Perceiving Gloss in Surfaces and Images

Adria Fores[1], James Ferwerda[1], Ingeborg Tastl[2], and John Recker[2]
[1] ***Munsell Color Science Laboratory, Center for Imaging Science, Rochester Institute of Technology***
[2] ***Hewlett-Packard Laboratories***

Abstract

Color Appearance Models are successfully used to model the color perception differences seen when the same stimuli are presented on different media, e.g. hard copy or a self-luminous display. It is currently unknown if the similar effects are present in gloss perception and if there is need for Gloss Appearance Models.

Gloss communication, and the higher level material appearance communication is becoming more important everyday with the increase in customized manufacturing and the need for the costumer to preview a final product while short-runs, time and cost constraints prohibit the use of hard-copy proofs.

Three experiments are proposed in order to analyze this phenomenon. The Gloss matching performance of observers on real objects is first going to be studied. Then, the same experiment will be repeated with synthetic images. Finally, a cross-media matching experiment will be performed, where the observers will have to match a real material with synthetic representations.

The same trend was observed in the experiment using only real objects and in the cross-media situation, where a high matching accuracy was obtained for low gloss samples, and the gloss of mid and high gloss samples was underestimated. The same accuracy for low gloss samples was obtained for the experiment with only synthetic images, but mid and high gloss samples were overestimated. The sensitivity of the observers was higher when only real samples were used, it decreased when the display was used due the lack of visual disparity and multiple viewing conditions, and it was lowest on the last experiment, influenced by the multiple media and the above limitations.

Introduction

Gloss communication, and the higher level material appearance communication is becoming more important every day with the increase in customized manufacturing and the need for the costumer to preview a final product while short-runs and cost constraints do not allow the use of hard-copy proofs.

Color Appearance Models were developed to account for the viewing conditions and its effects on the perception of color. The same color stimuli seen on a hard copy and on a self-luminous display produces different color perceptions. Color Appearance Models are successfully used and have been widely evaluated [10] to model those changes in appearance and enable to create the same color perception on different media.

It is currently unknown if the same effect is present in gloss perception, or if there is any need for Gloss Appearance Models. This project is designed to study if there exists a fundamental difference in cross-media gloss perception. Gloss communication could be improved with a transformation that accounted for the difference between the representation of a material seen on a display and the real material.

Gloss also varies in other dimensions than color. Vangorp et al. [13] studied the gloss perception dependence on an object's shape, and found that the material appearance perceived varied depending on the shape of the object. By using the uniform gloss space defined in Pellacini et al. [8] the authors modeled the shape dependence and were able to correct for it, being able to match the gloss appearance of two objects with different shapes.

In this project, in order to understand the gloss perception difference between real objects and synthetic objects seen on a display, three different matching experiments will be conducted.

In the first experiment, the observers will have to match real objects in a custom-built light booth. This will enable an understanding of the accuracy of the observers and their variability when performing the task with real objects.

In the second experiment, the observers will repeat the same task but they will perform it on a display with synthetic images representing the real objects. As with the previous experiment, this will enable an understanding of the accuracy and variability of the observers performing this task on another media. More interestingly, it will allow us to compare how the accuracy and variability varies from the real objects to the simulations.

In the third experiment, a cross-media matching experiment will be performed in which the observers will have to select the simulation of an object that matches a real object. This will allow us to understand the influence of the media used in the matching task.

Related work

Gloss is a perceptual attribute related to the physical phenomenon of the Bidirectional Reflectance Distribution Function (BRDF).

The BRDF is a 4-dimensional function that describes how light is scattered by a surface and it is defined by the following equation:

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

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

In their classic publication, Hunter and Harold [5] described six features that relate to the perception of Gloss:

Specular gloss This property models the specular reflection at different angles, commonly 20°, 30°, 45°, 60°, and 75°. The integration of the reflected light at a given aperture for a material in respect to a black glass defines the specular gloss.

Figure 1. *Setup used for the experiment. From left to right, 30-inch HP ZR30w display, custom-built light booth, and lazy susan used to provide easy access to the samples to the users. The color difference seen between the different media is due to the camera response.*

Lower angles are used to compare high-specular materials and higher angles are used to compare low-specular materials.

