 demonstrates the differences in appearance caused by the two lighting environments when applied to the same material.

To summarize, the combinations of three print orientations, three colors, and two post-processing options defined 18 physically printed materials. The 18 materials were scanned using X-Rite's TAC7 to get a Ward SVBRDF for each. Filtering the original SVBRDFs produced 18 corresponding ideal materials. Including the original scan, 8 SVBRDF variations were paired with each of the computed ideals. With two illumination con-

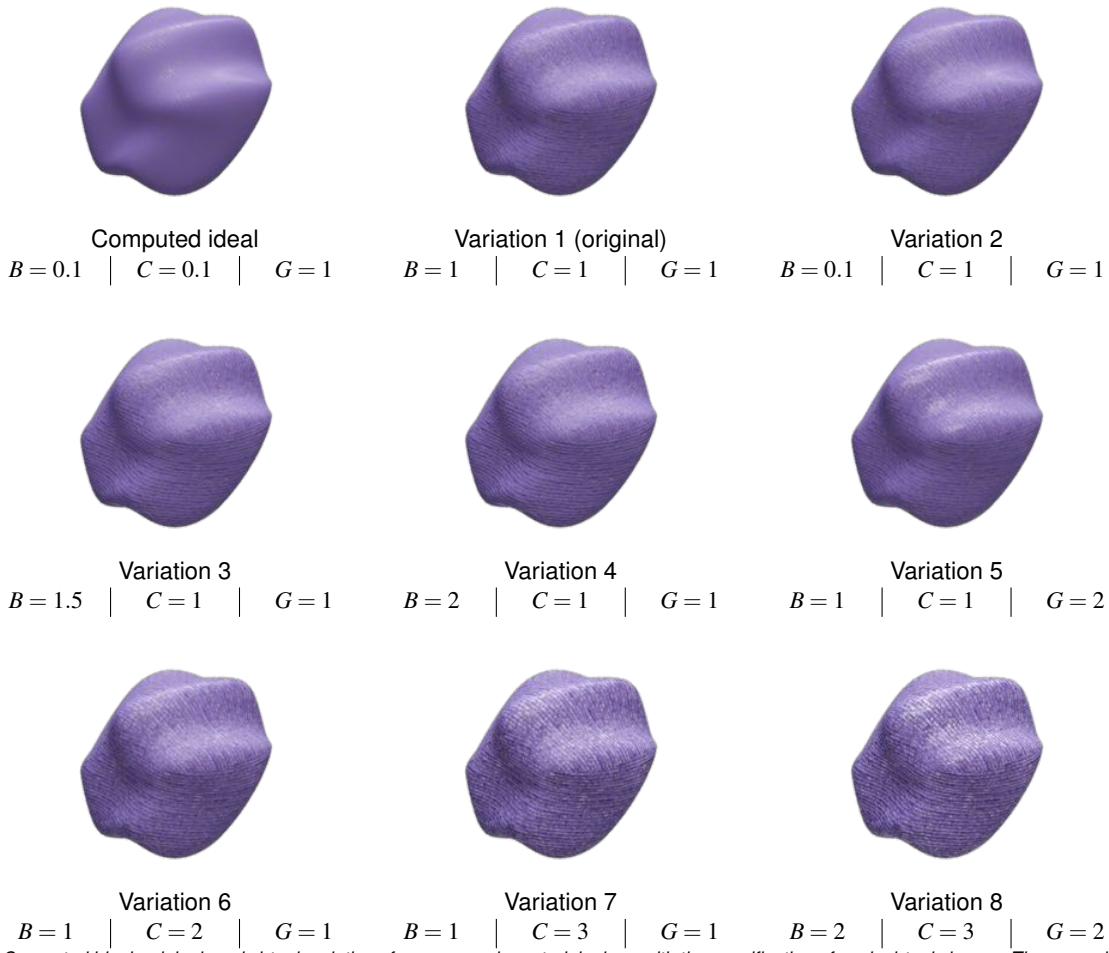

Figure 4. *Computed ideal, original, and virtual variations for an example material, along with the specification of each virtual change. The examples are based on the post-processed purple tile printed at 10°. The appearance attribute changes are specified by the values in* B *(bumpiness),* C *(color), and* G *(glossiness) fields.*

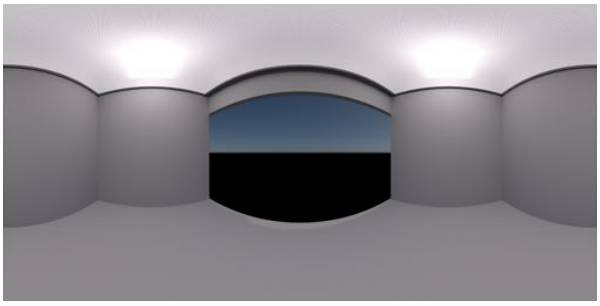

Figure 5. *Panoramic tone-mapped representation of the D65 light booth environment.*

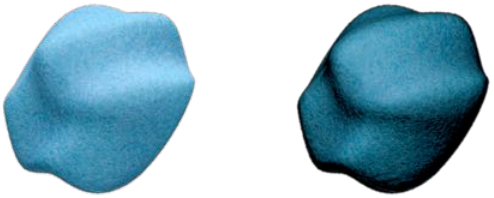

Figure 6. *Comparison between the two lighting conditions' effects on the appearance of a printed cyan material. Left: D65; right: spotlight.*

ditions, this produces 36 ideal appearance images and 288 non-uniform appearance images. While there are material and illumination changes, all images use the same geometry and are rendered from the same point of view. The psychophysics task to measure the perceived dissimilarity between every non-uniform appearance image and its corresponding ideal is described below.

