Toward a Perceptually Based Metric for BRDF Modeling

Adria Fores, James Ferwerda, and Jinwei Gu
Munsell Color Science Laboratory, Center for Imaging Science, Rochester Institute of Technology; Rochester, NY, USA

Abstract

Measured materials are used in computer graphics to enhance the realism of synthetic images. They are often approximated with analytical models to improve storage efficiency and allow for importance sampling. However, the error metrics used in the optimization procedure do not have a perceptual basis and the obtained results do not always correspond to the best visual match. In this paper we present a first steps towards creating a perceptually-based metric for BRDF modeling. First, a set of measured materials were approximated with different error metrics and analytical BRDF models. Next, a psychophysical study was performed to compare the visual fidelity obtained using different error metrics and models. The results of this study show that the cube root metric leads to a better perceptual approximation than other RMS based metrics, independently of the analytical BRDF model used. More benefit of using the cube root metric compared to the RMS based metrics is obtained for sharp specular lobes, and as the specular lobe broadens the benefit of using the cube root metric decreases. The use of the cube root error metric will improve the visual fidelity of renderings made using BRDF approximations and expand the usage of measured materials in computer graphics.

Introduction

Realistic material appearance modeling and rendering is an important but challenging problem in computer graphics, with many applications such as movie industry, advertising, video games, and virtual reality.

Due the complexity of modeling materials empirically, the data-driven approach has been successfully used in order to improve the representation of those materials, where measured materials are directly used for rendering. In order to increase storage efficiency and allow for importance sampling the material measurements are commonly approximated with analytical models.

An error metric is used to guide the optimization procedure into the best approximation of a measured material. However, the obtained results do not always correspond with the best visual match because the error metrics currently used do not have any perceptual basis.

A similar effect can be seen in color science, where the minimization of the RMS spectral difference does not correlate with the color difference minimization. This result is the consequence of not taking the observer into account. However, when the observer is considered by using the color matching functions and a *uniform* color space, the color difference can be correctly approximated and minimized. The challenge in this case is that no perceptual metrics exist to compare measured and approximated materials.

This paper presents the first steps towards creating a perceptually based metric for BRDF modeling. A psychophysical study was performed to compare the visual fidelity of images rendered using different error metrics and models for a set of materials. The results of this study show that the cube root metric leads to a better perceptual approximation than other RMS based metrics, independently of the analytical BRDF model used. More benefit of using the cube root metric compared to the RMS based ones is obtained for sharp specular lobes, and as the specular lobe broadens the benefit of using the cube root metric decreases. The use of the cube root error metric will improve the visual fidelity of renderings made using BRDF approximations and expand the usage of measured materials in computer graphics.

Related work

The BRDF (Bidirectional Reflectance Distribution Function) is a 4-Dimensional function that describes how light is scattered by a surface. It is defined by the following equation:

$$f(\omega_i, \omega_o) = \frac{L(\omega_o)}{E(\omega_i)} \tag{1}$$

where E defines the irradiance due the light source in the direction defined by ω_i, and L defines the radiance of a surface in the direction ω_o, where the directions are defined in spherical coordinates.

In order to understand which analytic BRDF models best approximate measured BRDF data, the 100 materials of the MERL database were approximated with 7 different analytical BRDF models in [11]. This study provided insights about the expressivity of the different analytical BRDF models. A key aspect in the approximation step is the error metric selection. In this case, the objective function used in the optimization step was the minimization of the RMS error metric weighted by the cosine of the incident light direction and the solid angle, in order to compensate for the reflectance increase towards grazing angles and the measurement sampling. The authors emphasize that the best fit according to their metric does not always correspond to the best visual match, which they found to be highly dependent on scene geometry and illumination.

A method to navigate through a uniform material appearance space was created in [12]. The pixel-by-pixel differences between synthetic images generated with different BRDF models were used to create this space. A precomputation step was used to generate all the images used in an interactive interface to aid the material design. This technique would probably give a good performance if used as error metric during the optimization process, but it would require the generation of a synthetic image in each iteration step of the optimization process, making it computationally expensive.

