MatMix 1.0: Using optical mixing to probe visual material perception

Fan Zhang
Perceptual Intelligence Laboratory, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, the Netherlands

Huib de Ridder
Perceptual Intelligence Laboratory, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, the Netherlands

Roland W. Fleming
Department of Psychology, Justus-Liebig-Universität Gießen, Giessen, Germany

Sylvia Pont
Perceptual Intelligence Laboratory, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, the Netherlands

MatMix 1.0 is a novel material probe we developed for quantitatively measuring visual perception of materials. We implemented optical mixing of four canonical scattering modes, represented by photographs, as the basis of the probe. In order to account for a wide range of materials, velvety and glittery (asperity and meso-facet scattering) were included besides the common matte and glossy modes (diffuse and forward scattering). To test the probe, we conducted matching experiments in which inexperienced observers were instructed to adjust the modes of the probe to match its material to that of a test stimulus. Observers were well able to handle the probe and match the perceived materials. Results were robust across individuals, across combinations of materials, and across lighting conditions. We conclude that the approach via canonical scattering modes and optical mixing works well, although the image basis of our probe still needs to be optimized. We argue that the approach is intuitive, since it combines key image characteristics in a "painterly" approach. We discuss these characteristics and how we will optimize their representations.

Introduction

Natural materials scatter light in various manners. Even if we limit ourselves to the main scattering characteristics of opaque materials, we probably still need about a dozen scattering types or canonical modes to represent most materials. Bidirectional reflectance distribution functions (BRDFs) provide a physical description of how opaque material surfaces scatter light. Knowing how light scatters from surfaces makes it possible to simulate materials using parametric BRDF models in computer renderings (Newell & Blinn, 1977; Cook & Torrance, 1982; Hapke, Nelson, & Smythe, 1998; Koenderink & Pont, 2003; Koenderink, Van Doorn, Dana, & Nayar, 1999; Nayar & Oren, 1995; Oren & Nayar, 1995; Phong, 1975; Torrance & Sparrow, 1967; Torrance, Sparrow, & Birkebak, 1966; van Ginneken, Stavridi, & Koenderink, 1998; Ward, 1992). Generally speaking, if the scattering properties or optical characteristics of materials can be accurately described, the so-called forward rendering problem can be solved. However, it is very unlikely that these optical characteristics correspond to the representation of the visual attributes in the brain (Fleming, 2014). In other words, we do not see BRDFs. On the one hand, a BRDF combined with various object shapes and lighting conditions can result in different images of the same material (we consider an image as the resulting optical structure projected on a picture or the retina). On the other hand, different combinations of BRDF, object shape, and lighting can result in similar images. In other words, the so-called inverse problem does not have a unique solution. Thus, images contain ambiguities of material, shape, and light (Belhumeur, Krieg-

Citation: Zhang, F., de Ridder, H., Fleming, R. W., & Pont, S. (2016). MatMix 1.0: Using optical mixing to probe visual material perception. *Journal of Vision, 16*(6):11, 1–18, doi:10.1167/16.6.11.

doi: 10.1167/16.6.11 Received December 15, 2015; published April 18, 2016 ISSN 1534-7362

man, & Yuille, 1999), and consequently material, light, and shape perception are confounded (Anderson, 2011). In the present study, we will exploit this metamerism and investigate key ingredients in the images that trigger our material perceptions.

Most investigations into material perception are confined to matte–glossy variations. That is, the perception of materials varying from matte to shiny has been intensively studied on perceived glossiness (Anderson, 2011; Anderson & Kim, 2009; Fleming, 2012; Fleming, Dror, & Adelson, 2003; Ho, Landy, & Maloney, 2006, 2008; Kim, Marlow, & Anderson, 2011; Marlow, Kim, & Anderson, 2012; Motoyoshi, Nishida, Sharan, & Adelson, 2007; Nishida & Shinya, 1998; Pellacini, Ferwerda, & Greenberg, 2000; Vangorp, Laurijssen, & Dutré, 2007; Wiebel, Toscani, & Gegenfurtner, 2015; Wijntjes & Pont, 2010). In our research, we want to address material perception not only within the matte–glossy continuum but also for as wide a range of natural materials as possible. To date, little is known about the visual perception of materials outside the matte–glossy dimension, such as velvetiness (Koenderink & Pont, 2003; Nishida, Sawayama, & Shimokawa, 2015; te Pas & Pont, 2005) or other material dimensions (Fleming, Wiebel, & Gegenfurtner, 2013; Sharan, Rosenholtz, & Adelson, 2014). The main problem seems to be the lack of a tool to test purely visually (without referring to physical parameters or attributes) and quantitatively what material is perceived, for a wide range of materials.

We hereby present MatMix 1.0, a novel material probe using optical mixing, which will be explained in the next section. MatMix 1.0 is meant to account for a wide range of opaque materials. We use optical mixing of four canonical scattering modes as a tool for quantitatively measuring visual perception of materials. In our main study, we integrated the probe into a MATLAB graphical user interface and conducted two matching experiments without (Experiment 1) and with (Experiment 2) variation of the illumination and viewpoint conditions. Images of real objects were used as a basis set. Before the experiments started, we expected the task to be difficult for inexperienced observers, as they would have to simultaneously manipulate four different scattering modes. Surprisingly, we found that all participants could handle MatMix 1.0 well, as indicated by the finding that they performed far above chance level within reasonable amounts of time. In an additional study, we replaced the images with renderings and conducted a similar matching experiment (Experiment 3). Again, participants performed far above chance level, demonstrating that the approach works well with both real and simulated materials. In the General discussion and the Conclusion we address the relationships between a few key image characteristics and the results.

MatMix 1.0: A novel material probe

Optical mixing: A painterly approach

Many arbitrary materials can be represented by linearly combining surface scattering distributions (Matusik, Pfister, Brand, & McMillan, 2003; Pellacini et al., 2000; Ward, 1992). Instead of directly combining reflectance functions, we propose to linearly superpose images of objects with the same shape but finished with different materials. This image-combination process, called optical mixing, was introduced by Griffin (1999), who also described the mathematics behind the optical-mixing method and showed that it could be used as a tool for visual-perception studies. In Brainard's lab, Griffin's partitive mixing method was applied to reduce the number of stimuli to be rendered for their experiments (Olkkonen & Brainard, 2010; Radonjić, Cottaris, & Brainard, 2015; Xiao & Brainard, 2008). Although applying image mixing was not the main purpose of those studies, it can still be concluded from them that implementing optical mixing in psychophysical studies is indeed feasible and efficient. However, it has not been implemented yet for variations other than matte–glossy.

The optical-mixing procedure shows an interesting analogy with how a painter renders materials in a scene. Most painters do not think about image statistics or BRDFs when they paint. Instead, their approach is more similar to optical mixing of key visual ingredients layer by layer. A frequently observed recipe for oil paintings (Wallert, 1999) is to first draw the contour of an object, then apply the matte layer (the diffuse body scattering), and finally add highlights or a bright contour to render specular or velvety elements (forward or asperity scattering). We reasoned that optical mixing of nonspherical objects of arbitrary scattering modes should work because it similarly combines key image ingredients that trigger our perceptions—even though it may be physically incorrect.

Pont, Koenderink, Van Doorn, Wijntjes, and te Pas (2012) generated optical mixtures of three canonical scattering modes (matte, velvety, and specular) by optically mixing real objects in a viewing box. The task for the observers was to rate perceived material qualities such as glossiness, warmth, hardness, and softness. In a follow-up study, observers performed the same task, but now the stimuli were optically mixed images of matte, velvety, and specular materials displayed on a screen (Pont, te Pas, & Wijntjes, 2014). They obtained robust and systematic ratings for material qualities as a function of the weights of the three modes in both experiments. On the basis of these studies, we hypothesized that observers should be able to match the perceived material of a certain object if

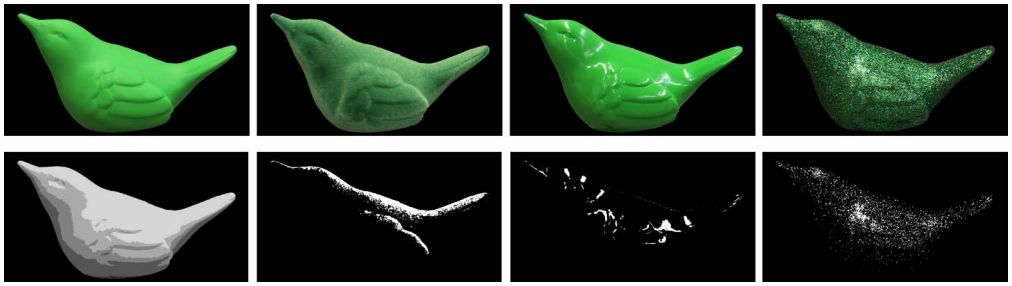

Figure 1. The top row shows the images of the birdlike object with the four materials representing the chosen canonical scattering modes: diffuse, asperity, forward, and meso-facet scattering (from left to right). These modes are represented by matte, velvety, specular, and glittery materials. The bottom row shows the prototypical image characteristics of each material. Note that the reflectance components are not only in different directions in BRDF space but result in characteristics of the proximal image in different regions too. For the image of the matte bird, the green channel was posterized from 255 to six levels. For the velvety, specular, and glittery bird images, we performed red-channel thresholding at the 50% level. These extremely simple processes resulted in smooth shading from the top to the bottom of the object for the matte object, bright contours for the velvet object, highlights at specular points for the specular object, and bright speckles all over the surface for the glittery object.

they have the opportunity to create a mixture with desired material attributes via control over the weights of the underlying canonical material modes in that mixture. This forms the basis of our proposed new materials probe, MatMix 1.0.

For MatMix 1.0, we employed four basis materials by finishing four bird-shaped objects with matte, velvety, specular, and glittery materials (Figure 1). These materials represent four canonical scattering modes, namely diffuse, asperity, forward and meso-facet scattering modes. The scattering distributions of these four canonical scattering modes together span a large part of the BRDF space. Because the main scattering directions of these scattering modes are different, the key characteristics in the images of the corresponding objects will end up in different locations on the object too. This means that the reflectance components not only are complementary in BRDF space but also allow the user to adjust different characteristics of the proximal image. With simple image analyses, the prototypical image characteristics of the four materials can be easily distinguished from each other, as shown in Figure 1. Note that there are many alternative image manipulations that would give similar results; the examples just serve to demonstrate the main idea.

The interface of MatMix 1.0

Inspired by audio-mixing desks, we built a user interface consisting of four sliders, a stimulus window, and a probe window (Figure 2). During each matching trial, the stimulus image and the probe image were simultaneously presented to the observers in the corresponding windows, with the stimulus on the left-hand side and the probe on the right-hand side. The four sliders were positioned directly underneath the probe window. In order to give purely visual information, we avoided the use of terms like "matte," "velvet," and so on in the interface. Instead, we put cropped images (the head parts of the bird images) in front of each slider, representing the material modes. The position of each slider bar represents the selected weight value per material mode, varying between 0 and 1.2.