Visual search task

We measured the appearance uniformity of the 288 non-uniform images by measuring the perceived dissimilarity to their matching ideal images. The perceived dissimilarity was measured using the visual search paradigm [8]. Arun's recent work [1] showed that the reciprocal of response time in a visual search

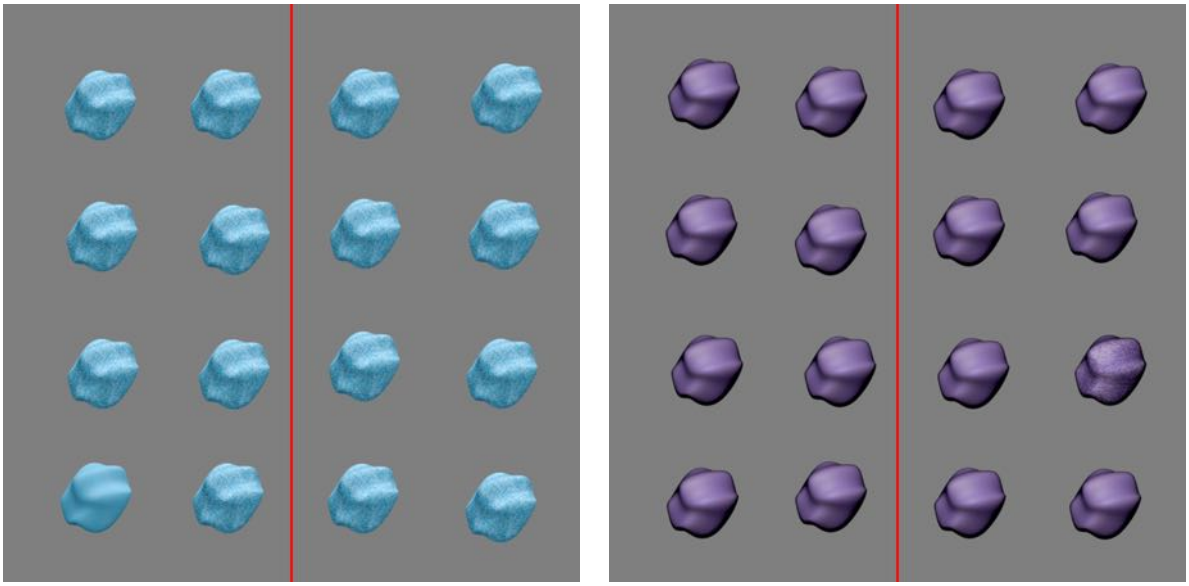

Figure 7. *Screenshot of the visual search task. The jittering changes with each presentation. The left example has a uniform target in the bottom left corner. The right example as a non-uniform target along the right edge.*

task behaves like the linear perceptual distance between the target and distractor items. Using a task inspired by Pramod et al.'s work with object silhouettes and visual search [23], we applied the search task to the appearance simulations described above. For our purposes, the target and distractor images are a specified non-uniform appearance image and its paired ideal image.

The visual search task consisted of multiple trials, each presenting a 4×4 grid of appearance images. The positions of the 16 images were jittered to prevent features in the rows or columns from aligning. One of the 16 images would be the target, in a randomized location in the grid, and the remaining 15 images would be the corresponding distractor image. For a given presentation of the jittered grid, the target image and distractor would be a non-uniform appearance and its ideal, or vice-versa. It has been shown in past visual search research that if a target A is shown amongst a distractor B, the reaction time does not necessarily equal that of searching for B in a field of A's. To determine if this was true for our appearance stimuli, each presentation randomly selected the ideal or non-uniform image as the target. Figure 7 shows an example screenshot of the visual search task, implemented using the PsychToolbox [3].

After the jittered grid was shown to a subject, they searched as quickly as possible for the target appearance image, described to them as the "odd one out". A vertical red line was rendered with the grid to divide the field into two halves. Once the target appearance image was found, the subject specified the side of the red line that contained the target. This was done using a keyboard press, Z for the left side and M for the right. Specifying only the side of the target's position kept the time cost of the input action as low as possible for the subject. While the subject still had to identify the target, if they were also expected to select the exact grid position via mouse, it would have been challenging to separate reaction time into the search time and selection time.

The non-uniform appearance images were shuffled independently for each subject. Additionally, each appearance image was repeated multiple times until the subject correctly identified the target three times. The repeated trials were inserted at random amongst the remaining trials. Picking the wrong side resulted in an audio tone being played and advancing to the next trial. If a subject failed to identify the target three times after five total presentations of that appearance, it was skipped so that the experiment could terminate (this did not occur in practice). Subjects were given up to 10 seconds to find the target amongst the distractors and make their decision. Reaching the time limit was treated the same as picking the incorrect side.

Study setup

Using the protocol described above, the task was performed in a dark room, with the interface presented on a 31.5 inch HP DreamColor Z32X UHD display. The display had a resolution of 3840×2160 pixels. The display was set to a factory calibration of the sRGB color space and all stimuli images were presented in sRGB. Each of the sixteen appearance images in a search field occupied an approximate viewing angle of $3.1°$.