In [13], a perceptual space of glossy materials represented by the Ward BRDF model was created. Two perceptual gloss dimensions were defined in this space: contrast gloss and distinctness-of-image gloss. These dimensions were used to reparameterize

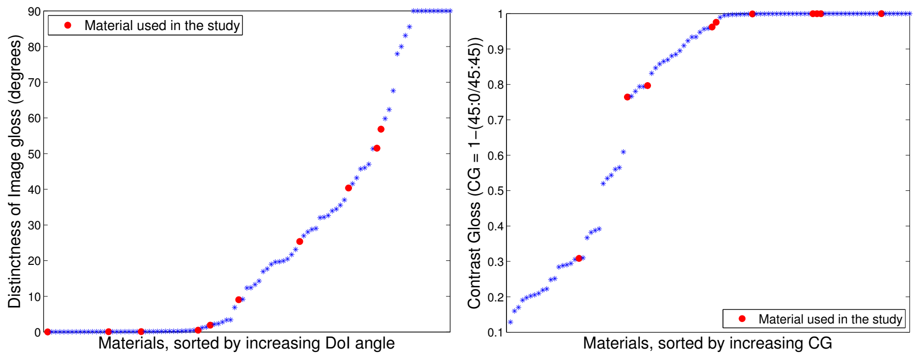

Figure 1. *From left to right, distinctness of image gloss and contrast gloss of the MERL database materials. Materials used in this study are shown in red.*

the Ward reflectance model parameters to a perceptually uniform space. This work is not directly applicable to the creation of an error metric to compare measured data and analytical models, because this perceptually uniform space is only defined for materials represented with the Ward BRDF model.

Overview

In this paper, the perceptual fidelity of different error metrics commonly used in BRDF modeling is studied. A paired comparison psychophysical experiment with reference was performed to evaluate the perceptual fidelity of the metrics. The reference was a rendered image using measured data and the observer was asked to select the closest approximation to the reference of the two test stimuli presented in each trial.

Multiple factors influence the visual fidelity of an approximation, not only the error metric. The materials approximated, the analytical BRDF models used, the lighting and geometry used in the evaluation scene, and the optimization procedure are key elements involved in the approximation and its evaluation.

For that reason, a set of materials and analytical BRDF models were also studied for each error metric and a scene that maximized the material discrimination was defined based on previous perceptual studies.

Stimuli

In this section the different components used to generate the images used in the psychophysical experiment are described.

Material data set

A set of 10 materials of the MERL Database [9] were used in this study. This database includes isotropic measurements of 100 materials, including painted surfaces, fabrics, metals, and plastics. It was created by imaging a sphere of a given material lit by a point light source with a camera for a dense set of incident directions. For each incident direction, a set of images with different exposures were merged to obtain an HDR image. This image-based method allows high angular resolution measurements and many radiance samples are recored in each image.

The selection of 10 materials used in Vangorp et al. [18] was used in this study, except for the *copper* material that was not found on the database and it was substituted for the *nickel*. The material selection spans different types of materials, colors, and gloss levels (*gold metallic paint2, aluminium, blue acrylic, alum bonze, nylon, nickel, blue metallic paint, pearl paint, light red paint, and silver metallic paint*), a subset of which is shown in the first row of Figure 3. To visualize the distribution of these materials, the distinctness of image gloss (DoI) and the contrast gloss (CG) of the 100 materials of the MERL Database are shown in Figure 1, where the selected materials are shown in red. The DoI is computed as the angle between the specular direction at 30° and a measurement at 0.3° from the specular direction. This angle is small for sharp specular lobes, and it increases when the specular lobe broadens. The perceptual spacing of the selected materials is well balanced as more materials are selected with small DoI angles, which is the region where we are more sensitive to (see Figure 1 left).

BRDF Models

To be able to generalize the fidelity obtained with different error metrics, three analytical BRDF models commonly used in the literature were selected for this study. The Ashikhmin-Shirley [2] and the Cook-Torrance [3] BRDF models were selected because they are widely used and also provided the best performance in [11]. The Ward BRDF model [19] was also used in this study, due to its wide use in vision science and perceptually based material modeling experiments [5, 13, 18]. The models' equations presented in [11] were implemented in Matlab to perform the optimization process and in the rendering engine used to generate the synthetic images:

- Ashikhmin-Shirley
$$K = \frac{m+1}{8\pi}\frac{(n\cdot h)^m}{(\omega_o\cdot h)max((n\cdot\omega_i),(n\cdot\omega_o))}Fresnel(F_0,\omega_o,h) \quad (2)$$
where n is the normal direction, h is the half way vector ($\frac{\omega_i+\omega_o}{2}$), and m models the shape of the specular lobe. The fresnel term is approximated using the Schlicks approximation [17], which depends on the parameter F_0:
$$Fresnel(F_0,\omega_o,h) = F_0 + (1-F_0)\cdot(1-(\omega_o\cdot h))^5 \quad (3)$$
- Cook-Torrance
$$K = \frac{1}{\pi}\frac{DG}{(n\cdot\omega_i)(n\cdot\omega_o)}Fresnel(F_0,\omega_o,h) \quad (4)$$
where the Beckmann distribution is used to represent the normal distribution probability for the micro-facets, D:
$$D = \frac{1}{\alpha^2 cos^4\delta}e^{-[(tan\delta)/\alpha]^2}, \quad \delta = acos(n\cdot h) \quad (5)$$
α describes the surface roughness of the material, and G is the geometric attenuation term, which describes the masking and shadowing effects between the microfacets.
$$G = min\left(1, \frac{2(n\cdot h)(n\cdot\omega_o)}{(\omega_o\cdot h)}, \frac{2(n\cdot h)(n\cdot\omega_i)}{(\omega_o\cdot h)}\right) \quad (6)$$

- Ward

$$K = \frac{1}{\sqrt{(n\cdot\omega_i)(n\cdot\omega_o)}} \cdot \frac{e^{-\tan^2\delta/\alpha^2}}{4\pi\alpha^2} \quad (7)$$

where α controls the width of the lobe.