Sheen This property models the specular reflection at grazing angles and it is defined at 85°.

Contrast gloss or luster Defines the difference between the highlight areas and its surrounding. This effect can be clearly seen in velvet cloth, which has distinct highlights and dark areas.

Absence-of-bloom gloss Also known as absence of haze, which is defined as the spread of the specular component of the reflected light from a glossy surface.

Distinctness-of-image gloss This property defines how well a material allows to distinguish the reflected background on the surface of the material. For example, a mirror will have a higher distinctness-of-image than a brushed metal as the mirror is going to sharply reflect the background, while the brushed metal will introduce some amount of blur to the reflected image of the background.

Surface-uniformity gloss This property defines how smooth a surface is, being able to perceive a non-uniform texture when the surface is rough.

The importance of gloss in the finishing of commercial products drove the creation of international standards concerning the measurements of some of those perceptual gloss attributes: Specular Gloss is defined in ISO 2813, ISO7668, ASTM D523, ASTM D2457, DIN 67530, and JIS 8741, Distictness-of-image gloss is defined in ASTM D5767, and Haze is defined in ASTM E430, and ISO 13803.

Setup

The setup created to perform the experiments presented in this paper can be seen in Figure 1. A custom-built light booth is used for the first experiment, and the combination of the light booth and a display are used for the rest of the experiments. The perspective between what is seen on the display (left) and what is seen in the light booth (right) does not match because the synthetic images are generated from a specific viewpoint and have the correct perspective only when the observer is located at a specific position and looking at the samples. Figure 2 shows the observer stimuli as seen in the three different experiments with the most and least glossy sample being presented.

In this section, the design decisions and the detailed information for each setup is explained.

Real Scene

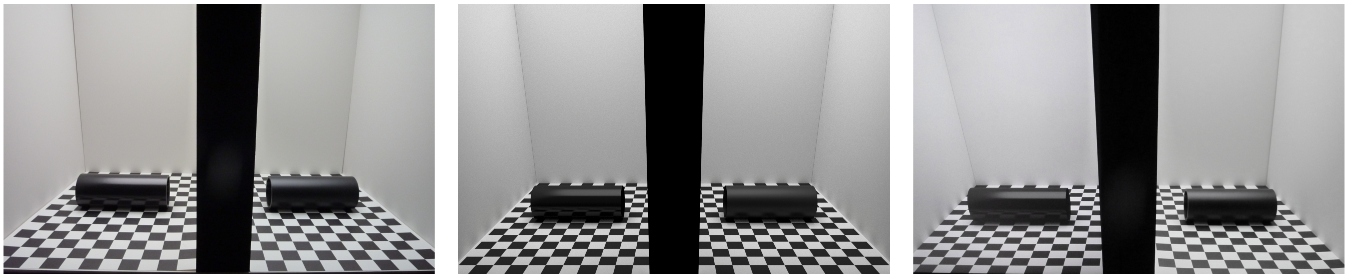
To perform the perceptual study on real objects, a scene that is easy reproducible when generating synthetic images and at the same time enhances the material discrimination was designed.

The custom-built light booth and the material samples used for this set of experiments can be seen in Figure 1 right. The light booth consist of a wood-structure with an opening of the same size as the 30-inch HP ZR30w display, the one used to display synthetic images, while the depth of the light booth is the same as the vertical edge of the same 30-inch display. A photo studio light source with CFL light bulbs was used to lit the scene. The inner and outer diffuser of the light source were used, and another diffuser was placed on top of the light booth. This light source provided constant chromaticity over angle and the peak luminance of the real scene was slightly lower than the peak luminance of the display ($330 cd/m^2$).

The light booth is split vertically in order to accommodate the two scenes used for the matching experiment. By having the regions physically separated we avoid comparisons when the two objects are too close together, which would improve the accuracy of observer judgements to the point that it would be difficult to relate results from a single media experiment to the cross-media experiment. The separation of the real scenes is equal to the monitor frame plus the light booth frame, in order to allow to have the same distance between samples in all scenes.