In order to reduce the duration of the experiment, the 288 non-uniform appearance variations were split into three blocks based on the orientation of the original tiles in the printer. The orientation was used as a between-subjects variable, while the remaining variables of color, post-processing, lighting, and SVBRDF modifications were within-subject. 21 volunteers with normal or corrected-normal vision participated, with 7 subjects assigned to each of the $0°$, $10°$, and $45°$ blocks. 8 subjects were female, with the remaining 13 male; subjects' ages ranged from 18 to 59 years old. Given the between-subjects split, each subject provided their perceived dissimilarity scores for 96 of the non-uniform appearances. Subjects on average completed the entire experiment in under 40 minutes and were always able to reach three successful target identifications for their assigned 96 appearances.

As part of the training, a motor-speed measurement task was

Symbol	Definition
A	A non-uniform appearance image.
$\mathcal{A}$	The set of 288 appearances evaluated in this study.
$\mathrm{I}(A)$	The ideal appearance corresponding to the non-uniform appearance, A.
$\mathcal{S}(A)$	The set of subjects that evaluated the appearance, A.
$r_s^{(i)}$	The i^{th} measured reaction time for the subject's motor skills task.
$t_{A,s}^{(i)}$	The i^{th} measured response time from the subject, $s \in \mathcal{S}(A)$, for the appearance, A.
d	A perceived non-uniformity value, which can annotated similarly to t.

Figure 8. *Notation for the raw inputs and outputs of the visual search task.*

performed. Using the same controls as the actual visual search task, subjects would pick the side of the screen that a circle would randomly appear on. The perceptual aspect of such a task is minimal, so the response time is dominated by the subject's reaction time and the time to press the physical key. This can be used to calculate the search time for the actual visual search task by subtracting the subject's minimum circle response time from their appearance response times. The results of this collected data are presented in the next section.

Results

Using the notation defined in Figure 8, response times t were repeatedly collected for each appearance in $\mathcal{A}$ and its matching ideal in $\mathrm{I}(\mathcal{A})$. For a given appearance, A, and specific subject, $s \in \mathcal{S}(A)$, we write their i^{th} non-uniformity measurement for A as $d_{A,s}^{(i)}$. Each appearance in $\mathcal{A}$ was evaluated by seven subjects a total of three correct times, therefore $|\mathcal{S}(A)| = 7$ and $i \in \{1,2,3\}$. We formally defined the non-uniformity as the perceived distance or dissimilarity between A and $\mathrm{I}(A)$, i.e. $d = |A - \mathrm{I}(A)|$.

It is worth noting that as d approaches 0, the appearance becomes more uniform and a value of 0 implies that a printed material appears perceptually indistinguishable to its ideal. Based on the work of Arun [1] and Pramod et al. [23], we can calculate the non-uniformity measurement, d, with respect to a subject(s), appearance(A), and specific presentation(i):

$$\begin{aligned} d_{A,s}^{(i)} &= |A - \mathrm{I}(A)| \\ &= \frac{1}{t_{A,s}^{(i)} - r_s} \end{aligned} \tag{1}$$

where r_s is a subject's fastest reaction time:

$$r_s = \min_i r_s^{(i)}$$

We define a subject's aggregate measurement as:

$$d_{A,s} = \underset{i}{\text{median}}\, d_{A,s}^{(i)} \tag{2}$$

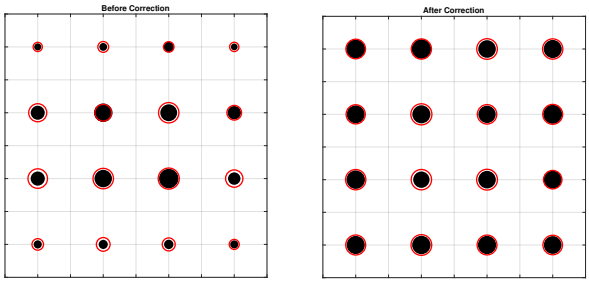

Figure 9. *The median, normalized dissimilarity score by position in the 4×4 visual search field. The solid black circle represents the case where the target was the uniform appearance; the red circle visualizes that the target was the non-uniform appearance image.*

and the final non-uniformity value for an appearance as:

$$d_A = \underset{s \in \mathcal{S}(A)}{\text{median}}\, d_{A,s} \tag{3}$$

Justifications for aggregating over a subject's repeated trials (Equation 2), and aggregating between subject's measurements (Equation 3) are presented in the next two subsections. Finally, main and interaction effects visible in all of the computed d_A scores are explored in the last subsection.

Eccentricity correction

Eccentricity, in this context, refers to the search target's distance from the center of the screen. Given that the target was in one of sixteen positions in a 4×4 grid, there were two levels of possible eccentricity. The target was either in the central four positions or in the outer twelve. It is reasonable to expect that it may take longer to find the target in the outer ring because there are more grid cells to consider and they are farther from the center focus point.

To better understand this, we plotted the median dissimilarity value d after linearly normalizing away the effects of subject, print angle, color, surface finish, material variation, and lighting environment. In effect, this is removing the impact of every directly controlled experimental variable. However, given how the study was set up, the target could either be the uniform ideal appearance or the non-uniform appearance. These two states were normalized and corrected for separately.

The median normalized dissimilarity per grid position is shown on the left of Figure 9. If the eccentricity had no impact on the response times then the size of the circles would be approximately the same. However, it is clear that there is a systemic decrease in the radius of the outer circles, for both target types. Since the size of the circle correlates to dissimilarity, the smaller circle implies a higher response time, matching our hypothesis.