Error Metrics

Error metrics represent the difference between a measured material and an approximation, as color differences represent the difference between two colors, and its minimization leads to the best approximation of a measured material. The error metric is computed across each pair of incident and outgoing directions for each color channel for the measured and approximated material. Three error metrics used in the literature were evaluated in this study: the root mean square error (RMS), the RMS weighted by the cosine of the incident direction, and the cube root of the cosine weighted metric:

- Root mean square (RMS)

$$E = \sqrt{\frac{\sum (M(\omega_i,\omega_o) - A(\omega_i,\omega_o,p))^2}{n}} \quad (8)$$

- Cosine weighted RMS

$$E = \sqrt{\frac{\sum (M(\omega_i,\omega_o)cos\theta_i - A(\omega_i,\omega_o,p)cos\theta_i)^2}{n}} \quad (9)$$

- Cube root cosine weighted RMS

$$E = \sqrt{\frac{\sum \left((M(\omega_i,\omega_o)cos\theta_i - A(\omega_i,\omega_o,p)cos\theta_i)^2\right)^{1/3}}{n}} \quad (10)$$

where the difference between the measured BRDF M and the approximation A obtained using a given BRDF model with the parameters p is computed across the n pairs of incident and outgoing directions.

The RMS is the simpler error metric, in which the distance between each of the points of the measured data and the approximation obtained with the analytical BRDF model is computed.

The weighting factor used in the cosine weighted RMS is added to compensate for the reflectance increase towards the grazing angles when the incident direction goes from the normal direction at 0° to 90° in θ_i.

RMS metrics tend to overemphasize the importance of the BRDF peaks in the mirror direction and deemphasize the off-peak values. For that reason, the empirically derived cube root metric is sometimes used for trying to correct this effect.

There is no consensus in the literature about metric selection, and every researcher tends to apply corrections from their previous experience. In Ngan et al. [11], the log and cube root compressive metrics were not used as the authors found that the specular highlights became too blurry, so they used the cosine weighted RMS with a solid angle correction term. On the other hand, Matusik [7] emphasized the need of compression to obtain a good approximation of glossy materials and used a log function as error metric. The cube root metric was used in this paper as the log function behaves badly near zero, and as it has been commonly used in the literature. The exploration of compressive metrics with exponents similar to the ones used in gamma functions could be an interesting avenue of future work if those are found to better model the perception of material differences.

Two major types of corrections are usually applied in the error metrics: physical and empirical. Physical corrections try to correct or normalize for physical changes in the light-material interaction or measurement process. For example, the above mentioned reflectance increase towards grazing angles or solid angle corrections applied for different directions. Empirical corrections are derived from trial and error, and do not have any physical basis. For example, the cube root or log compression in the data are a form of empirical correction.

It's important to note that no special consideration is commonly given to the information from multiple color channels. Thus, the information of different channels is considered as another set of points in this work, without the use of any color difference equation. This fact may lead to hue errors that can become highly perceptible on rendered images.

Table 1. **Starting values, and lower and upper boundaries for the parameters used in the non-linear optimization procedure.**

		Ward	Cook-Torrance		Ashikhmin-Shirley	
	ρ_s	α	m	F_0	m	F_0
Starting	0.5	0.01	0.02	0.3	5	0.3
Lower	0	0.001	0.001	0.02	0.001	0.02
Upper	1	0.5	1	1	50000	1

Fitting BRDFs

A core task of this project is the fitting process, in which the parameters of a BRDF model are optimized to minimize a given error metric for a given material. The analytical BRDF model parameters are highly non-linear and the result obtained depends on the initial values used for the optimization process.

We used a diffuse lobe and two specular lobes to approximate each measured material. Ngan et al. [11] stated that the fit quality was much improved with the addition of the second specular lobe, probably because of the multiple layer finish of the materials of the MERL database.