To enhance the material discrimination, a checkerboard pattern was placed on the bottom of the light booth, allowing to see more or less distinctive reflections of the checkerboard pattern depending on the glossiness of the objects used in the experiment.

The object shape selected for this experiments was a cylin-



(a) Real vs. Real (picture) (b) Display vs. Display (image) (c) Display vs. Real (picture)

Figure 2. *Viewing conditions of the three different experiments, displaying the most glossy sample (left) and least glossy sample (right).*

der. Several advantages are found on a cylinder over other shapes: they are easy to manufacture, easy to represent on synthetic images due to their analytical definition, easy to wrap paper around them (see later), and were found to provide a high material discrimination in Vangorp et al. [14]. Cylinders of diameters 4 and 6 centimeters were created and evaluated. The ones with a diameter of 6 centimeters were finally used as the lower curvature allowed a better discrimination between samples. This was because the spread of the specular lobe and the reflections of the environment were occurring over a larger area and were easier to perceive.

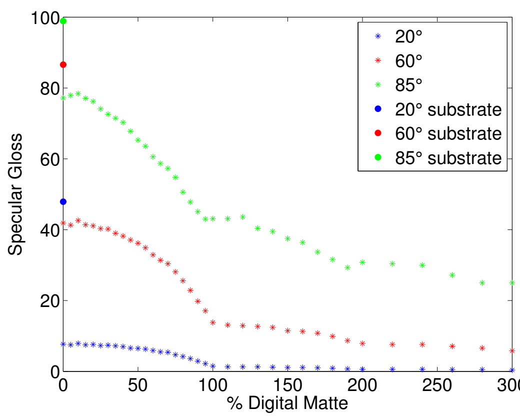


Figure 3. 20°*, *60°*, and *85° *Specular Gloss of the samples used in the experiment and the glossy substrate used.*

The key aspect of the physical setup is the material selection. A set of materials that only varied in gloss was obtained, while keeping other appearance attributes like texture and color constant. In order to achieve this goal, the Digital Matte feature of the HP Indigo 5000 Digital Press was used. The Digital Matte is a varnish that decreases the gloss of the surface on which it is applied. For that reason, the glossy HP Photo Paper was selected as starting point. Then, a first layer of 100% black ink was applied on top of the paper in order to increase the contrast and enhance gloss perception. Then, a varying amount of Digital Matte (0-300%) was applied on top of the black ink in order to obtain 36 samples of different gloss levels. The amounts of Digital Matte used were visually selected in order to approximately have the samples equally spaced in terms of perceived gloss. Figures 3 and 4 shows the Specular Gloss and the Haze of the created samples measured with a Elcometer 6015 DOI Haze Meter, respectively. In spite of also measuring the distinctness-of-image gloss, it is not

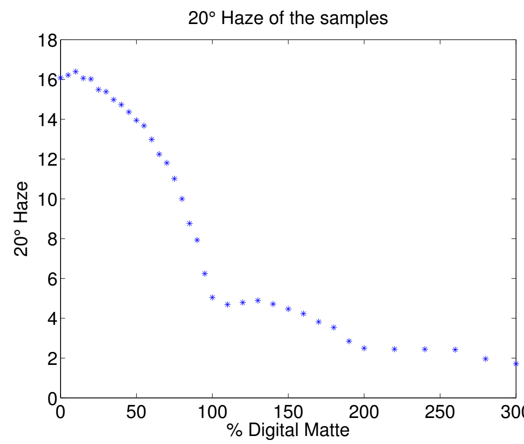


Figure 4. 20° *Haze of the samples used in the experiment.*

reported in this document as it it didn't allow to differentiate the samples used, mainly because of its broad specular lobe.

Synthetic images

An accurate representation of the materials, geometrical objects, and lighting of the custom-built light booth needed to be obtained in order to represent the real scene in synthetic images. The scene described before was carefully designed taken into account that it had to be used afterwards for rendering and for that reason simple geometric objects like a rectangular shape and cylinders were used instead of other selections.