The interface was developed using features of graphical user interfaces in MATLAB R2014a, and

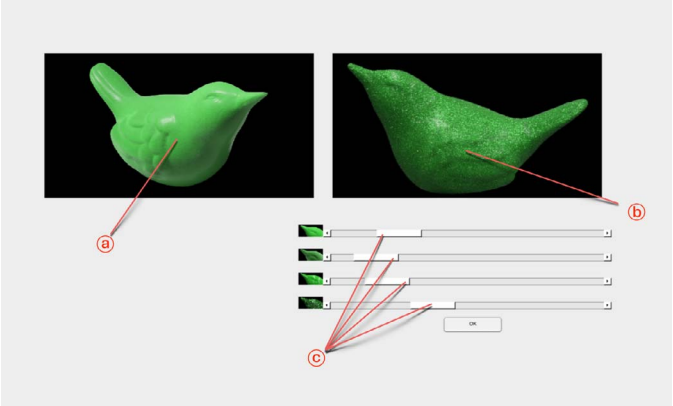

Figure 2. The interface of MatMix 1.0. (a) Stimulus image. (b) Material probe, generated by linear weighted superposition of the four images representing the canonical scattering modes. (c) Four sliders, with the position of each slider bar representing the selected weight value per material mode, ranging from 0 to 1.2 (left to right). The icon on the left of each slider visualizes the corresponding material component. The task of the observers was to change the material of the probe to match the stimulus. They could take as much time as they needed. Observers could click the OK button below the sliders to finish the matching procedure. Here, it is obvious that the two materials do not match.

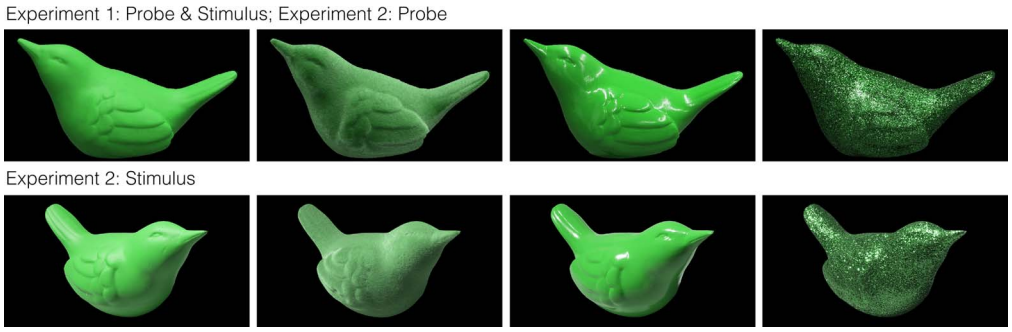

Figure 3. Basis images. From left to right, the columns represent the matte, velvety, specular, and glittery modes, respectively. The images in the top row were taken under office lighting and were used as basis images for the probe in the main study. These images were also used as the basis for the stimulus images in Experiment 1. The images in the bottom row were taken in studio lighting and from a different viewing angle than the first set, and were used as the basis for the stimulus in Experiment 2.

presented to observers on a linearly calibrated Apple 15-in. Retina display.

Basis images

The surfaces of four physical objects with identical shapes were finished with matte, velvety, specular, and glittery materials. The bird-shaped objects were purchased in a shop. They were originally made of ceramic and had exactly the same shape. The matte and specular birds were created by spray-painting them with matte and glossy paint, respectively (both color RAL 6018). The glittery bird was created by repeatedly sprinkling green glitter over a layer of spray glue. The velvety bird was finished by a factory using a technique called flocking (color RAL 6018). These materials represent diffuse, asperity, forward, and meso-facet scattering modes, respectively. We took photos of the objects under office lighting and under studio lighting from different viewing angles. The office lighting consisted of multiple fluorescent tubes in the ceiling of a room without daylight. The studio lighting consisted of a halogen spotlight from the left side of the object. The camera settings were kept constant per lighting condition and we used raw imaging in order to photometrically gauge the basis images. Furthermore, to allow superposition of the basis images, we placed each object in exactly the same position. To do so, we drew their cast shadows and base outlines on their groundings as references. Next we adjusted the white balance of the raw images using Adobe Photoshop so that the highlights were all white. We did this in the same manner for all images per lighting condition. Then we segmented the images using the shared contours of the birds and made the background black for all images. Last, in order to avoid color interactions, we set the hue value to 0.33 (green) for all images using MATLAB. Because the birds were pure green, this transformation had a negligible influence on the images (Figure 3). The saturation of the colors was not adjusted, because the saturation as a function of lighting and viewing angles can vary strongly. For instance, specular reflections lower the saturation of highlights. This effect depends on the type of scattering (Klinker, Shafer, & Kanade, 1987; Koenderink et al., 1999; Koenderink & Pont, 2008; Shafer, 1985; Wolff, 1994). This is why the different modes have substantial differences in saturation.

The probe: MatMix 1.0

The probe is a linearly superposed optical mixture of the basis images. The mixing process can be illustrated by Equation 1:

$$I_{\text{probe}} = w_{\text{m}} \cdot I_{\text{m}} + w_{\text{v}} \cdot I_{\text{v}} + w_{\text{s}} \cdot I_{\text{s}} + w_{\text{g}} \cdot I_{\text{g}}, \quad (1)$$

where subscripts {m, v, s, g} denote the four scattering modes matte, velvety, specular, and glittery, representing the four canonical scattering modes (diffuse, asperity, forward, and meso-facet scattering); $\{w_{\text{m}}, w_{\text{v}}, w_{\text{s}}, w_{\text{g}}\}$ are the weight values corresponding to the positions of the slider bars, ranging from 0 to 1.2 (see Figure 2); and $\{I_{\text{m}}, I_{\text{v}}, I_{\text{s}}, I_{\text{g}}\}$ are the basis images under office lighting (top row in Figure 3) for Experiments 1 and 2 in the main study. The linearly mixed image I_{probe} plus the interface forms the probe MatMix 1.0, which allows real-time dynamic and interactive variation of a visual presentation of material through adjustments of the slider bars.

Main study (optical mixing with images of real objects)

Introduction

In the main study we tested the material probe we developed. The study consisted of two experiments,

Stimulus	w'_m	w'_v	w'_s	w'_g
1	1	0	0	0
2	0	1	0	0
3	0	0	1	0
4	0	0	0	1
5	0.5	0.5	0	0
6	0.5	0	0.5	0
7	0.5	0	0	0.5
8	0	0.5	0.5	0
9	0	0.5	0	0.5
10	0	0	0.5	0.5
11	0.33	0.33	0.33	0
12	0.33	0.33	0	0.33
13	0.33	0	0.33	0.33
14	0	0.33	0.33	0.33
15	0.25	0.25	0.25	0.25

Table 1. Overview of the weight combinations of the four material modes that were used to generate the stimulus images. There were 15 stimuli in total.

which mainly differed in the illumination and viewpoint conditions under which the photos of real objects were taken. In Experiment 1, we created the stimuli by mixing the basis images shown in the top row in Figure 3. In Experiment 2, we created the stimuli by mixing the basis images shown in the bottom row in Figure 3. Thus, in Experiment 1 the stimuli and probe were mixed from the same basis, while in Experiment 2 the stimuli were mixed from a different basis than the probe.

Method

Stimuli

We tested 15 weight combinations of the four scattering modes, as shown in Table 1. The basis images were linearly superposed, implementing Equation 1 in the form

$$I_{\text{stimulus}} = (w'_m + x_m) \cdot I_m + (w'_v + x_v) \cdot I_v + (w'_s + x_s) \cdot I_s + (w'_g + x_g) \cdot I_g, \quad (2)$$

where $\{w'_m, w'_v, w'_s, w'_g\}$ are the weights of the scattering modes; $\{x_m, x_v, x_s, x_g\}$ are randomly generated offsets in a range from −0.1 to 0.1 that were added to the nonzero weights only; and $\{I_m, I_v, I_s, I_g\}$ are the stimulus basis images shown in Figure 3 (top row for Experiment 1, bottom row for Experiment 2). The resulting linearly mixed image is the stimulus image I_{stimulus}. The complete set of stimulus images for Experiments 1 and 2 is shown in Figure 4.

Observers

There were eight paid inexperienced participants in total (four men and four women, aged 23 to 30), with normal or corrected-to-normal vision. All of them participated first in Experiment 1, and a few days later in Experiment 2. Participants read and signed a consent form before the experiments. The experiments were conducted in agreement with the Declaration of Helsinki and local ethical guidelines and approved by the Human Research Ethics Committee of the Delft University of Technology.

Procedure

The positions of the slider bars (i.e., the initial weights of the probe) were randomly initialized in each trial. In Experiments 1 and 2, each stimulus weight combination in Table 1 was repeated three times. Three repeats combined with 15 different weight combinations, making each experiment 45 trials in total. The trials were presented in pseudorandom order. At the start of the experiment, the interface (Figure 2) was shown to the observers. The observers were instructed that their main task was to move the sliders to adjust the material of the bird in the top right window (probe) until it appeared to be made of the same material as the bird in the top left window (stimulus), and that they could take as much time as they needed. Once observers finished a matching trial, they pressed the "OK" button, after which only the stimulus and probe images were presented on the screen. The observers were asked to indicate to what extent they were satisfied with the matching result. After they pressed the "Next" button, the next matching trial started. Three trials were performed as practice trials before the experiment formally started. In the practice trials, participants were told that they could move the slider bars by dragging the mouse or pressing the left and right arrow keys on the keyboard. Moving the slider bars by dragging the mouse resulted in bigger steps, while pressing the arrow keys resulted in smaller steps and more gradual changes in the probe.

Results

Overview

In order to test the usability of the method and evaluate the general matching results per experiment, we will first fit one single linear equation to the complete set of weights of the stimuli and probe adjustments. Then we will analyze the satisfaction ratings and the durations of the matching processes. After that we will look into the details of the interactions of the four canonical material modes to analyze the perceptual effects in detail.