For all subsequent analysis, we remove the eccentricity bias from trial responses in the outer ring by linearly shifting their normalized dissimilarities to match the inner ring. The corrected medians are shown in Figure 9 (right), demonstrating that the effects of position have substantially been reduced. After this shift, all normalized dissimilarities have the effects of the experimental variables restored by undoing the original linear normalizations. With the bias removed, the specific position of the target appearance is irrelevant for the subsequent analysis.

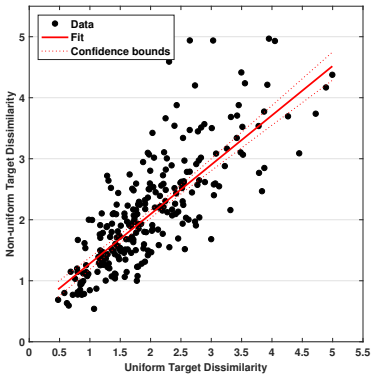

Figure 10. *Dissimilarity scores for appearances, where each axis corresponds to a different target condition. Each data point represents a subject's appearance judgment when viewing the appearance under both target conditions. The confidence bounds represent slopes for the 95% confidence interval.*

Reciprocity

Past visual search literature has documented that the measured response time does not always respect reciprocity, i.e. the response time for finding A amongst B does not equal that of finding B in A. For example it is easier to find a mirrored letter in a field of regularly oriented letters than it is to find the regular in the field of mirrored letters [33]. Using the notation introduced above, this would mean that $d = |A - \mathrm{I}(A)|$ does not necessarily equal $d' = |\mathrm{I}(A) - A|$. To determine if that was the case, subjects were given tasks that randomly chose between finding the ideal or the non-uniform appearance.

Figure 10 plots the dissimilarity scores for each appearance, where the X axis represents the appearance's non-uniformity when the target was the ideal appearance, and the Y axis is the appearance's non-uniformity when the target was the non-uniform appearance. Each black marker represents a subject's evaluation of an appearance under both conditions; if a subject only viewed an appearance in either target condition it was excluded from the plot. Barring the outliers above the linear fit, there is reasonable consistency in dissimilarity measurements for the two target conditions.

A robust fit yields a linear model, $d' = md + b$, between the two conditions, where $m = 0.8094$ and $b = 0.4734$ ($R^2 = 0.6528$). If the two target choices did not affect the response time, the model would be expected to have $m = 1$ and $b = 0$. The constant shift in dissimilarity is, however, consistent with Arun's analysis for conditions that do not exhibit perfect reciprocity. Following the same approach as above for removing the eccentricity bias, response dissimilarity scores when the target was the non-uniform appearance are adjusted so that the post-correction correlation between the two target options is 1.

We justify the removal of the eccentricity and reciprocity biases as aligning different perceptually linear spaces. Each combination of eccentricity level and target choice represent perceptual judgments in a slightly different space. However, by applying linear transformations to them, they can be aligned so that the effects of the other experimental variables can be evaluated with more robustness. All subsequent analysis of the measured values has eccentricity and reciprocity corrected for.

Subject consistency

As stated earlier, for each appearance, A, there were three presentations evaluated successfully by seven subjects. Ordering effects and consistency of a subjects' response can be measured with a repeated measures ANOVA for each stimuli. 4 of 288 stimuli (when considering all three print orientations) have a p-value < 0.05. The remaining images cannot reject the null hypothesis that there was no substantial change from the repetitions. This suggests subjects were self-consistent when viewing the stimuli multiple times and, for the remainder of the results section, we use the median response time of a subject's repeated trials to calculate their personal dissimilarity score per appearance (see Equation 2).

We use Cronbach's α to evaluate the consistency between the seven subjects in each of the print angle stimuli blocks. While not a statistic, it is a measure of internal consistency between a set of abstract items and repeated evaluations of said items. A subject is considered to be an item and each appearance dissimilarity measurement is a sample of the item's behavior, so that inter-subject consistency is determined by α. Each between-subjects block is evaluated separately: $\alpha_0 = 0.9221$, $\alpha_{10} = 0.9062$, and $\alpha_{45} = 0.9120$. A score greater than 0.9 has generally been considered as excellent consistency in past psychophysical research, while values greater than 0.7 represent reasonable inter-consistency.

Given this, we assume that subjects have high inter-person consistency, which implies that people are making similar perceptual decisions. Thus, we can combine all subject evaluations for an appearance to get a better estimate of its perceived dissimilarity. This was shown in Equation 3 as the median of the subject measurements. The aggregate scores, d_A, are used in the next subsection to explore the impact of the independent variables in our study. A breakdown of the dissimilarities for all variables is included at the end of the paper in Figure 16.

Analysis

The main effects of the manipulated variables on the d_A values for every $A \in \mathcal{A}$ are shown in Figure 11. Each bar in a variable's chart corresponds to the average d_A for all appearances matching the particular variable's value. Error bars represent the standard error within that set. There was a fifth independent variable that was between subjects, for the tile orientation when the material was printed. Additionally, there were four within-subjects independent variables that modified a stimuli's appearance: printed color, post-processing treatment, lighting environment, and synthetic SVBRDF variations. Additionally, because subjects provided perceptually linear scales, d_A scores across the different appearances and subject blocks can be examined without additional rescaling or normalization.