The analytical form used to approximate measured materials is the following:

$$K = \rho_d \; diffuse + \sum_{i=1}^{2} \rho_s \; specular(p) \quad (11)$$

where ρ_d is the diffuse albedo (RGB scalars), $diffuse$ is a Lambertian lobe, $specular$ is a particular analytical BRDF model (Equation 2, 4 or 7), ρ_s is the specular albedo (RGB scalars), and p are the parameters of the $specular$ analytical BRDF model. Note that the same analytical BRDF model is used for both specular lobes, but each specular lobe has different ρ_s and p.

We used the following optimization technique: First, the diffuse albedo was set using a 45:0 measurement. This helps to increase the stability of the optimization process, as less parameters need to be optimized. Then, one specular lobe having a single scalar as a specular albedo and the BRDF model parameters were non-linearly optimized. The same procedure is then performed again adding the second specular lobe, and using as initial values for the first specular lobe the ones previously found. Finally, the specular albedos are converted from single scalars to RGB triplets, and are non-linearly optimized while the other BRDF parameters are kept constant.

The stability of the optimization procedure is greatly increased by adding one specular lobe at a time and by initially using a single scalar as specular albedo. A single scalar is used in order to reduce the number of parameters to optimize, and because the objects' highlights are perceived to be of a similar color to the light source, this is a good first approximation. Later on, the specific optimization of the specular albedos allows for slight scale changes for each channel.

The *fmincon* constrained non-linear optimization MATLAB routine was used to perform the optimization step. The initial values, and lower and upper boundaries for each of the parameters optimized were carefully set to span the meaningful range for each parameter (see Table 1).

Scene description and rendering

To generate synthetic images a material, object, and lighting need to be defined. In this case, the materials are either the tabulated measured data of the real physical material or one of the approximations obtained.

In Ngan et al. [11], a sphere and a environment map mainly composed of colored lights were used to visually evaluate the rendering results. The influence of shape on the perception of material reflectance was studied in [18], where the ability to discriminate if two different geometric objects had the same reflectance or not was analyzed in a psychophysical experiment. The fact that every 3D modeling application uses a sphere as a sample material was one of the reasons driving this work, and the authors found that the sphere was one of the least discriminating shapes for judging materials. One of the shapes that gave the best discriminating accuracy was a blob-like shape, which contained both concave and convex regions. This blob-like shape was selected for this study in order to maximize the material discrimination. The Eucalyptus Groove light probe from Paul Debevec was used in this study because it was found to be the environment map with real world statistics providing the best material discrimination in [5]. This light probe also allows to evaluate the color of an object without the need to perform any chromatic adaptation.

The Physically Based Ray Tracer (PBRT) [14] was used to generate the synthetic images with a resolution of 400x400 pixels. This ray tracer natively supports the format of the measured materials' data and takes advantage of parallel execution, which reduces the computational time required.

The global Reinhard et al. [16] HDR tone mapping operator was applied to a composite HDR image containing all the images used in the experiment (parameters: key=0.18, and phi=1.0). Then, a 2.2 gamma correction was applied to the tone mapped images. The implementation provided in [6] was used.

Experiments

The perceptual fidelity of images created using different combinations of materials, models, and metrics were evaluated by performing a two-alternative forced-choice (2AFC) psychophysical experiment with reference. The reference was a rendered image using measured data, and the observer was asked to select the closest approximation to the reference of the two stimuli presented in each trial. The interface used for the experiment can be seen in Figure 2.

The first experiment compared each possible combination of the error metrics and the three analytical BRDF models for each of the ten materials to the reference image.

The reference image was included in the trial selection in a second experiment in order to evaluate the distance between the approximations and the measured data. In this case, the approximations obtained with the three analytical BRDF models using the error metric that gave the best result in the first experiment and the measured data were compared to the reference image. The camera position on the reference image was rotated 15° around the object to avoid pixel-by-pixel comparisons by the observers.

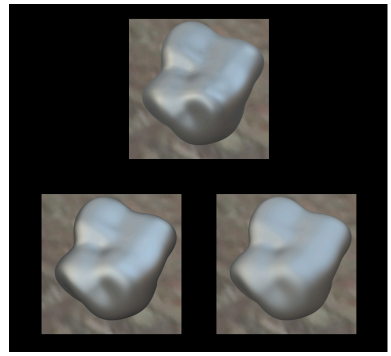

Figure 2. *Interface used for the 2AFC experiment with reference, developed with Psychtoolbox. The reference image is shown on the top, and the two approximations are shown on the bottom.*

A total of 360 trials were done for each of the 15 observers that participated in the first experiment, and 60 trials were done for each of the 20 observers that participated in the second experiment. All the observers had normal color vision and normal or corrected to normal visual acuity.