Matching results

The general results of the matching experiments were evaluated by solving the linear factor matrix **A** in

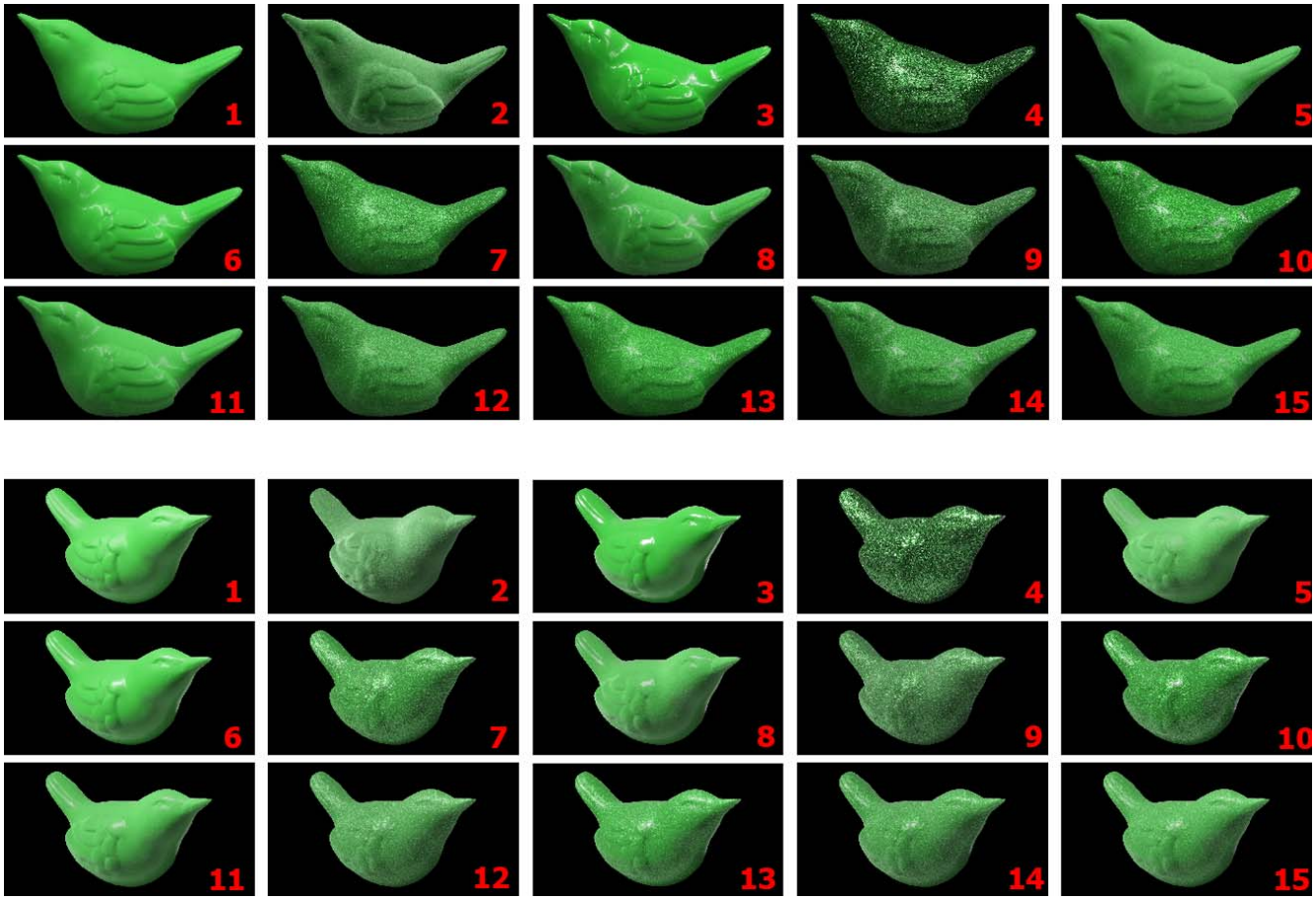

Figure 4. The stimuli. The top set represents the test stimuli in Experiment 1. The bottom set represents the test stimuli in Experiment 2. The randomly generated offsets $\{x_m, x_v, x_s, x_g\}$ were set to 0 to generate these images. The numbers in the images correspond to the stimulus numbers in Table 1.

Equation 3:

$$[\mathbf{Y}]_{4\times 360} = [\mathbf{A}]_{4\times 4}\cdot[\mathbf{X}]_{4\times 360} + [\mathbf{E}]_{4\times 360}, \quad (3)$$

where

$$[\mathbf{X}] = \begin{bmatrix} w'_m \\ w'_v \\ w'_s \\ w'_g \end{bmatrix}, \qquad [\mathbf{Y}] = \begin{bmatrix} w_m \\ w_v \\ w_s \\ w_g \end{bmatrix},$$

and the residuals

$$[\mathbf{E}] = \begin{bmatrix} e_m \\ e_v \\ e_s \\ e_g \end{bmatrix}.$$

For each trial, one column in matrix **X** represents the weights of the four scattering modes in the stimulus image, and the corresponding column in matrix **Y** represents the weights of the four scattering modes in the probe image, i.e., the values represented by the positions of the four sliders set by the participant. We consider all eight participants together. Thus, there are 45 trials for 8 participants = 360 columns in matrix **X**, matrix **Y**, and matrix **E** (the residuals). The 4×4 linear factor matrix **A** was solved using a least-squares fit in MATLAB, and then the matrix **E** was simply calculated as the difference between **Y** and **A** · **X**. If observers were to move all sliders so that the weights in matrix **Y** would be exactly equal to the corresponding weights in matrix **X** (i.e., the matching would be veridical), then **A** would be a 4×4 identity matrix and **E** would be a zero matrix.

The resulting matrix **A** of Experiment 1 is surprisingly close to an identity matrix (see Table 2). To be more specific, the nondiagonal values are 0.18 or lower and close to 0, and the diagonal values are 0.78, 0.89, 0.91, and 1.08 for the matte, velvety, specular, and glittery modes, respectively. In the resulting matrix for Experiment 2 the first three diagonal elements decreased to 0.65, 0.69, and 0.63 for the matte, velvety, and specular modes, respectively. The diagonal value for the glittery mode is 1.09, which is similar to that of Experiment 1. The nondiagonal values that represent the interactions between the scattering modes are larger for Experiment 2 than for Experiment 1. To be more specific, $\{w_m, w'_v\}$—the value between w_m and w'_v in

	w'_m	w'_v	w'_s	w'_g
Experiment 1				
w_m	**0.78**	0.14	0.16	0.00
w_v	0.18	**0.89**	0.03	0.00
w_s	0.18	0.04	**0.91**	0.04
w_g	−0.02	0.09	0.02	**1.08**
Experiment 2				
w_m	**0.65**	0.25	0.32	0.02
w_v	0.14	**0.69**	0.10	0.00
w_s	0.30	0.24	**0.63**	0.19
w_g	0.01	0.12	−0.00	**1.09**

Table 2. Linear factor matrices **A** for Experiments 1 and 2.

matrix **A**—was 0.14 in Experiment 1, which means that occasionally velvety contributions in the stimuli w'_v were perceived to match with a matte contribution in the probe w_m. The value increased from 0.14 to 0.25 in Experiment 2, showing that the chance increased of perceiving velvety contributions in the stimuli to match a matte contribution in the probe. Similarly, for the combination $\{w_m, w'_s\}$ the value increased from 0.16 to 0.32; for $\{w_s, w'_m\}$ it increased from 0.18 to 0.30; for $\{w_s, w'_v\}$ it increased from 0.04 to 0.24; and for $\{w_s, w'_g\}$ it increased from 0.04 to 0.19. Thus, overall, a comparison of the off-diagonal elements between the two experiments shows that the interactions between perceptions of matte, velvety, and specular modes became stronger when stimulus and probe were under different lighting and viewing conditions.

Another measure of general performance is the ratio between the sum of the four diagonal values in matrix **A** and the sum of all values in matrix **A**. This ratio can vary from 0 to 1, with veridical behavior at 1 (identity matrix) and chance level at 0.25 (all values in matrix **A** being equal). For each individual, we solved the linear factor matrix **A** with 45 trials per observer per experiment and calculated the ratios. As shown in Figure 5, in Experiment 1 the ratios for the observers were 0.80, 0.85, 0.72, 0.83, 0.80, 0.77, 0.70, and 0.80 ($M = 0.78$, $SD = 0.05$). In Experiment 2 these ratios were 0.47, 0.70, 0.58, 0.76, 0.51, 0.55, 0.64, and 0.69 ($M = 0.61$, $SD = 0.10$). Overall, all observers performed far above chance level.

We also analyzed how close the residuals (matrix **E**) were to 0. We first took the absolute values of the 4×360 matrices, and then calculated the mean of all elements in each 4×45 matrix, for each observer. The results were quite similar between observers per experiment. As shown in Figure 6, in Experiment 1 the means of the residuals' absolute values for the eight observers

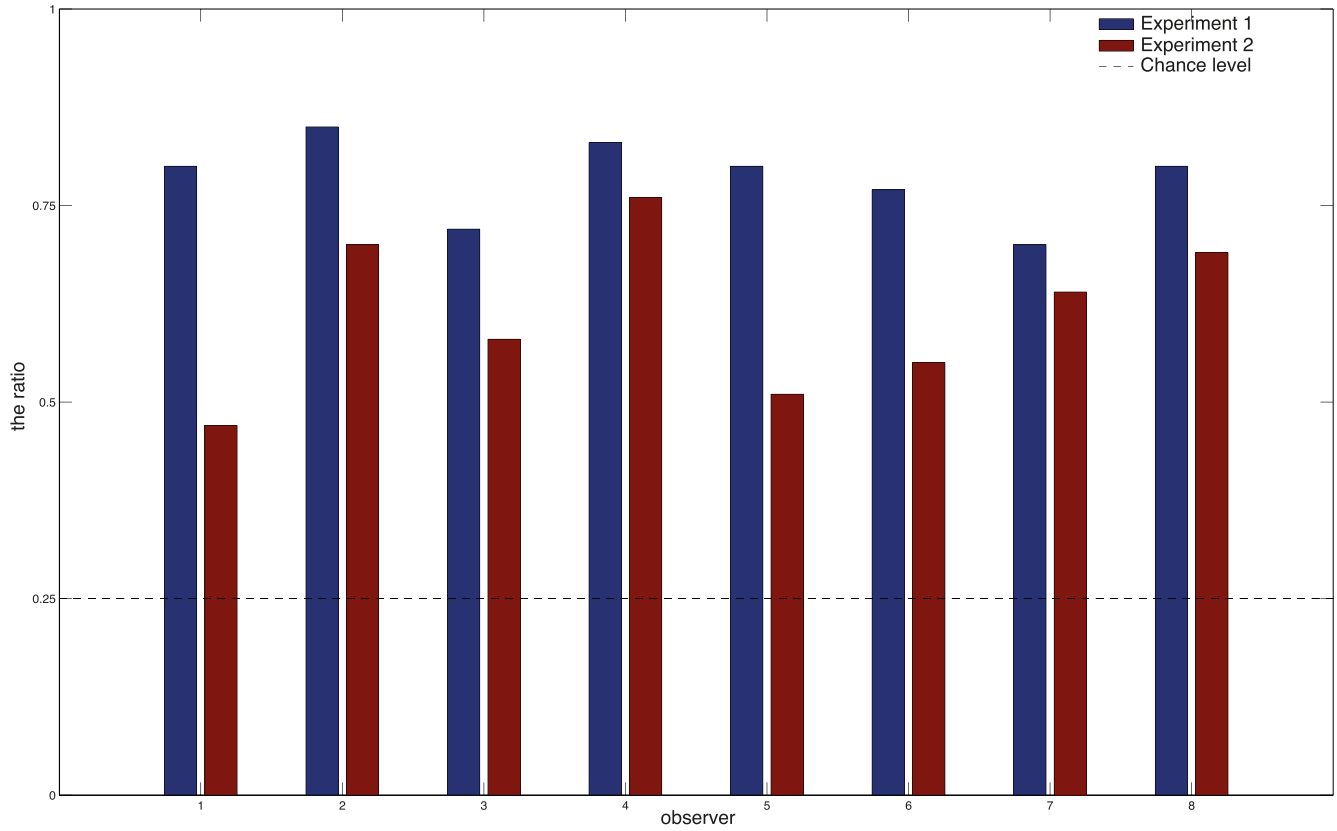

Figure 5. The ratio between the sum of the four diagonal values in matrix **A** and the sum of all values in matrix **A**. All eight observers performed far above chance level in Experiment 1 (blue) and Experiment 2 (red).