From these summaries, it is clear that many of the variables had noticeable effects on the perceived appearance uniformity. Of particular significance to 3D printing, the three printed tile orientations had significant changes to their aggregated dissimilarity scores. The 10° orientation, with its regular striations, had the worst uniformity, which is not surprising. Even though the 0° and 45° orientations featured similar, stochastic non-uniformities, their effects on dissimilarities differed. Thus, orientation with respect to the printer's Z axis can play a role in the perceived uniformity, at least when printing a flat tile. When 3D printers produce more complex objects composed of multiple surface orientations,

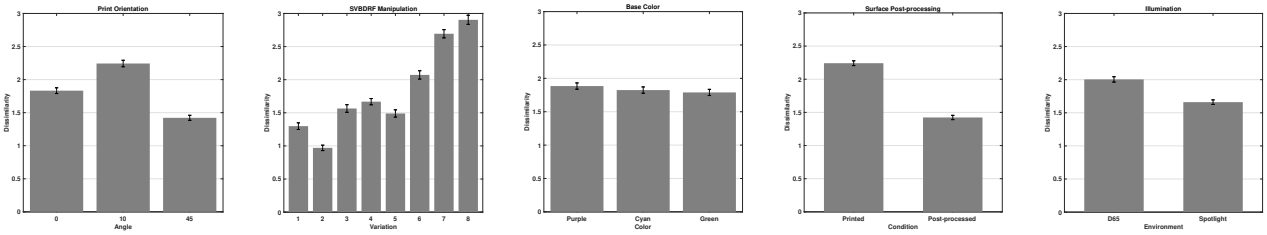

Figure 11. *Main effects of the four within-subjects and one between-subjects independent variables. Dissimilarity values are averaged over the other variables, error bars represent standard error.*

it has yet to be seen if perceived uniformity will behave in the same manner.

It does not appear as though the base diffuse color significantly changed the measured dissimilarity scores. This is not to say that color does not impact uniformity judgments, but that it is more likely that the 3D printer we utilized had reasonable color printing capabilities over the 1cm $\times$ 1cm printed region. The effects of surface post-processing, lighting, and SVBRDF manipulations are examined below.

Interaction effects

The main effect of post-processing the printed tiles yielded significant improvements in the perceived uniformity. Once again, this is not a surprising discovery since the post-processing resulted in smoother and more color-saturated surfaces. In the left side of Figure 12, we break down how the post-processing changes were affected by the SVBRDF variations. Of particular note are how the dissimilarities changed between the two finish conditions for variations 3, 4, and 5. These three variations represent the intermediate levels of SVBRDF changes: variations 3 and 4 are progressive bumpiness increases, and variation 5 is a change to glossiness.

In the print-only condition, which has a higher base level bumpiness compared to the post-processed condition, the two bumpy variations do not have significantly different dissimilarities. This suggests that non-uniformity due to bumpiness may reach a threshold. Under the post-processed condition, variation 3 and 4 both increase relative to the original but also maintain an almost linear trend with each other, having not yet reached that bumpiness threshold. Consistent with past literature on mesoscale rough surfaces' effects on glossiness [20], the bumpier print-only condition has less of a change due to glossiness increase compared to the post-processed variant.

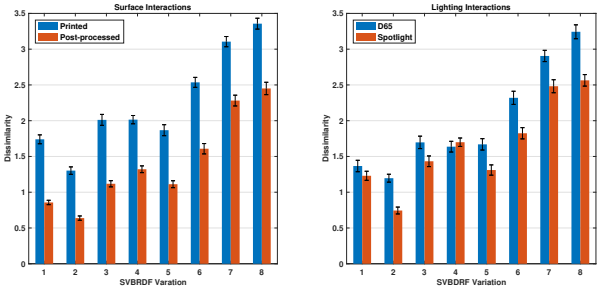

Figure 12. *Interaction effects between the SVBRDF modifications and surface post-processing.*

Lighting, however, had the opposite effect from what was anticipated. The spotlighting appeared to improve perceived uniformity, but we had expected its directional component would highlight bumps and cast more shadows, reducing the uniformity. From Figure 12 (right), it can be seen that the spotlight condition consistently improved the perceived uniformity of an appearance image. Comparing the two bumpiness variations, 3 and 4, we see that the general lighting from the D65 light booth makes it hard to distinguish the changes in mesoscale height. This makes sense since light from additional directions will fill out otherwise shadowed regions. The spotlight condition demonstrates the expected trend between variations 3 and 4. However, due to the overall upset of our lighting hypothesis, we explore it in more detail below.

Reviews of the results for two different lighting environments used for the rendering of the blobs seem to indicate that the dissimilarities for a spotlight source were smaller than for a diffuse D65 lighting condition modeled after a light booth. At first glance this seems counter intuitive. However, comparing the lightness distributions of a blob, with a cyan material that was printed at 0 degrees and rendered with the diffuse and spot illumination (see Figure 13), it is obvious that the overall lightness was much higher for the D65 illumination. Observing non-uniformities due to color or bumpiness is generally much harder in darker, shadow areas, which are present for the blob rendered with the spotlight. On the other hand, the lighter color distribution of the D65 blob enables non-uniformities to be recognized quite easily.