The light source used in the physical scene was carefully measured in order to correctly simulate it. The light source was measured with a PR-650 spectroradiometer at 10° intervals from the normal direction up to 80°. A constant chromaticity over angle was found, and the luminance fall-off measured was approximated with a cubic polynomial. This approximation was then implemented into the rendering engine.

Finally, the most important aspect to represent in the synthetic images is the material appearance of the different samples. In order to get the highest angular sampling possible within the available resources, the samples were measured with the Eldim EZContrast 160R at the Centre de Recherche sur la Conservation des Collections in Paris, France. This device allows to measure all the outgoing directions (up to 80° in θ) at once given a single incident direction. Figure 5 shows the diagram of the instrument, in which the light reflected on the sample over all the directions is

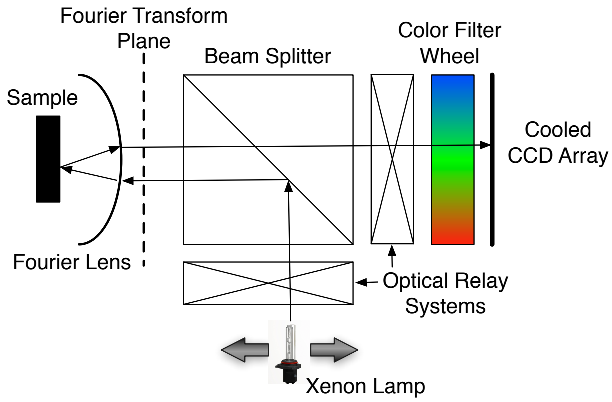


Figure 5. *Design schematic of the Eldim EZContrast 160R.*

projected onto a plane with a special fourier lens and that plane is imaged by the device sensor. A beamsplitter is used in the middle in order to set the incident light direction, which is projected back to the sample at the selected incident direction using the fourier lens. Five transmission filters are used in front of the CCD sensor to obtain a good approximation of CIE XYZ, and for each measurement a different exposure is performed for each filter.

While all the outgoing directions are measured at the same time, a separate measurement is required for each incident direction. Due to the time required for each measurement and the number of samples to measure, only the following incident directions were measured for each sample: $\theta = 5°, 15°, 30°, 45°, 60°$ and $70°$ with $\phi = 0°$. It's important to note that the material will be considered isotropic, even thought paper substrates are made of fibers that produce a light anisotropy. Each of the measurements was later on calibrated using the following equation:

$$f_{sample}(\omega_i, \omega_o) = \frac{i_{sample}(\omega_i, \omega_o) - i_{black}(\omega_i, \omega_o)}{(i_{ref}(\omega_i, \omega_o) - i_{black}(\omega_i, \omega_o))\pi} \quad (2)$$

where f_{sample} is the BRDF defined at the direction (ω_i, ω_o). The i_{sampe} is the measurement of the sample, the i_{black} is the measurement with a black trap in front of the measurement port, and the i_{ref} is the measurement of a PTFE created from teflon powder. The i_{black} measurements were required due to the stray light produced by the fourier lens in the measurement device, however it was only measured for incident directions up to $30°$, as the stray light is considered to be negligible for higher incident directions.

Two options are present when measured materials are considered for rendering. The first one and more commonly used is to approximate the measurements with an analytical BRDF model and the other technique is to render the measured data directly by means of interpolation. The option to approximate the measurements with an analytical model was selected in this paper. As a future work, it would be interesting to explore the direct rendering of the measured data, for example using the technique proposed in Stark et al. [12].

In order to approximate each material a Lambertian lobe was used to represent the diffuse component, and it was set to the 45:0 measurement, and the parameters of a specular BRDF model were non-linearly optimized. The performance of the Ward [16], the Ashikhmin-Shirley [1], and the Cook-Torrance [2] BRDF models to approximate the samples used in our study was evaluated. At the same time, the behavior of different error metrics was also evaluated: the RMS cosine weighted error metric and the cube root cosine weighted metric.