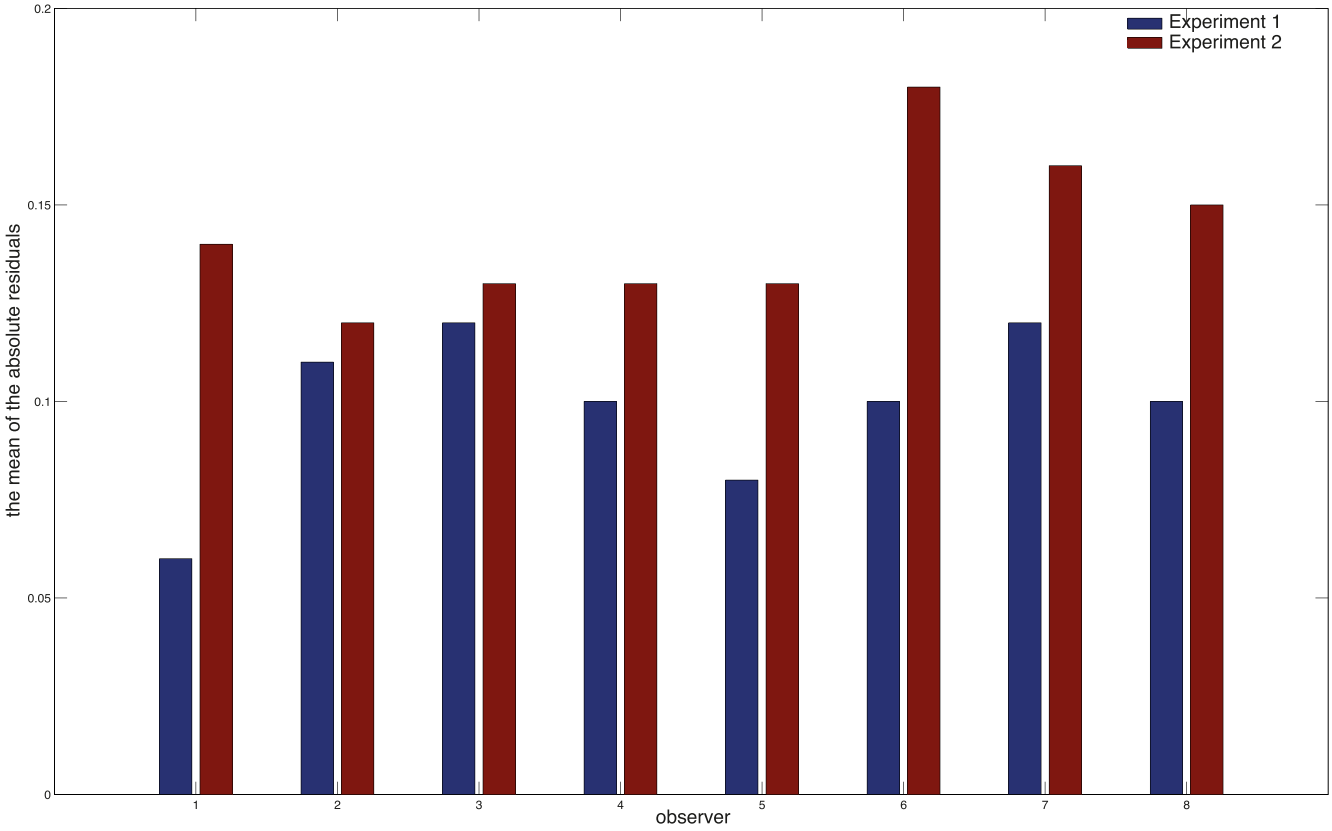

Figure 6. The mean of the absolute residuals of each observer in Experiment 1 (blue) and Experiment 2 (red).

were 0.06, 0.11, 0.12, 0.10, 0.08, 0.10, 0.12, and 0.10 ($M = 0.10$, $SD = 0.02$). In Experiment 2 these values became 0.14, 0.12, 0.13, 0.13, 0.13, 0.18, 0.16, and 0.15 ($M = 0.14$, $SD = 0.02$). We can conclude that the least-squares fit method properly solved the linear Equation 3.

Durations and satisfaction ratings

In Figure 7 we plotted the mean duration of the matching trials over all observers as a function of trial number. We fitted the data for both experiments simultaneously by means of multiple linear regression with one dummy variable to directly compare the slopes and establish a possible shift between the two regression lines. The first five trials of each experiment were excluded in the linear regression because we found the duration data in those trials to vary wildly, probably because observers were still exploring the possibilities of the interface. After the first five trials, the pattern of trial durations stabilized (see Figure 7). For Experiment 1, the slope of the regression line was found to deviate significantly from 0 (-1.11 ± 0.21, $p < 0.001$). The difference between the two slopes was also significant (0.58 ± 0.3, $p = 0.05$) resulting in a slope of -0.53 for Experiment 2. The offset for Experiment 1 (99.85 ± 6.0, $p < 0.001$) was higher than that for Experiment 2 (76.35; difference equals -23.49 ± 8.46, $p < 0.001$). These results imply that the duration for Experiment 1 started at a higher level than for Experiment 2, and afterward the durations of both experiments systematically decreased with trial number, converging to the same level at the final trials. In conclusion, the main effect is a gradual but small decrease of trial duration as a function of trial number. On average, the duration was slightly above 1 min per matching trial.

The satisfaction ratings were defined to range from 0 (*not satisfied* with the matching) to 1 (*satisfied* with the matching). Subsequently, we took the average over all observers per trial. Excluding the first five trials, data were again fitted by multiple linear regression with one dummy variable (Figure 8). The only significant effects were the offset for Experiment 1 (0.81 ± 0.02, $p < 0.001$) and the difference between the two offsets (-0.09 ± 0.03, $p < 0.001$). Both slopes (0.001 ± 0.01, $p = 0.17$) did not significantly deviate from 0. We can conclude that the participants generally found the matching task feasible, as the average satisfaction is quite high, but that changing the illumination and viewpoint conditions significantly decreased the satisfaction ratings.

Sum of weights

Here we analyze the sum of the four weights in the probe—i.e., the sum of the four slider values—per trial.

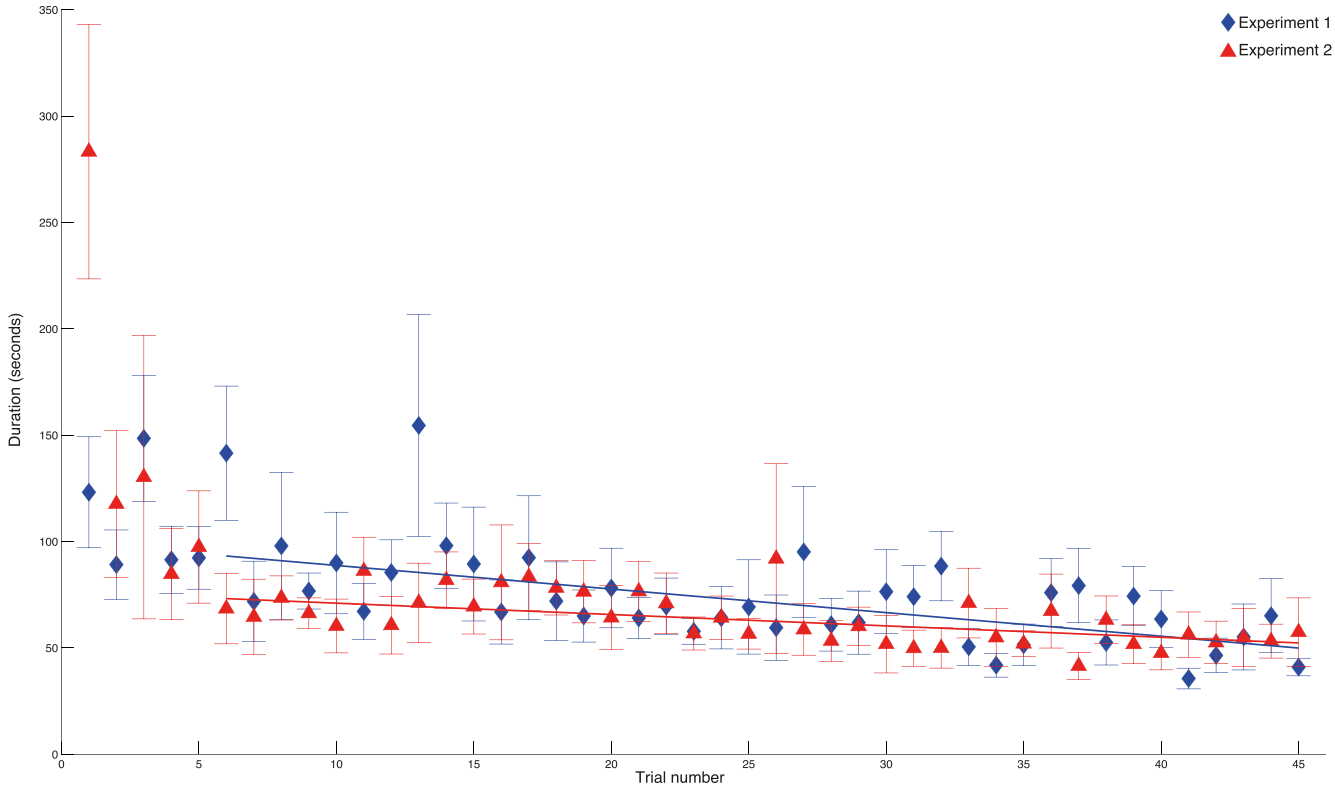

Figure 7. Mean trial duration as a function of trial number, averaged across all observers, for Experiment 1 (blue) and Experiment 2 (red). Error bars of each data point represent one standard error of the sample mean.