Figure 14 visualizes the probability density estimates for the blobs under the two different light environments. The median lightness value is 61 for D65, versus 42 for the spot illumination. Furthermore the dynamic range for the D65 blob is much smaller (34) than for the spot light (63). A higher dynamic range, which is also nicely visible by comparing the cumulative distribution functions of the two light environments always leads to images that convey the impression of a 3D object much stronger than one with

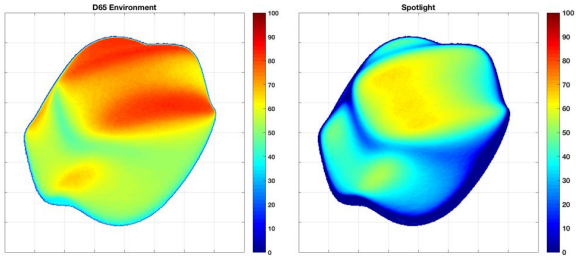

Figure 13. *False color rendering of L* from the appearances in Figure 6.*

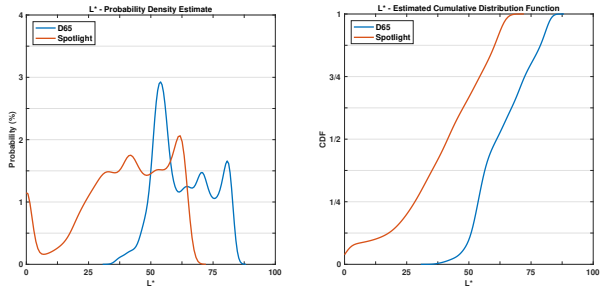

Figure 14. *L* distribution summary of the two appearance images shown in Figure 13.*

lower dynamic range. Photographers like Ansel Adams used it for striking photographs. The lower dynamic range image looks flatter, and consequently the viewer perceives it more as 2D surface on which spatial variations are easily discerned.

Examining the lightness distributions and using the second order statistics of skewness, which has previously been related to material appearance perception [18], it can be observed that the values are quite different: the probability density estimate for the point source has a higher absolute skewness (0.6) then the D65 probability density estimate (0.2). All of that indicates a higher spatial variation (noisier or grainier image), which makes it harder to perceive changes. This explains the differences concerning the obtained perceived dissimilarities.

Moving forward and trying to evaluate the influence of different lighting environments it might be a good idea to at least increase the intensity of the spot light so that the median values are more similar. Adjusting the histograms of lightness distributions so that the dynamic ranges are closer might also be worth considering. Having said that the question remains whether, after all those changes, the resulting images are still perceived to be under a realistic spot light illumination.

SVBRDF variation effects

The data visualized in Figure 15 is the same as in Figure 11, except that it breaks apart the SVBRDF variations based on bumpiness, color contrast, and glossiness. Bumpiness, color contrast and gloss can be seen as the axes of a 3D space, where the ideal stimuli is placed at the origin. The different stimuli that were used in the psychophysical tests were scaled from the original material in each of the three axes (see Figure 4 for the steps used). The centers of the circles in this figure represent those stimuli. For example bumpiness was varied by either a factor of 0.1, 1, 1.5 or 2. The color contrast of the stimuli varied by a scalar of 1, 2 or 3. Gloss varied only by a factor of 1 or 2. Only 2 samples had a gloss that was 1 step away from the original. Those two samples (variations 5 and 8) are indicated with a red border. The dissimilarities resulting from the psychophysical test are represented by the size of the circles.

Changes in color contrast resulted in bigger dissimilarities than changes in roughness. This holds true for the chosen units. Using a stimuli like variation 7 that already has a high color contrast and increasing both the bumpiness and the gloss doesn't have much impact on the perceived dissimilarity. The figure also shows that just changing color contrast results in some dissimilarity, but then gradually increasing the bumpiness increases the dissimilarity further. In summary, this figure shows the type of SVBDRF variations used in the visual search tasks and their corresponding dissimilarities. Each SVBDRF variation axis can change uniformity, and they can additively combine, although we have not yet explored if this relationship is linear.

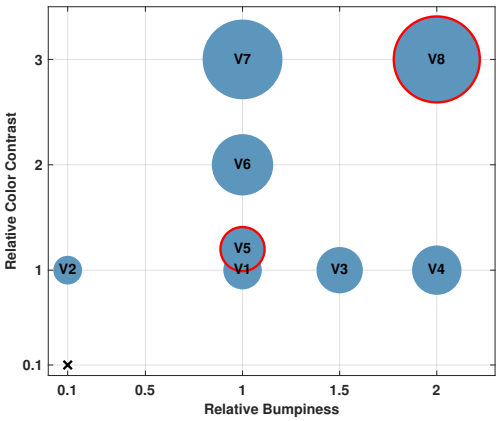

Figure 15. *Breakdown of bumpiness, color contrast, and gloss SVBRDF variations. The red border indicates an increase in glossiness. The black × represents the idealized variant.*

Conclusion

In this paper we have begun studying the perception of appearance uniformity of 3D printed materials. Uniformity is an important visual quality when 3D printing is used in art and consumer products. We have defined appearance uniformity as a subset of the appearance similarity problem. To our knowledge, this is the first user study to apply the visual search task to study appearance perception. Previously it has been used to explore attention and more primitive visual phenomena such as contrast gradients and object silhouettes. In our work, it has successfully been used to model perceptual differences between appearances with complex, spatially-varying materials, while remaining a simple and fast-paced task.