The materials used in this work were well approximated with the Ashikhmin-Shirley BRDF model and the RMS cosine weighted error metric. This selection differs from the one presented in Fores et al. [4], in which several specular lobes were needed to faithfully represent the materials of the MERL Database [6]. This difference is probably due to the fact that the Ashikhmin-Shirley BRDF model can better represent the material samples used in this study than the ones of the MERL Database.

The Ashikhmin-Shirley analytical BRDF model is defined by the following equation:

$$K = \frac{m+1}{8\pi} \frac{(n \cdot h)^m}{(\omega_o \cdot h) max((n \cdot \omega_i), (n \cdot \omega_o))} Fresnel(F_0, \omega_o, h) \quad (3)$$

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 [11], which depends on the parameter F_0:

$$Fresnel(F_0, \omega_o, h) = F_0 + (1 - F_0) \cdot (1 - (\omega_o \cdot h))^5 \quad (4)$$

The RMS cosine weighted metric is defined by the following equation:

$$E = \sqrt{\frac{\sum (M(\omega_i, \omega_o) cos\theta_i - A(\omega_i, \omega_o, p) cos\theta_i)^2}{n}} \quad (5)$$

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 paper samples used in this work had a small relief, as happens with most paper substrates. This feature of the samples was constant across samples due the fabrication technique and we decided not to represent it in the simulation, mainly because it was not perceived from the viewing distance of the observers.

The Physically Based Ray Tracer (PBRT) [9] was used to generate the synthetic images presented to the observers from the specific observers' viewpoints. The rendering is performed after transforming the XYZ tristimulus values of the materials and the light source to the sharpen cone responses obtained with the CAT02 matrix, following the technique described in Ward [15].

Finally, the display was characterized using a PR-650 spectroradiometer and the Day model [3], presenting a good additivity and obtaining a mean $CIEDE2000$ of 0.36 when displaying the colors of the 24 patches of the Macbeth Color Checker.

In spite of the careful measurements and approximations of the samples and light source, the luminance of the highlights on the cylinders slightly differ when different media were compared (see Figure 6). The measurement's accuracy, and specially the approximations of the specular component of the materials are probably the main facts of the difference seen. In order to address this limitation, the specular component of the materials was scaled to account for the difference between the measurements of the cylinders in the light booth and the cylinders displayed on the monitor. This scaling process helped to better match the luminance of the samples in the real scene on the simulated images.

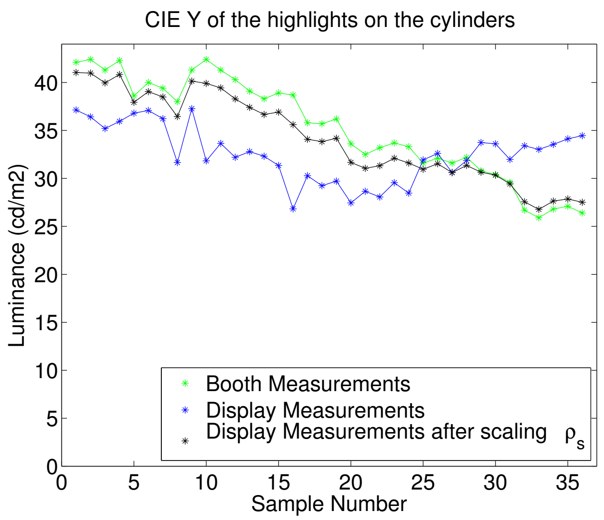


Figure 6. *Luminance of the highlights on the cylinders measured on the light booth and the display using a PR-650.*

Experiments

The same matching task was performed in all the experiments. The observers were asked to find the match to a sample reference given another full set of 36 samples to choose from.

In each experiment, the observers had to match 8 samples. The samples were carefully selected to be just noticeably different from the other samples selected to be matched, while keeping some samples on the low and high end.

The three experiments were conducted with the setup seen in Figure 1. The experiment was conducted in the dark, in the single media experiments the apparatus not in use was turned off. Fifteen observers with normal color vision and normal or corrected to normal visual acuity participated in the experiments.