The experiments were performed in a darkened room with a controlled viewing conditions on a 30-inch Apple Cinema Display. Unfortunately, neither the material measurements nor the environment map were color calibrated, and it was not possible to obtain colorimetric data to input a calibrated display. Hence, only the additivity of the display was evaluated, presenting a good additivity. An avenue of future work would be to obtain accurate colorimetric data for both, the materials and the environment maps using the technique described in [4], which would allow to better asses the goodness of the color approximation. A lower dynamic range display was selected for this experiment to replicate the common viewing conditions in which synthetic images are visualized.

Results
Fitting

The renderings for 7 of the 10 materials can be seen in Figure 3. The use of a compressive metric (i.e. cube root) seems to improve the approximation of high gloss materials over the RMS based metrics (see bottom row of Figure 3). For low gloss, all the metrics seem to produce a similar renderings. The RMS based metrics seem to overfit the specular lobe for high gloss materials. The blue acrylic material was not well approximated for any combination of error metrics and models, the diffuse component was approximated well when the compressive metric was used, but the specular lobe was overfit by the RMS based metrics.

The fidelity of a material approximation is usually evaluated by showing BRDF plots, with the values given by an error metric, and rendered images of the measured data and its approximation.

A disconnect exists between the values given by error metrics and the visual fidelity of an approximation because error met-

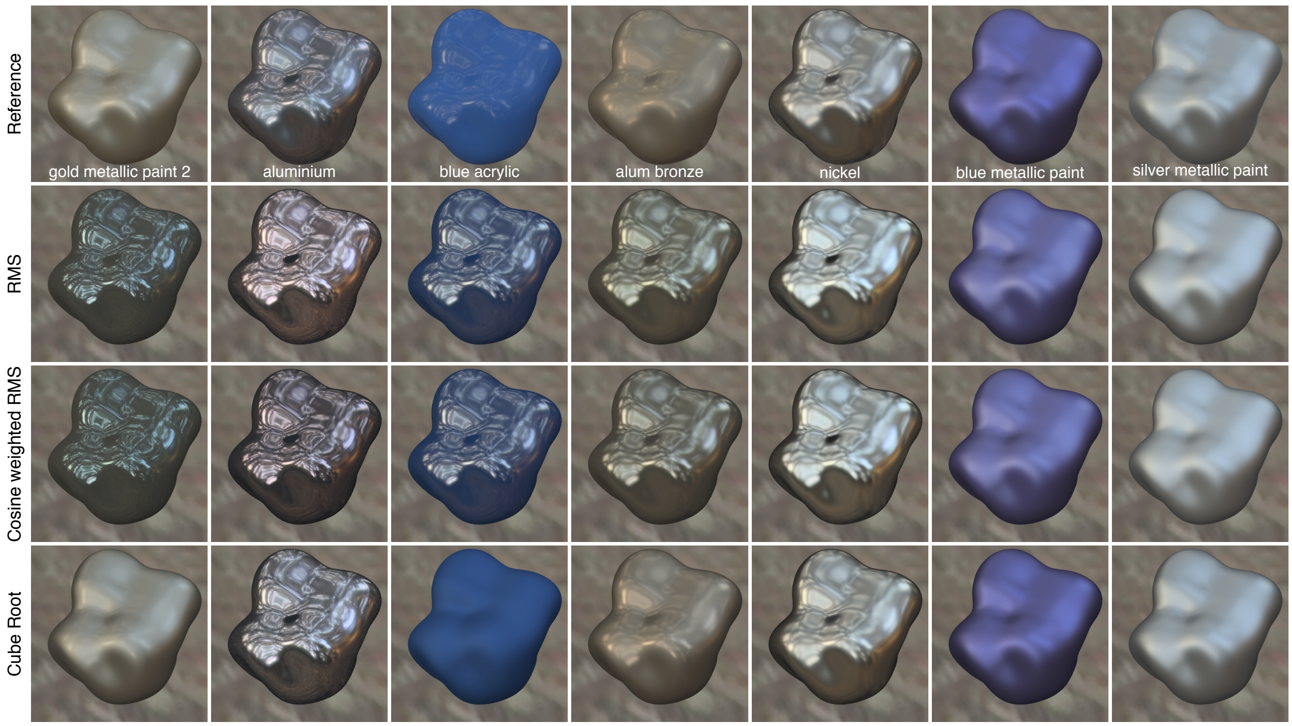

Figure 3. *Reference and approximations obtained for 7 of the 10 materials used in the study using the Ward BRDF model. A better visual fidelity is obtained with the cube root error metric for high gloss materials, while the RMS based metrics seem to over fit the specular lobe. For low gloss, all the metrics seem to produce a similar visual fidelity.*

rics currently used are not perceptually based. For example, an approximation that is off in hue can have a lower error value than another approximation, while the latter may be closer to the measured material if the rendered images are compared.