After analyzing the collected data, we have shown that the reciprocal of response time in the visual search task remains a consistent perceptual measure across subjects and material variations. We have identified a number of effects that influence the appearance uniformity of a material. Orientation within the printer bed was shown to be significant, which emphasizes the need to include perceptual analysis of any algorithm for packing and arranging objects in a printer. Post-processing the printed object can greatly improve uniformity, which can be a useful addition to any printing process. We have also demonstrated that lighting and SVBRDF variations can interact in ways to emphasize or hide non-uniformities. The perceptual uniformity measurements for the real world materials stands as its own contribution to the study of general appearance similarities.

Acknowledgments

Partially funded by the University of Minnesota's Doctoral Dissertation Fellowship.

References

[1] Arun, S. P. Turning visual search time on its head. *Vision Res.* 74 (2012), pp. 86–92. DOI: 10.1016/j.visres.2012.04.005.

[2] Ashikhmin, M. and Shirley, P. *An Anisotropic Phong BRDF Model*. Tech. rep. 2000. DOI: 10.1080/10867651.2000.10487522.

[3] Brainard, D. H. The Psychophysics Toolbox. *Spatial Vision* 10.4 (1997), pp. 433–436. DOI: 10.1163/156856897X00357.

[4] Brunton, A. et al. Pushing the Limits of 3D Color Printing. *ACM Trans. Graph.* 35.1 (2015), pp. 1–13. DOI: 10.1145/2832905.

[5] Cook, R. L. and Torrance, K. E. A Reflectance Model for Computer Graphics. *ACM Trans. Graph.* 1.1 (1982), pp. 7–24. DOI: 10.1145/357290.357293.

[6] Dalal, E. N. and Natale-Hoffman, K. M. The Effect of Gloss on Color. *Color Res. Appl.* 24.5 (1999), pp. 369–376. DOI: 10.1002/(SICI)1520-6378(199910)24:5<369::AID-COL8>3.3.CO;2-1.

[7] Dong, Y. et al. Appearance-from-motion. *ACM Trans. Graph.* 33.6 (2014), pp. 1–12. DOI: 10.1145/2661229.2661283.

[8] Duncan, J. and Humphreys, G. W. Visual Search and Stimulus Similarity. *Psychol Rev.* 96.3 (1989), pp. 433–458.

[9] Fleming, R. W. et al. Real-world illumination and the perception of surface reflectance properties. *Journal of Vision* 3.5 (2003), pp. 347–368. DOI: 10.1167/3.5.3.

[10] Gardner, A. et al. Linear Light Source Reflectometry. *ACM Trans. Graph.* 22.3 (2003), pp. 749–758. DOI: 10.1145/882262.882342.

[11] Ghosh, A. et al. A Basis Illumination Approach to BRDF Measurement. *Int J Comput Vis* 90.2 (2008), pp. 183–197. DOI: 10.1007/s11263-008-0151-7.

[12] Hergel, J. and Lefebvre, S. Clean color. *Computer Graphics Forum* 33.2 (2014), pp. 469–478. DOI: 10.1111/cgf.12318.

[13] Ludwig, M. and Meyer, G. Brightness Perception of Surfaces with Mesoscale Structures. *JIST* 61.2 (2017), pp. 1–14. DOI: 10.2352/J.ImagingSci.Technol.2017.61.2.020504.

[14] Marlow, P. J. et al. The Perception and Misperception of Specular Surface Reflectance. *Current Biology* 22.20 (2012), pp. 1909–1913. DOI: 10.1016/j.cub.2012.08.009.

[15] Matusik, W. et al. A Data-Driven Reflectance Model. *ACM Trans. Graph.* 22.3 (2003), pp. 759–769. DOI: 10.1145/882262.882343.

[16] Mcallister, D. K. *A Generalized Surface Appearance Representation for Computer Graphics*. PhD thesis. 2002.

[17] Monnier, P. Detection of multidimensional targets in visual search. *Vision Res.* 46 (2006), pp. 4083–4090. DOI: 10.1016/j.visres.2006.07.032.

[18] Motoyoshi, I. et al. Image statistics and the perception of surface qualities. *Nature* 447 (2007), pp. 206–209. DOI: 10.1038/nature05724.

[19] Muryy, A. A. et al. Specular reflections and the estimation of shape from binocular disparity. *Proc. Natl. Acad. Sci.* 110.6 (2013), pp. 2413–2418. DOI: 10.1073/pnas.1212417110.

[20] Padilla, S. et al. Perceived roughness of $1/f^{\beta}$ noise surfaces. *Vision Res.* 48.17 (2008), pp. 1791–1797. DOI: 10.1016/j.visres.2008.05.015.

[21] Pellacini, F. et al. "Toward a Psychophysically-Based Light Reflection Model for Image Synthesis". In *SIGGRAPH*. 2000, pp. 55–64. DOI: 10.1145/344779.344812.

[22] Pereira, T. and Rusinkiewicz, S. Gamut Mapping Spatially Varying Reflectance with an Improved BRDF Similarity Metric. *Computer Graphics Forum* 31.4 (2012), pp. 1557–1566. DOI: 10.1111/j.1467-8659.2012.03152.x.

[23] Pramod, R. T. and Arun, S. P. Object attributes combine additively in visual search. *Journal of Vision* 16.5 (2016), pp. 1–29. DOI: 10.1167/16.5.8.

[24] Reiner, T. et al. Dual-Color Mixing for Fused Deposition Modeling Printers. *Computer Graphics Forum* 33.2 (2014), pp. 479–486. DOI: 10.1111/cgf.12319.