Experiment 1: Real vs. Real

In the first experiment, the observers had to match real objects in the custom-built light booth. The reference samples to be matched by the observers were placed, one at a time, in the left side of the light booth. For each of them, the observer had to select the sample that matched the reference sample.

Observers wore latex gloves to avoid damaging the appearance of the samples, as the grease in the skin would rapidly dull the appearance of the paper samples. To easily browse the 36 material samples easily we built a lazy susan, a circular surface with bearings underneath that allows to rotate the surface freely (see Figure 1 right). A set of dowels were placed along the circle, allowing to set and secure the 36 samples used in the experiment. This setup allowed the observer to efficiently change the sample to be inspected inside the light booth. The samples were sorted from most glossy to least glossy (in terms of % Digital Matte) along the circle.

Before starting the experiment, every observer was trained in how to use the setup. First, the goal of the experiment, to better understand gloss perception in different media was explained. Second, the samples were presented to the observer, telling them that cylinders were wrapped with paper of different gloss levels and that they were sorted from most to least glossy. Then, a demonstration was given by placing the most and least glossy samples inside the light booth while explaining how to place and align the samples in the small rubber pieces that kept the cylinders in place inside the light booth. Finally, specific instructions to maintain the accuracy of the experiment were given to the observer: only one sample of the matching set was allowed to be taken from the lazy susan at a time, the observer was asked to place the cylinder inside the light booth and rotate it to avoid seeing the seam, to keep the hands off the light booth when making decisions, and to always make the final decision with the sample inside the light booth. Observers were allowed to make a first guess and navigate the gloss range by looking at the reflection seen on the samples in the lazy susan. Once it was clear that the observer understood the setup and the task to conduct, a trial sample that was not recorded was given to the observer in order to let him accommodate to the setup before the start of the experiment.

Results Experiment 1

The data obtained from the Real vs. Real experiment can be seen in Figure 7, in which the mean and standard deviation matching performance of the observers in respect to 60° Specular Gloss and 20° Haze for each of the 8 samples is shown. The small black circles on the x axis show the measured perceptual properties for all the samples used in the study. The Specular Gloss and Haze measurements of the samples used for this study are highly correlated, thus the matching performance and the fitted linear equation parameters are almost identical. Still, that information will be reported in this document for reference. Because of that similarity, the generic term gloss will be used to refer to both perceptual features, Specular Gloss and Haze, when explaining the results.

The black diagonal line shows the 1:1 correspondence if the observers were selecting the sample with equal gloss when performing the matching experiment. It can be seen that for low gloss materials the observers are accurate with their selections and for higher gloss materials the observers tend to select a sample with a lower measured gloss. The fact that the samples are not equally spaced in terms of gloss produces the difference in terms of standard deviation that is seen across the range. At the same time, the cluster of samples in the high end of the gloss range and the large standard deviation obtained on that area probably means that the discrimination there was harder than in the low gloss range, or the possibility that the samples in that area are closer perceptually.

A linear equation was fit to the mean observer responses and it can be seen that it correctly models the data, modeling the accurate matching performance for low gloss samples and the gloss underestimation for high gloss samples.

Experiment 2: Display vs. Display

In the second experiment, the observers repeated the same task performed in the first experiment, but in this case the experiment was performed on the monitor by displaying synthetic images representing the real objects. The synthetic images were rendered from the same point of view where the observer looked at the real scene in the first experiment, and the camera was tilted down to match the height where the cylinders are located.

In this case, the observers were able to navigate across the gloss range to perform the matching task by using the left and right arrow keys from the keyboard. The spacebar was pressed by the observer when the match was found, and also directed the