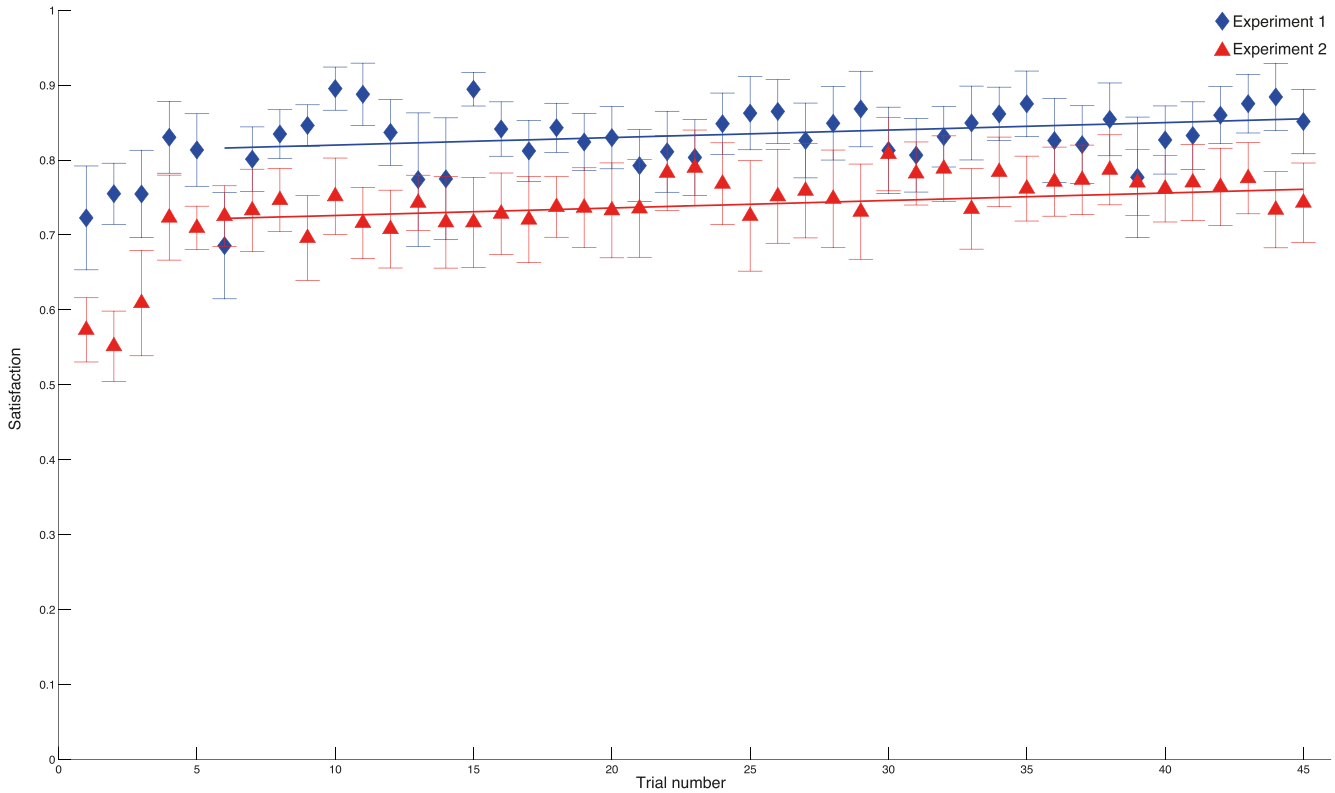

Figure 8. Average satisfaction ratings over all observers as a function of trial number, for Experiment 1 (blue) and Experiment 2 (red). Error bars of each data point represent one standard error of the sample mean.

Our interface and mixing algorithm is based on additive mixing. The sums theoretically can vary from 0 to 4.8 (each slider ranges from 0 to 1.2). However, if the observers were to adjust the image as a partitive mixture constraining the overall brightness of the probe, the sum would be 1 (Griffin, 1999). Similar to what we did when analyzing the durations and the satisfaction ratings, we considered only the last 40 trials per experiment for all eight observers, so in total there were 320 values per experiment. We found that the averages of these sums were 1.13 ± 0.06 and 1.18 ± 0.09 in Experiments 1 and 2, respectively. Because of the randomly generated offsets $\{x_m, x_v, x_s, x_g\}$ in the stimuli, the sums of the weights in the stimuli were very close to 1 but not exactly equal to 1. The averages of the sums in the stimuli were actually 1.00 ± 0.01 in Experiment 1 and 0.99 ± 0.00 in Experiment 2. We calculated the differences between the sums in the probe and the sums in the stimuli and found that these differences significantly deviated from 0 (one-sample t test, $p < 0.001$ for both experiments), with the sums of the weights in the probe being larger than those of the stimuli in both experiments. We also found a significant difference between the two experiments (paired two-sample t test, $p < 0.001$), with the average sum of the weights in the probe of Experiment 2 being larger than that of Experiment 1.

Interactions between scattering modes

In Figure 9 we visualized the interactions between each combination of two scattering modes by means of ellipses representing 1 *SD* of bivariate normal distributions fitted to the 24 data points (8 observers × 3 repetitions) for each stimulus. Every data point represents the settings of two of the four sliders in the probe in one trial. For clarity of presentation the data points themselves were rendered invisible in the plots. Each subplot contains 6 ellipses, which are the results of three different weight combinations in the stimuli in the two experiments. The crosses depict the corresponding stimulus weight combinations. This provides a means to visualize the extent to which participants would trade off—or confuse—the weights of different reflectance modes.

To give an example, in the top left subplot the blue ellipses depict the variations of the weights of the matte and velvety modes in the probe for matches to stimulus number 5 (half matte and half velvety in the stimulus, as in Table 1 and Figure 4). The solid plot represents the result in Experiment 2, which is centered close to the veridical value (the blue cross). The dashed plot represents the result in Experiment 1, which is slightly shifted upward—i.e., in these trials the matte slider was set around its veridical value, while the velvety slider setting was set larger than the weight of the velvety mode in the stimulus. This indicates that in our office lighting, the half-matte, half-velvety mixture was perceived as a match to mixtures of half-matte and more-than-half-velvety components.

Another way of interpreting Figure 9 is to look at how the ellipses are oriented and shifted from their veridical centers. To be more specific, in both Experiment 1 (dashed lines) and Experiment 2 (solid lines), the matte and specular contributions strongly interacted with each other, as seen by the ellipses oriented and shifted diagonally in the middle left subplot. For the velvety and matte (top left) and velvety and specular contributions (middle), we also find diagonal shifts for Experiment 2, while for Experiment 1 there are primarily horizontal or vertical shifts. The glittery contributions were all set around their veridical values in both experiments, and the ellipses in the three subplots at the bottom primarily shifted horizontally from their veridical centers. To conclude, in Experiment 1 we found interactions primarily between the matte mode and the specular mode. In Experiment 2 the matte, velvety, and specular modes interacted strongly with each other. The glittery mode remained quite independent in both experiments.

Validation study (optical mixing with rendered images)

To cross-validate the method, we conducted an additional experiment with MatMix 1.0 in which we used computer-rendered images as the basis images for the mixtures of the stimuli and the probe. In Figure 10 we show the basis images of Experiment 3. To generate these basis images, we built a 3-D model of a bird-shaped object in Blender (Figure 11). We then applied four different "materials" to the object in MaxwellRender (Figure 12). The parameters of the materials in MaxwellRender can be obtained from its online material library. We carefully adjusted the parameters of the four materials to represent matte, velvety, specular, and glittery finishes. High-dynamic-range image-based lighting was used as the illumination environment in rendering. For the basis images of the probe, we used Debevec's "Grace Cathedral" environment map (Debevec, 1998). For the basis images of the stimuli, we used Debevec's "Eucalyptus Grove" environment map.

Experiment 3 was conducted at the University of Giessen, Germany. Five paid, inexperienced participants with normal or corrected-to-normal vision participated in the experiment. Participants read and signed a consent form before the experiment. The experiment was done in agreement with the Declaration of Helsinki and local ethical guidelines.

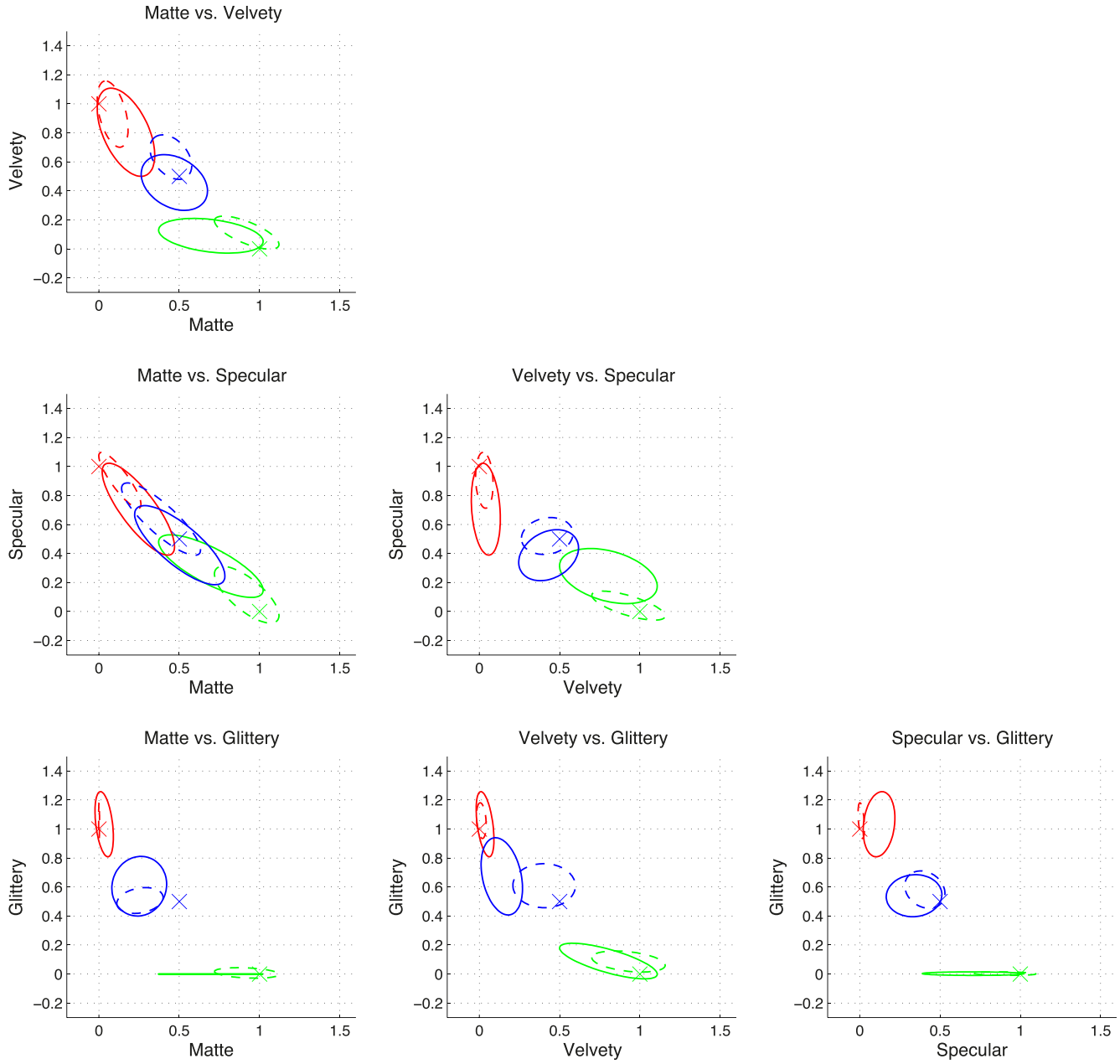

Figure 9. A visualization of the interactions between each combination of two scattering modes. Different colors correspond to different weight combinations in the stimuli, which are depicted by the crosses. The ellipses represent 1 *SD* of bivariate normal distributions fitted to the data. Dashed plots represent the data of Experiment 1 and solid plots represent the data of Experiment 2.