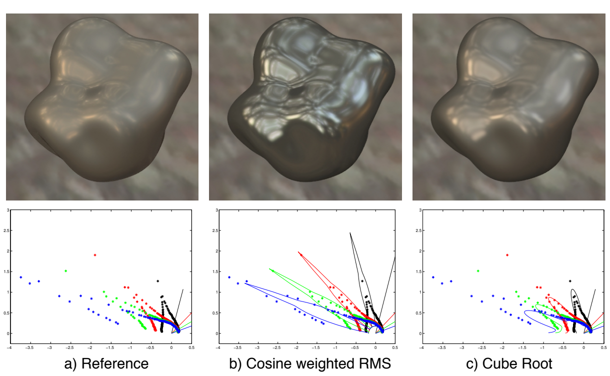

Figure 4. *From left to right, alum-bronze reference material, cosine weighted RMS, and cube root approximations using the Ward BRDF model. The second row shows the cube root compressed BRDF plots with the measured data and its approximations for the green channel and given different incident directions. The evaluation of an approximation using only the BRDF plots can be misleading.*

BRDF plots are commonly used to evaluate the goodness of an approximation, where the in-plane measured and approximated data are displayed. Again, a disconnect exists between BRDF plots and the visual fidelity of an approximation. Figure 4 shows the rendered images and the BRDF plots of a reference material and two approximations. If the BRDF plot is used to evaluate the goodness of the approximation, the cosine weighted RMS metric approximation would be selected as best. However, by looking at the rendered images, it's clear that the metric providing the best visual rendering is the cube root, in spite of the differences seen in the BRDF plots.

Psychophysical experiment

Thurstone's law of comparative judgment (case V) was used to derive interval scales given the data from the psychophysical experiments. The confidence intervals were computed using the empirical formula derived from Monte Carlo simulations of paired comparison experiments in [10].

The interval scales obtained for the first experiment with each material and error metric given a different BRDF model are shown in Figure 5. The materials are sorted by increasing DoI angle. For the Ward BRDF model (Figure 5a), the cube root metric is always preferred by the observers in comparison to the RMS based metrics. The sharper the specular lobe, the more beneficial the use of the cube root metric is. The confusion seen in the blue-acrylic material could be explained with different criteria among observers, where some observers probably gave more weight to the highlights and others to the diffuse component (see Figure 3). Once the specular lobe broadens, the benefit of using the cube root metric decreases, but still better visual fidelity is perceived by the observers when this metric is used. Without being significant, a small benefit is observed if the cosine weighted RMS metric is used in place of the RMS metric for the Ward BRDF model.