[25] Riviere, J. et al. Mobile Surface Reflectometry. *Computer Graphics Forum* 35.1 (2015), pp. 191–202. DOI: 10.1111/cgf.12719.

[26] Tunwattanapong, B. et al. Acquiring Reflectance and Shape from Continuous Spherical Harmonic Illumination. *ACM Trans. Graph.* 32.4 (2013), pp. 1–12. DOI: 10.1145/2461912.2461944.

[27] Vangorp, P. et al. The Influence of Shape on the Perception of Material Reflectance. *ACM Trans. Graph.* 26.3 (2007), pp. 1–10. DOI: 10.1145/1239451.1239528.

[28] Wang, W. M. et al. Improved Surface Quality in 3D Printing by Optimizing the Printing Direction. *Computer Graphics Forum* 35.2 (2016), pp. 59–70. DOI: 10.1111/cgf.12811.

[29] Wang, W. et al. Saliency-Preserving Slicing Optimization for Effective 3D Printing. *Computer Graphics Forum* 34.6 (2015), pp. 148–160. DOI: 10.1111/cgf.12527.

[30] Ward, G. J. "Measuring and Modeling Anisotropic Reflection". In *SIGGRAPH*. 1992, pp. 265–272. DOI: 10.1145/142920.134078.

[31] Wills, J. et al. Toward a Perceptual Space for Gloss. *ACM Trans. Graph.* 28.4 (2009), pp. 1–15. DOI: 10.1145/1559755.1559760.

[32] Wolfe, J. M. "Visual Search". In *Attention*. 1998, pp. 1–41.

[33] Wolfe, J. M. Asymmetries in visual search. *Perception & Psychophysics* 63.3 (2001), pp. 381–389. DOI: 10.3758/BF03194406.

[34] Wyszecki, G. and Stiles, W. S. *Color Science*. 2nd ed. Concepts and Methods, Quantitative Data and Formulae. Wiley-Interscience, 1982.

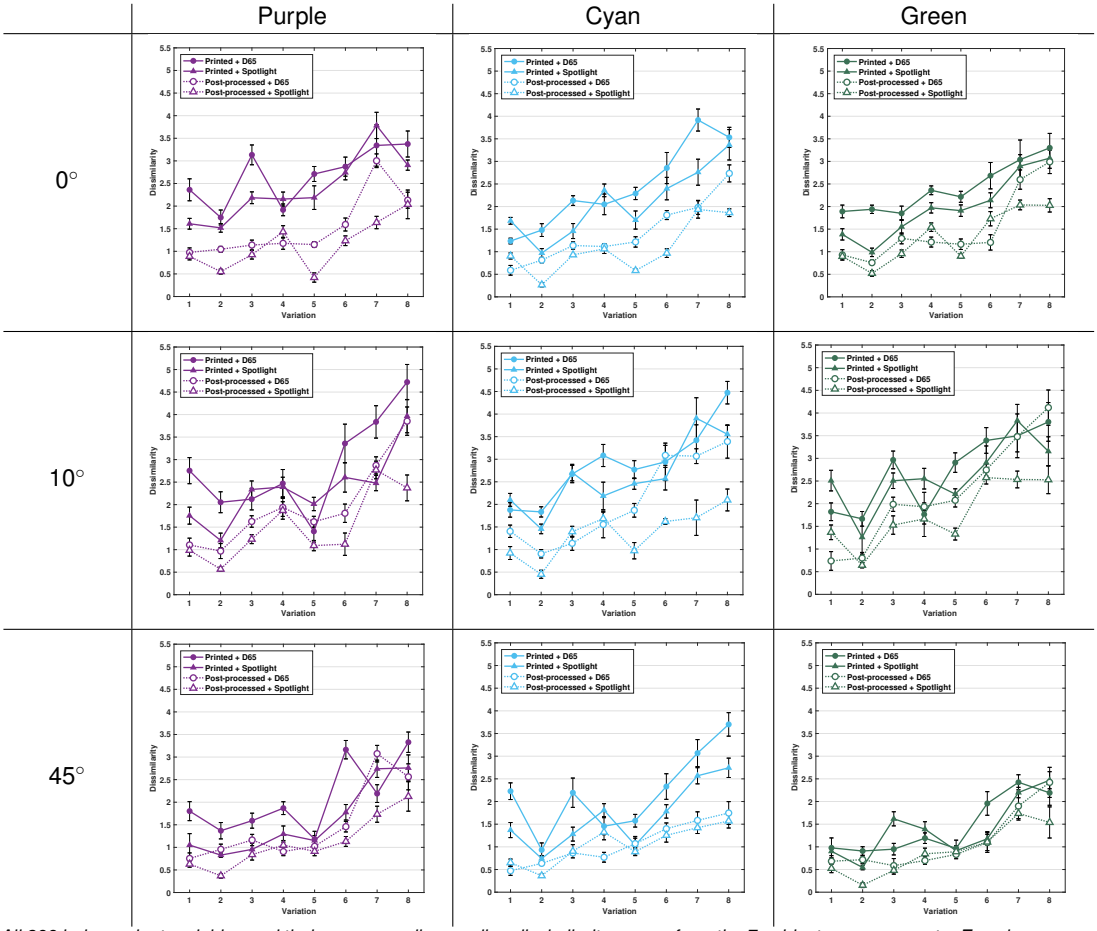

Figure 16. *All 288 independent variables and their corresponding median dissimilarity scores from the 7 subject measurements. Error bars represent standard error within subject ratings.*