We asked the observers to perform the matching task with MatMix 1.0 using the rendered images in Figure 10 instead of the photographs in Figure 3. Observers spent 50–100 s per trial, which is similar to the durations in Experiments 1 and 2. The satisfaction ratings were 0.76 on average, which is similar to what we found in Experiment 2. Thus, using renderings as the basis images in MatMix 1.0 does not influence the time costs or the satisfaction ratings of observers in the matching experiment.

The linear factor matrix **A** for Experiment 3 is shown in Table 3. It is very close to an identity matrix, except for the values that represent the perception of the velvety mode. The ratios between the sum of the diagonal values and all values for the five observers were 0.79, 0.71, 0.62, 0.70, and 0.60 ($M = 0.68$, $SD = 0.08$), and thus far above chance level (0.25). The nondiagonal values, specifically 0.46 for $\{w_{\mathrm{m}}, w'_{\mathrm{v}}\}$ and 0.35 for $\{w_{\mathrm{v}}, w'_{\mathrm{m}}\}$, indicate that the perception of the velvety mode strongly interacted with the perception of the matte mode in Experiment 3. The residuals were all close to 0. The averages of the absolute value of the residuals were 0.16, 0.12,

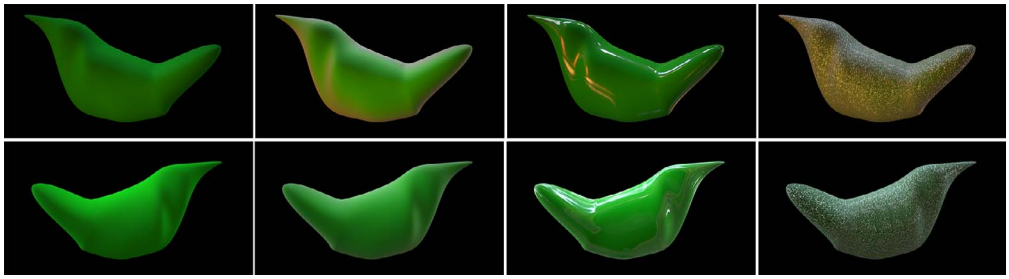

Figure 10. The basis images for Experiment 3. The images in the top row were used as basis images for the probe. The images in the bottom row were used as basis images for the stimuli. From left to right, columns represent matte, velvety, specular, and glittery modes, respectively.

0.14, 0.16, and 0.16 for the five observers, and thus similar to those of Experiment 2.

To sum up, we find that (a) MatMix 1.0 could be implemented by replacing the basis images of the stimuli and the probe with rendered images and (b) with renderings as the basis images, observers can still perform the task well. However, we found increased interactions between the matte mode and the velvety mode. This probably reflects limitations in the current simulations of such reflectance properties.

General discussion

A major finding in this study is our demonstration that the interface (Figure 2) and the probe enable accurate and robust measurements of material perception. Although observers were asked to manipulate four canonical reflectance modes simultaneously, they could do the task within a reasonable amount of time and felt satisfied about their matching results. Moreover, the general matching results were found to be far

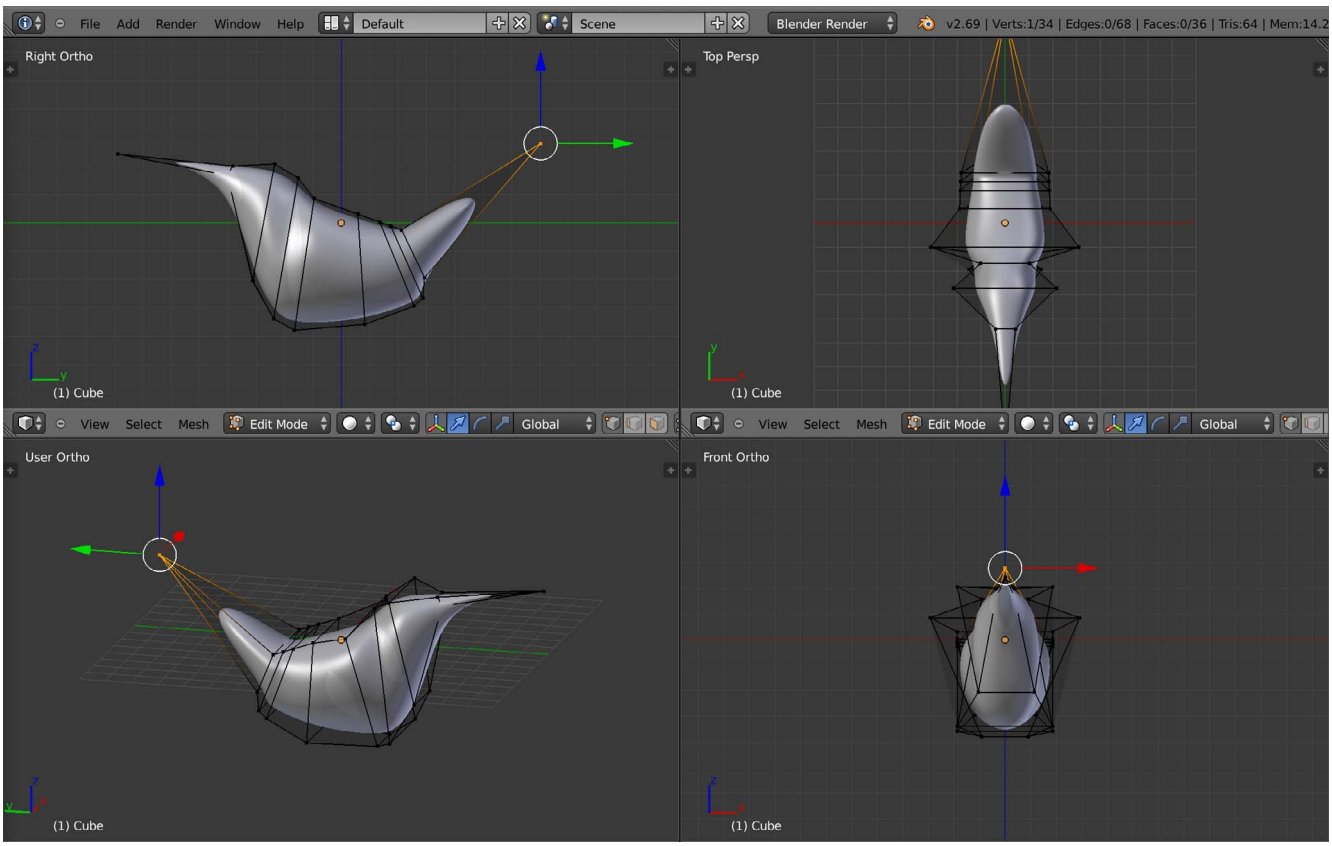

Figure 11. A screenshot of Blender during the 3-D modeling process. The model is mirror symmetric. Note that the model is a simplified version of the shape we used in Experiments 1 and 2.

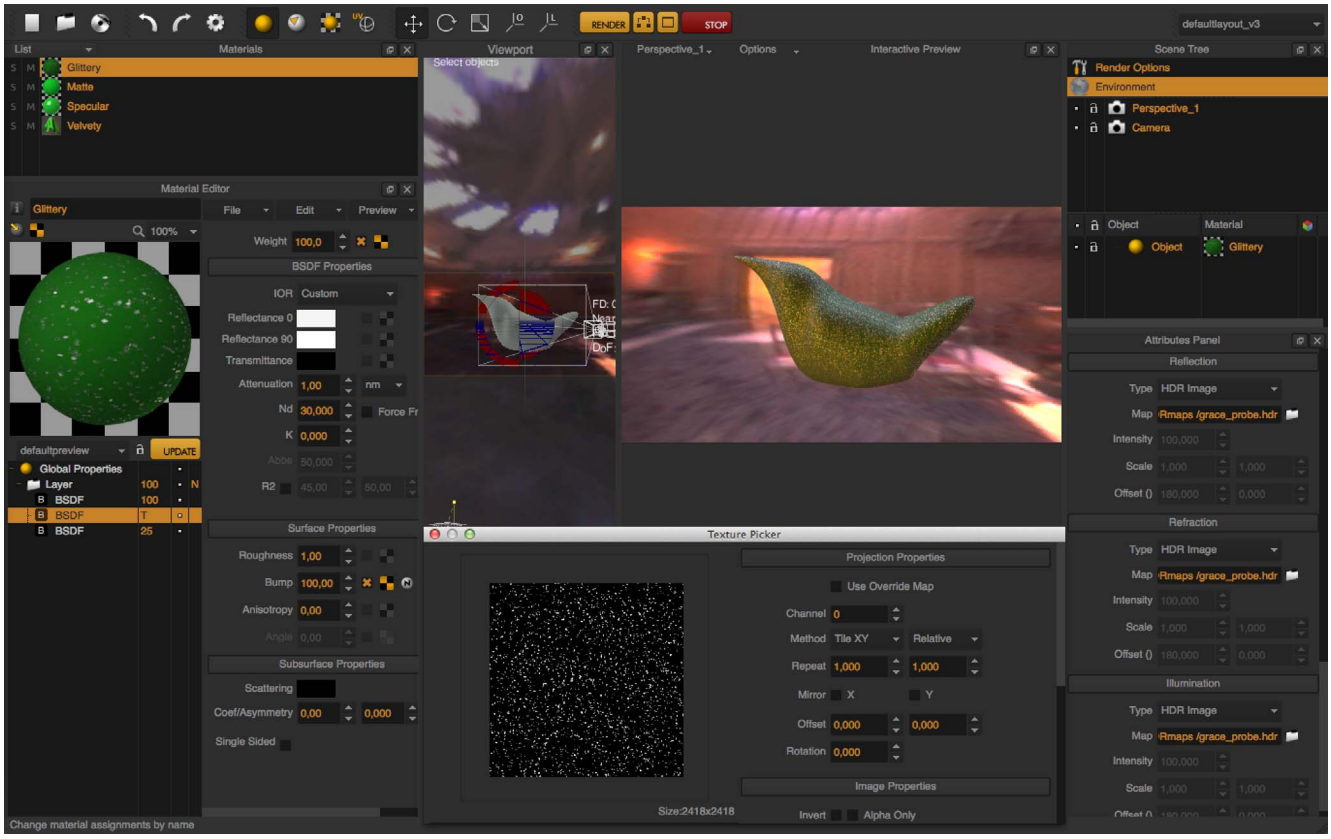

Figure 12. A screenshot of MaxwellRender during the rendering process. In this figure, glittery material was assigned to the object, and the Grace Cathedral image was selected as the environment map.

above chance level in all experiments. In Experiment 1, the illumination and viewpoint conditions were the same for both stimuli and probe images. Observers might have simply compared the two images on a pixel-to-pixel basis and searched for the perfect match. In order to avoid this possibility, we implemented Experiments 2 and 3. Results showed that observers were able to match the materials even if the stimulus and probe images did not correspond. The similarities of the results of Experiments 2 and 3 further convinced us that observers were not just doing a best possible image match. Specifically, in Experiment 3 the light fields of stimulus and probe were quite different, but results were similar to those of Experiment 2. This suggests that the observers were indeed matching perceived materials.