The scalings obtained for the Ashikhmin-Shirley and Cook-

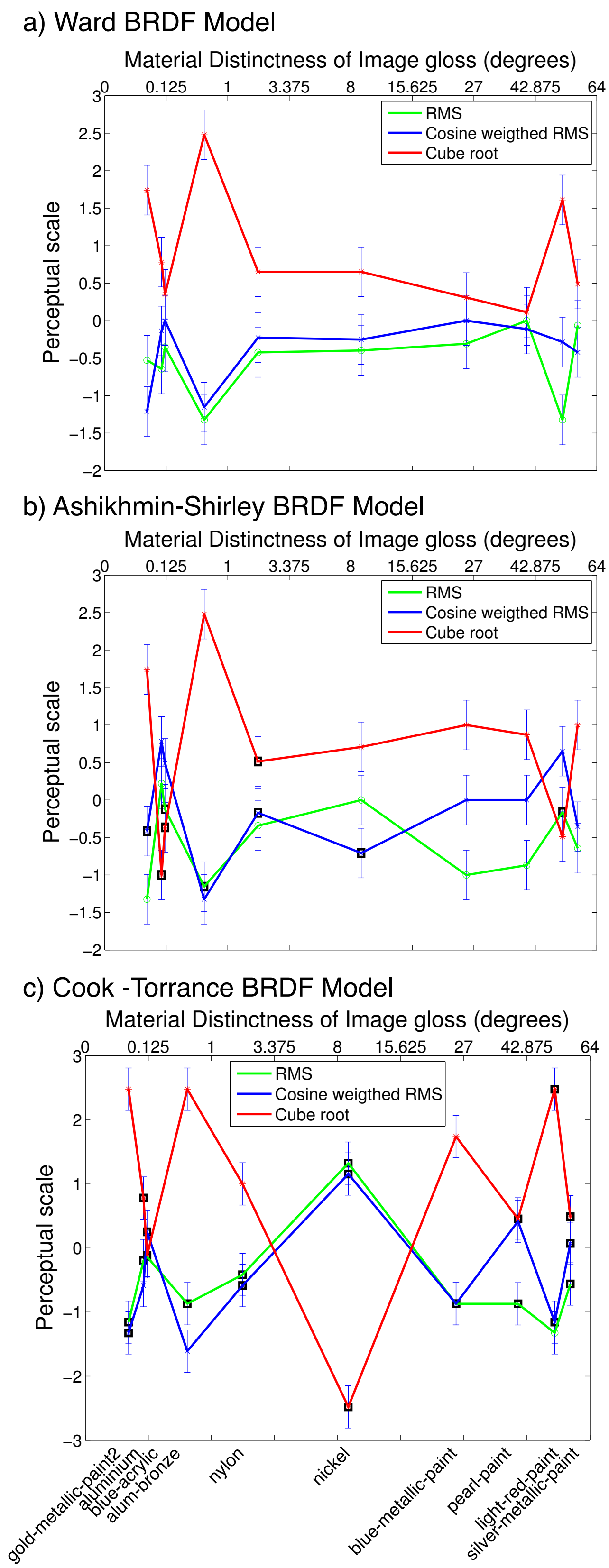

Figure 5. *Error metric interval scaling across materials sorted by increasing DoI angle for the a) Ward, b) Ashikhmin-Shirley, and c) Cook-Torrance BRDF models.*

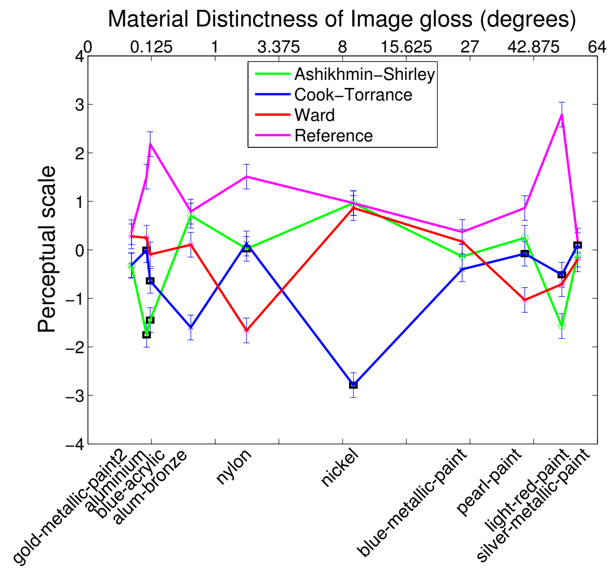

Figure 6. *BRDF models and measured data (Reference) interval scaling across materials sorted by increasing DoI angle for the cube root error metric.*

Torrance BRDF models are shown in Figures 5b and 5c, respectively. For those BRDF models, the cube root metric is almost always preferred by the observers. However, this is not the case for all the materials and some crossovers appear for the lines connecting the error metrics performance. The main reason of the crossovers is the high number of *outliers* obtained through the optimization procedure for those BRDF models. Outliers are those approximations in which a local minimum was reached by the optimization procedure, and are represented with a black rectangle. A simple technique was used to determine when a local minima was found: for each approximation performed, the parameters found for the same material and the same analytical model with the other error metrics were used with the initial error metric. If the use of those parameters with the initial error metric produced a lower error value than the one given in the optimization, a local minima was found. The reason why outliers were found for the Cook-Torrance and the Ashikhmin-Shirley BRDF models is probably because those BRDF models have two parameters that need to be optimized for each specular lobe, while the Ward BRDF model only has one parameter to optimize (see Table 1).

The cube root metric was selected for the second experiment as it was found to give the overall best approximations in the first experiment. The measured data and the three BRDF models using that metric were compared against the reference image. The interval scales obtained are shown in Figure 6. The approximations obtained using the cube root error metric were confused with the measured data for 5 of the 10 materials studied, 4 materials with the Ward model, 3 materials with Ashikhmin-Shirley model, and 1 material with Cook-Torrance model. For some materials, the two former models were able to surpass the Ward model, probably due the better representation of the increased reflection towards the grazing angles provided by the fresnel term incorporated in those models. However, the lack of convergence of those two models did not allow a faithful evaluation of which model better approximates measured data.

Discussion

Our key finding is the higher visual fidelity obtained using the cube root metric compared to the RMS based metrics for the studied materials. The improvement in visual fidelity using the cube root metric compared to the RMS based metrics is higher for sharp specular lobes and decreases as the specular lobe broadens.

The better performance of a compressive metric can be related to perception, where a similar compression is applied to the lightness channel in CIELAB, and tone mapping operators compress HDR images to be displayed on low dynamic range displays. It would be interesting to repeat the experiment using a high dynamic range display, as it is known that limiting the image dynamic range does change the apparent gloss of surfaces depicted in images [15].

In spite of the similar trends obtained for the Cook-Torrance and Ashikhmin-Shirley BRDF models, the high number of outliers when compared to the Ward model limits the generalization of the conclusions that can be drawn from our experiments when other BRDF models are considered. To try to reduce the number of outliers, a simpler optimization technique was performed: multiple sets of starting values for the BRDF models' parameters were generated either randomly or by using a tabular representation, and the set of parameters that gave the lowest error was the one selected. However, the visual fidelity obtained with those techniques and/or the number of outliers obtained were worse than the obtained with the optimization technique previously stated. The global optimization technique in [20] could help to avoid the problems seen while approximating multiple-lobes.

Another approach that we are currently pursuing is the use of hybrid analytical models [1, 8]. One of the problems with the approximation of measured data is that the shape of the analytical BRDF models' lobes is different from the measured data. A hybrid analytical model uses a small set of data points of the measured data to represent the microfacet distribution, which combined with an analytical description of the shadowing and masking terms and the fresnel function successfully approximates measured materials. As this model better approximates the shape of the specular lobe the error metric that would give the best visual fidelity is probably going to be different than the one obtained for analytical BRDF models.

Acknowledgments

This work was supported in part by NSF IIS-1064412 to James Ferwerda, RIT OVPR 15804 to Jinwei Gu, and a gift from the Hewlett-Packard Laboratories.

References

[1] Michael Ashikhmin and Simon Premoze. Distribution-based BRDFs. Technical report, The University of Utah, March 2007.

[2] Michael Ashikhmin and Peter Shirley. An anisotropic phong BRDF model. *J. Graph. Tools*, 5:25–32, February 2000.

[3] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Trans. Graph.*, 1:7–24, January 1982.

[4] Benjamin A. Darling, James A. Ferwerda, Roy S. Berns, and Tongbo Chen. Real-time multi-spectral rendering with complex illumination. In *19th Color and Imaging Conference, San Jose, California, USA*, November 2011.

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

[6] Rafał Mantiuk, Grzegorz Krawczyk, Radosław Mantiuk, and Hans-Peter Seidel. High dynamic range imaging pipeline: Perception-motivated representation of visual content. In *Human Vision and Electronic Imaging XII*, volume 6492, San Jose, USA, February 2007. SPIE.

[7] Wojciech Matusik. *A Data-Driven Reflectance Model*. PhD thesis, Massachusetts Institute of Technology, 2003.

[8] Wojciech Matusik, Boris Ajdin, Jinwei Gu, Jason Lawrence, Hendrik P. A. Lensch, Fabio Pellacini, and Szymon Rusinkiewicz. Printing spatially-varying reflectance. *ACM Trans. Graph.*, 28:128:1–128:9, December 2009.

[9] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Transactions on Graphics*, 22(3):759–769, July 2003.

[10] Ethan D. Montag. Empirical formula for creating error bars for the method of paired comparison. *Journal of Electronic Imaging*, 15(1):010502, 2006.

[11] Addy Ngan, Frédo Durand, and Wojciech Matusik. Experimental analysis of BRDF models. In *Proceedings of the Eurographics Symposium on Rendering*, pages 117–226. Eurographics Association, 2005.

[12] Addy Ngan, Frédo Durand, and Wojciech Matusik. Image-driven navigation of analytical BRDF models. In *Symposium on Rendering*, pages 399–407, Nicosia, Cyprus, 2006. Eurographics Association.

[13] Fabio Pellacini, James A. Ferwerda, and Donald P. Greenberg. Toward a psychophysically-based light reflection model for image synthesis. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 55–64. SIGGRAPH '00, ACM, 2000.

[14] Matt Pharr and Greg Humphreys. *Physically Based Rendering, Second Edition: From Theory To Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2nd edition, 2010.

[15] Jonathan B. Phillips, James A. Ferwerda, and Stefan Luka. Effects of image dynamic range on apparent surface gloss. In *17th Color Imaging Conference*, pages 193–197, November 2009.

[16] E Reinhard, M Stark, P Shirley, and J Ferwerda. Photographic tone reproduction for digital images. *ACM Transactions on Graphics*, 21(3):267–276, 2002.

[17] Christophe Schlick. An inexpensive brdf model for physically-based rendering. *Computer Graphics Forum*, 13:233–246, 1994.

[18] Peter Vangorp, Jurgen Laurijssen, and Philip Dutre. The influence of shape on the perception of material reflectance. *ACM Trans. Graph.*, 26, July 2007.

[19] Gregory J. Ward. Measuring and modeling anisotropic reflection. In *Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, pages 265–272, New York, NY, USA, 1992. SIGGRAPH '92, ACM.

[20] Chanki Yu, Yongduek Seo, and Sang Wook Lee. Global optimization for estimating a multiple-lobe analytical brdf. *Comput. Vis. Image Underst.*, 115(12):1679–1688, December 2011.