Unlike in Griffin's study (1999), the weights of the material modes do not necessarily add up to 1 when the mixing is performed. Compared to Griffin's partitive mixing method, MatMix 1.0 implements additive mixing. As a result, observers had the freedom to manipulate each of the material modes independently, so that changing the weight of one material does not affect the weight of the others. Theoretically, the sum of the four slider settings could range from 0 to 4.8. We calibrated the luminances of all basis images per set in the same manner so that their relative luminances corresponded with the physical relations. So in order to generate a probe image that was neither too bright nor too dark, the sum of the four weights should be around 1. We found that the sums were somewhat higher than 1, which might be an overall effect of the velvety and glittery basis images having a somewhat lower luminance than the matte and specular basis images. An alternative approach could be equalizing the average luminance of all basis images. However, since the lightness of the resulting images is dependent on material, shape, and illumination, it is more logical to calibrate the physical inputs of the different materials by applying the same camera settings. Nevertheless, the finding that the sum was close to 1 suggests that participants can approximately match the overall magnitude of reflectance (or albedo) while simulta-

	w'_m	w'_v	w'_s	w'_g
w_m	**0.94**	0.46	0.24	−0.07
w_v	0.35	**0.62**	0.20	0.04
w_s	0.05	0.08	**1.01**	−0.07
w_g	0.00	0.03	0.02	**1.00**

Table 3. Linear factor matrix **A** for Experiment 3.

neously reporting precise differences in the quality of the reflectance.

We found systematic shifts in material perception between Experiments 1 and 2 (Figure 9), showing how the perception of material was influenced as the object orientation changed and the lighting changed from office light to studio light. These shifts can be related to the changes of the values in matrix **A**. For example, in Figure 9 the subplot for matte versus specular (middle left) shows that matte–specular interactions increased in Experiment 2 compared to those in Experiment 1, which corresponds to an increase of the nondiagonal values $\{w_m, w_s'\}$ and $\{w_s, w_m'\}$ in Table 2. Material–lighting interactions have been addressed by many researchers (Dror, Willsky, & Adelson, 2004; Fleming et al., 2003; Hunter, 1975; Marlow et al., 2012; Motoyoshi & Matoba, 2012; Olkkonen & Brainard, 2010, 2011; Pont & te Pas, 2006; te Pas & Pont, 2005). In a recent study we combined our canonical material modes with three canonical lighting modes, and in this manner we were able to systematically investigate material–lighting interactions for a broader range of materials and lightings (Zhang, de Ridder, & Pont, 2015). We found systematic effects that depended on lighting *and* material.

However, whether the type of interface we used is the most suitable one remains to be seen. MatMix 1.0 was designed and tested with a limited basis set consisting of four materials. In the future we want to include more material modes to span a wider gamut of the BRDF space, such as backscattering, split-specular scattering, and so on. In order to do this well, we need knowledge about which canonical materials have to be included to cover the perceptual space of natural opaque materials, and about how redundancies between modes could elicit formal ambiguities. But in order to generate such knowledge we would need an extended probe. Moreover, the interface needs to be optimized using knowledge about the perceptual space (e.g., using nonlinear rescaling of the sliders to make the adjustment steps perceptually uniform). These issues are currently still chicken-and-egg problems. We will approach these issues in future research via several iterations in typical design loops (van Boeijen, Daalhuizen, Zijlstra, & van der Schoor, 2013): redesign (extend interface with extra modes), test and evaluate (via formal psychophysical experiments), and adjust the design (on the basis of the experimental outcomes). Other techniques, such as psychophysical scaling methods (Knoblauch & Maloney, 2008; Maloney & Yang, 2003) may also aid with the scaling and selection of the reflectance components.

In the Introduction we made an analogy between optical mixing and painting. In order to analyze our results qualitatively in terms of image characteristics, we did some simple image analysis of the basis images of Experiments 1 (Figure 1), 2 (for results, see Figure 13A), and 3 (for results, see Figure 13B, C). In general, similar to what was shown in Figure 1, we find smooth shading to be typical for the diffuse scattering component (matte material), bright contours for the asperity scattering component (velvety material), highlights at specular points for the forward scattering component (specular material), and bright speckles all over the surface for meso-facet scattering (glittery material). Such key image characteristics may well form the main triggers for general material perception, in a weighted-mixture manner. So across illuminations and object orientations, the diffuse mode typically yields smooth variations, whereas the asperity mode tends to yield bright contours, the specular mode localized highlights, and the glitter mode high-spatial-frequency variations in the image. We find that this indeed allows the user to adjust different aspects of the proximal image. For highlights, many authors have already shown how their specific characteristics influence perception of glossiness (see e.g., Anderson, 2011; Giesel & Zaidi, 2013; Motoyoshi et al., 2007). Perception of velvetiness, glitter, and sparkle concern undeveloped topics. We argue that such understanding of separate modes, together with our findings about how these characteristics combine and interact, will eventually lead to in-depth understanding of any opaque material.

In Experiment 3, strong interactions between the velvety and matte modes were found. In Figure 13C, the rendered basis images of the velvety mode are just very subtly different from those of the matte mode. This might be due to both the rendering functions and the illumination environment (Giesel & Zaidi, 2013).

A connected question is how to represent each scattering mode properly. For instance, it might be better to mix only the highlights of the glossy bird to represent the specular mode, instead of using the green glossy bird which actually also includes a diffuse mode (see specular mode in Figure 13). Similarly, it might be better if only the bright contours were added to represent the velvety mode. This might avoid some interactions between the matte mode and other modes, and thus make them more independent. An analogy is that in a painterly approach, after drawing the contour, the body color is usually painted first in a diffuse manner, after which highlights are added to make the material look glossy or bright contours added to make it look velvety (Wallert, 1999). Additionally, in future studies we want to investigate whether color variation will affect material perception. Currently, green is used disproportionately in material-perception research, for no clear reason (Fleming et al., 2003; Marlow et al., 2012; Marlow & Anderson, 2013). Thus, in a novel version of our MatMix probe we will include color variations accordingly. This will also allow optical mixtures of differently colored modes. For example, specular plastic materials have white highlights, while

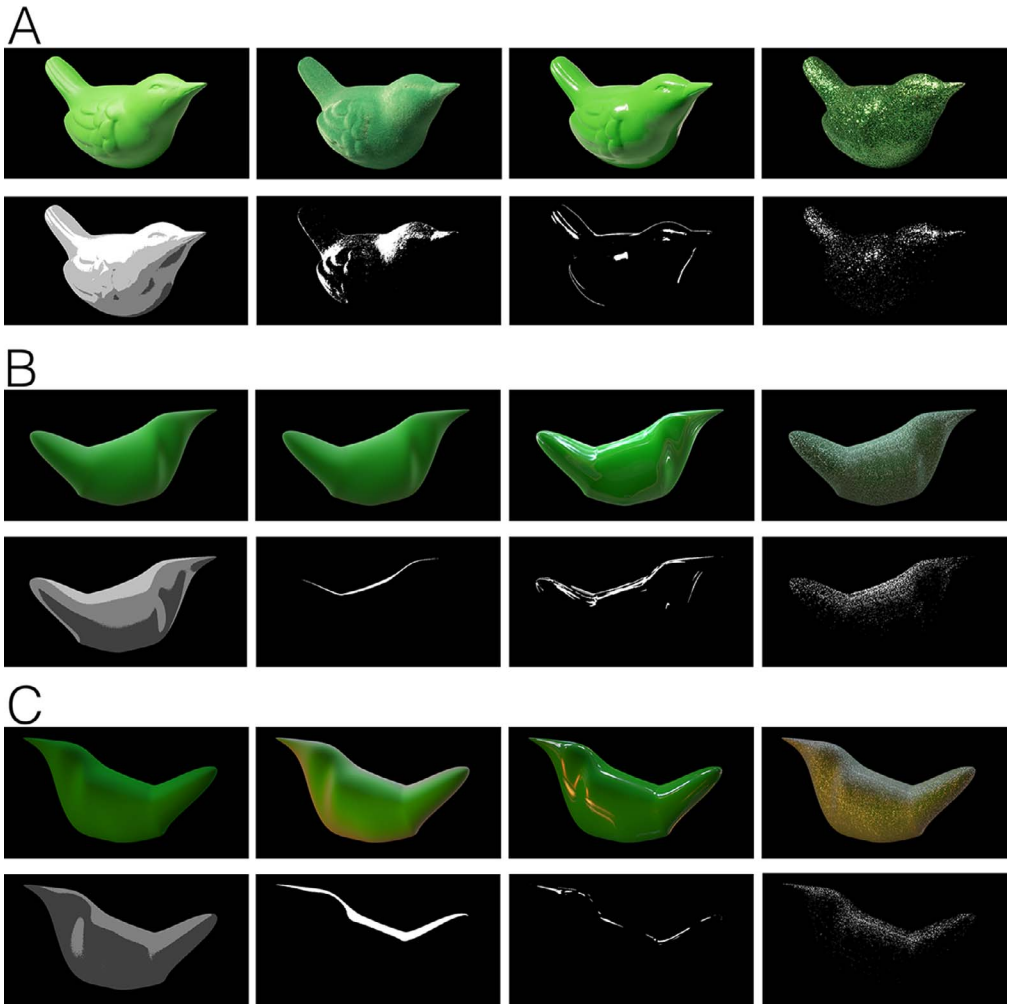

Figure 13. The basis images and some related image characteristics. In each subfigure, the first row shows the basis images; the bottom row shows the prototypical image characteristics of each material: the green channel of the basis images after posterization from 255 to six levels for the matte mode, the red-channel thresholding at a somewhat arbitrary 50% level for the velvety, specular, and glittery modes, respectively. Columns from left to right: Representations of matte, velvety, specular, and glittery modes. (A) Photographed set in studio lighting. (B) Rendered image set in Debevec's "Eucalyptus Grove" image-based lighting. (C) Rendered image set in Debevec's "Grace Cathedral" image-based lighting.

metals have highlights in the color of their diffuse reflectance, and thus we need color variations in the specular modes to cover both plastics and metals.

Many computer-graphics systems also include sliders to allow the user to alter the material parameters. However, MatMix 1.0 is different in several respects from the standard approach found in computer-graphics interfaces. First, we are able to combine (photographs of) real materials, which can exhibit subtle effects that cannot yet be modeled by computer graphics. Second, even experts find computer-graphics interfaces sometimes overwhelming, especially when there is no real-time feedback on appearance. In typical computer-graphics interfaces there are a large number of parameters to adjust, and it is often not intuitive how they are related to the proximal-image result. In contrast, in our approach the basis is limited to a smaller number of canonical visuals, and we have shown that the task is natural and intuitive for inexperienced observers. Third, the purpose of the method is to probe perceptual judgments rather than to design materials from scratch. Thus, the observer will typically have a reference object whose appearance they are trying to match. Finally, the bases are selected to provide perceptually intuitive means for altering proximal-image properties rather than parameters of a physical model, which may not have any distinctive perceptual correlate.

Conclusion

We tested a novel approach to probe material perception in a quantitative and purely visual manner.

In the main study (Experiments 1 and 2), we took photos of real materials under different illuminations and implemented them as basis images in MatMix 1.0. In an additional study (Experiment 3), we rendered four materials similar to the ones we used in the first two experiments and used them as basis images to perform another matching experiment with our probe. We found that (a) no matter how difficult the task was or how satisfied participants were, it cost them on average around the same amount of time per matching trial; (b) the participants were matching the probe to the stimuli on the basis of perceived materials instead of simply matching the two images pixel to pixel; and (c) participants performed well above chance level. In conclusion, it was found that the participants were able to handle the MatMix 1.0 interface well, and our MatMix 1.0 probe was shown to form a robust and intuitive method to test visual material perception. We believe that our painterly optical-mixing approach is promising, because it reflects how weighted mixtures of key ingredients for material representations can trigger our perceptions.

Keywords: material perception, material probe, MatMix 1.0, BRDF, reflectance

Acknowledgments

This work has been funded by the EU FP7 Marie Curie Initial Training Networks (ITN) project PRISM, Perceptual Representation of Illumination, Shape and Material (PITN-GA-2012-316746). Special thanks to pi-lab members Maarten Wijntjes, Ling Xia, and Tatiana Kartashova for all the helpful discussions; to Jan Jaap van Assen, Steven Cholewiak, Vivian Paulun, Patrick Spröte, and all other people at the University of Giessen for their help during FZ's 9-week PRISM secondment; and to Dario Lanza and Abraham Martin at Next Limit Technologies for the question-and-answer session about the rendering software Maxwell-Render.

Commercial relationships: none.
Corresponding author: Fan Zhang.
Email: f.zhang-2@tudelft.nl.
Address: Perceptual Intelligence Laboratory, Faculty of Industrial Design Engineering, Delft University of Technology, Delft, the Netherlands.

References

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

Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of Vision, 9*(11):10, 1–17, doi:10.1167/9.11.10. [PubMed] [Article]

Belhumeur, P. N., Kriegman, D. J., & Yuille, A. L. (1999). The bas-relief ambiguity. *International Journal of Computer Vision, 35*(1), 33–44.

van Boeijen, A., Daalhuizen, J., Zijlstra, J., & van der Schoor, R. (Eds.). (2013). *Delft design guide: Design methods*. Amsterdam, The Netherlands: BIS Publishers.

Cook, R. L., & Torrance, K. E. (1982). A reflectance model for computer graphics. *ACM Transactions on Graphics, 1*(1), 7–24.

Debevec, P. E. (1998). Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proceedings of SIGGRAPH* (pp. 189–198). New York: ACM Press.

Dror, R. O., Willsky, A. S., & Adelson, E. H. (2004). Statistical characterization of real-world illumination. *Journal of Vision, 4*(9):11, 821–837, doi:10.1167/4.9.11. [PubMed] [Article]

Fleming, R. W. (2012). Human perception: Visual heuristics in the perception of glossiness. *Current Biology, 22*(20), R865–R866.

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

Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision, 3*(5):3, 347–368, doi:10.1167/3.5.3. [PubMed] [Article]

Fleming, R. W., Wiebel, C., & Gegenfurtner, K. (2013). Perceptual qualities and material classes. *Journal of Vision, 13*(8):9, 1–20, doi:10.1167/13.8.9. [PubMed] [Article]

Giesel, M., & Zaidi, Q. (2013). Frequency-based heuristics for material perception. *Journal of Vision, 13*(14):7, 1–19, doi:10.1167/13.14.7. [PubMed] [Article]

van Ginneken, B., Stavridi, M., & Koenderink, J. J. (1998). Diffuse and specular reflectance from rough surfaces. *Applied Optics, 37*(1), 130–139.

Griffin, L. D. (1999). Partitive mixing of images: A tool for investigating pictorial perception. *Journal of the Optical Society of America A, 16*, 2825–2835.

Hapke, B., Nelson, R., & Smythe, W. (1998). The opposition effect of the moon: Coherent backscatter and shadow hiding. *Icarus, 133*(1), 89–97.

Ho, Y. X., Landy, M. S., & Maloney, L. T. (2006). How direction of illumination affects visually

perceived surface roughness. *Journal of Vision, 6*(5): 8, 634–648, doi:10.1167/6.5.8. [PubMed] [Article]

Ho, Y. X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science, 19*(2), 196–204.

Hunter, R. S. (1975). *The measurement of appearance*. New York: John Wiley.

Kim, J., Marlow, P., & Anderson, B. L. (2011). The perception of gloss depends on highlight congruence with surface shading. *Journal of Vision, 11*(9): 4, 1–19, doi:10.1167/11.9.4. [PubMed] [Article]

Klinker, G. J., Shafer, S. A., & Kanade, T. (1987). Using a color reflection model to separate highlights from object color. In *Proceedings - First international conference on computer vision*. New York: IEEE.

Knoblauch, K., & Maloney, L. T. (2008). MLDS: Maximum likelihood difference scaling in R. *Journal of Statistical Software, 25*(2), 1–26.

Koenderink, J., & Pont, S. (2003). The secret of velvety skin. *Machine Vision and Applications, 14*(4), 260–268.

Koenderink, J. J., & Pont, S. C. (2008). Material properties for surface rendering. *International Journal for Computational Vision and Biomechanics, 1*(1), 43–53.

Koenderink, J. J., Van Doorn, A. J., Dana, K. J., & Nayar, S. (1999). Bidirectional reflection distribution function of thoroughly pitted surfaces. *International Journal of Computer Vision, 31*(2–3), 129–144.

Maloney, L. T., & Yang, J. N. (2003). Maximum likelihood difference scaling. *Journal of Vision, 3*(8): 5, 573–585, doi:10.1167/3.8.5. [PubMed] [Article]

Marlow, P. J., & Anderson, B. L. (2013). Generative constraints on image cues for perceived gloss. *Journal of Vision, 13*(14):2, 1–23, doi:10.1167/13.14.2. [PubMed] [Article]

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

Matusik, W., Pfister, H., Brand, M., & McMillan, L. (2003). Efficient isotropic BRDF measurement. In *Proceedings of Eurographics/SIGGRAPH workshop rendering* (pp. 241–248). Aire-la-Ville, Switzerland: Eurographics Association.

Motoyoshi, I., & Matoba, H. (2012). Variability in constancy of the perceived surface reflectance across different illumination statistics. *Vision Research, 53*(1), 30–39.

Motoyoshi, I., Nishida, S. Y., Sharan, L., & Adelson, E. H. (2007, May 10). Image statistics and the perception of surface qualities. *Nature, 447*(7141), 206–209.

Nayar, S. K., & Oren, M. (1995, Feb 24). Visual appearance of matte surfaces. *Science, 267*(5201), 1153–1156.

Newell, M. E., & Blinn, J. F. (1977). The progression of realism in computer generated images. In *Proceedings of the 1977 Annual Conference* (pp. 444–448). New York: ACM.

Nishida, S. Y., Sawayama, M., & Shimokawa, T. (2015). Material-dependent shape distortion by local intensity order reversal. *Journal of Vision, 15*(12): 940, doi:10.1167/15.12.940. [Abstract]

Nishida, S. Y., & Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *Journal of the Optical Society of America A, 15*(12), 2951–2965.

Olkkonen, M., & Brainard, D. H. (2010). Perceived glossiness and lightness under real-world illumination. *Journal of Vision, 10*(9):5, 1–19, doi:10.1167/10.9.5. [PubMed] [Article]

Olkkonen, M., & Brainard, D. H. (2011). Joint effects of illumination geometry and object shape in the perception of surface reflectance. *i-Perception, 2*(9), 1014–1034.

Oren, M., & Nayar, S. K. (1995). Generalization of the Lambertian model and implications for machine vision. *International Journal of Computer Vision, 14*(3), 227–251.

te Pas, S. F., & Pont, S. C. (2005). A comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres. In *Proceedings of the 2nd symposium on applied perception in graphics and visualization* (pp. 75–81). New York: ACM.

Pellacini, F., Ferwerda, J. A., & Greenberg, D. P. (2000). Toward a psychophysically-based light reflection model for image synthesis. In *Proceedings of the 27th annual conference on computer graphics and interactive techniques* (pp. 55–64). New York: ACM Press/Addison-Wesley Publishing Co.

Phong, B. T. (1975). Illumination for computer generated pictures. *Communications of the ACM, 18*(6), 311–317.

Pont, S. C., Koenderink, J. J., Van Doorn, A. J., Wijntjes, M. W. A., & Te Pas, S. F. (2012). Mixing material modes. *Proceedings of SPIE, 8291*, 82910D, doi:10.1117/12.916450.

Pont, S. C., & te Pas, S. F. (2006). Material–illumination ambiguities and the perception of solid objects. *Perception, 35*(10), 1331–1350.

Pont, S., te Pas, S., & Wijntjes, M. (2014). Qualities of

optically mixed real materials and photographs—Towards a material probe. *Journal of Vision, 14*(10): 458, doi:10.1167/14.10.458. [Abstract]

Radonjić, A., Cottaris, N. P., & Brainard, D. H. (2015). Color constancy supports cross-illumination color selection. *Journal of Vision, 15*(6):13, 1–19, doi:10.1167/15.6.13. [PubMed] [Article]

Shafer, S. A. (1985). Using color to separate reflection components. *Color Research & Application, 10*(4), 210–218.

Sharan, L., Rosenholtz, R., & Adelson, E. H. (2014). Accuracy and speed of material categorization in real-world images. *Journal of Vision, 14*(9):12, 1–24, doi:10.1167/14.9.12. [PubMed] [Article]

Torrance, K. E., & Sparrow, E. M. (1967). Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America A, 57*(9), 1105–1112.

Torrance, K. E., Sparrow, E. M., & Birkebak, R. C. (1966). Polarization, directional distribution, and off-specular peak phenomena in light reflected from roughened surfaces. *Journal of the Optical Society of America A, 56*(7), 916–924.

Vangorp, P., Laurijssen, J., & Dutré, P. (2007). The influence of shape on the perception of material reflectance. *ACM Transactions on Graphics, 26*(3), 1–9.

Wallert, A. (1999). *Still lifes: Technique and style—The examination of paintings from the Rijksmuseum*. Amsterdam: Rijksmuseum.

Ward, G. J. (1992). Measuring and modeling anisotropic reflection. *Computer Graphics, 26*(2), 265–272.

Wiebel, C. B., Toscani, M., & Gegenfurtner, K. R. (2015). Statistical correlates of perceived gloss in natural images. *Vision Research, 115*(B), 175–187.

Wijntjes, M. W., & Pont, S. C. (2010). Illusory gloss on Lambertian surfaces. *Journal of Vision, 10*(9):13, 1–12, doi:10.1167/10.9.13. [PubMed] [Article]

Wolff, L. B. (1994). Relative brightness of specular and diffuse reflection. *Optical Engineering, 33*(1), 285–293.

Xiao, B., & Brainard, D. H. (2008). Surface gloss and color perception of 3D objects. *Visual Neuroscience, 25*(3), 371–385.

Zhang, F., de Ridder, H., & Pont, S. (2015). The influence of lighting on visual perception of material qualities. *Proceedings of SPIE, 9394*, 93940Q, doi:10.1117/12.2085021.