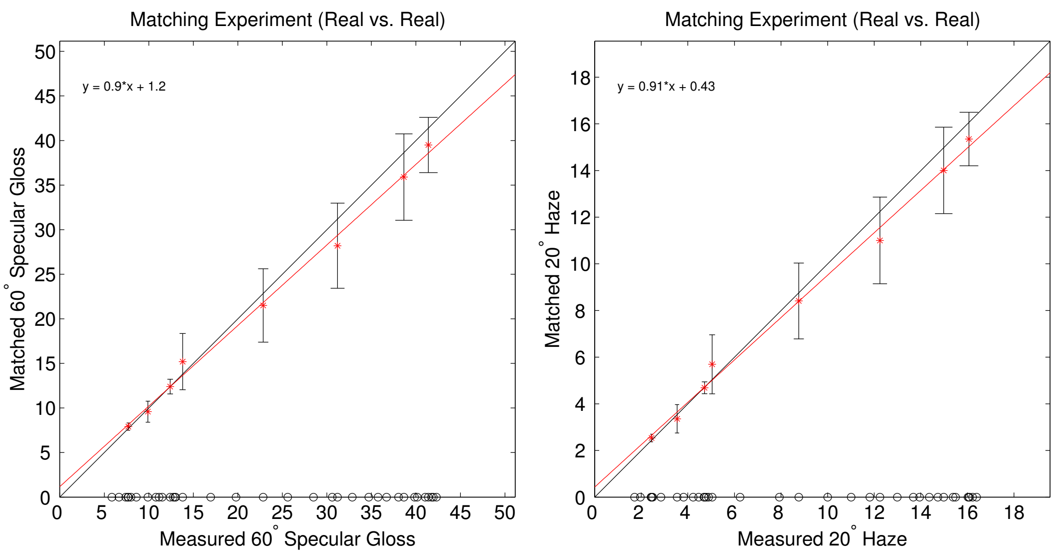


Figure 7. *Real vs. Real experiment results. Gloss matching performance relative to* 60° *Specular Gloss and* 20° *Haze. Sample distribution on the X axis.*

observer to a black screen indicating the number of samples left to match. As in the first experiment, the sample materials that the observer was able to inspect were sorted by gloss level, from the most glossy to the least glossy, and the system alerted the observer with a visual indication on the screen if they reached the end of the range in a given direction.

Results Experiment 2

Figure 8 shows the results obtained from the Display vs. Display experiment. It can be seen that for low gloss materials the observers are accurate with their selections, samples in the mid range of gloss evaluated are matched by glossier materials, and the glossiest sample was matched to a less glossy material. For the low and high gloss materials of the studied range the observers had the same perception seen in the Real vs. Real experiment, but the opposite effect was seen on the samples in the middle range.

Experiment 3: Display vs. Real

In the third experiment, the observers repeated the same task performed in previous experiments, but in this case the reference sample was placed in the light booth and the sample the observer had control over was seen on the monitor. In this case, just half of the light booth was visible to the observer while the other part of the scene was physically blocked, at the same time, only a single scene was being shown in the rendered image on the display.

For this experiment, the observer was located between the display and the light booth, and the synthetic images were rendered from that same point of view, and the camera was tilted down to match the height where the cylinders are located.

In this case, a part from different media being evaluated at the same time, there was another major difference. The observers were asked to only use their dominant eye, while closing the other. By doing the experiment with monocular vision, the perspective of the real scene and the synthetic image matched. This might have influenced the experiment, in the same way that binocular cues were eliminated from the second experiment, where no stereo was used. Conducting this cross-media experiment with binocular information in all circumstances would be a challenging experimental design problem, as the technique used to split the image that goes to each eye using glasses would probably also affect the perception of the real scene. For example, the use of polarized glasses would influence and modify the specular reflections seen on the real objects, while shutter glasses would dramatically reduce the luminance of the real scene to the point that it might become hard to perform the experiment and it could also produce a flickering effect on the real scene. Still, the study of the influence of stereo vision in cross-media gloss perception would be an interesting topic for further research. The first and second experiments were not run with monocular vision as we wanted to evaluate the real life performance of observers.

Results Experiment 3

Figure 9 shows the results obtained from the Display vs. Real experiment. An accurate observers' matching ability is seen for low gloss materials, while a slight gloss underestimation is seen for mid and high gloss materials.

A linear equation was fit to the mean observer responses and it can be seen that it correctly models the data, modeling the accurate matching performance for low gloss samples, with the slight gloss overestimation observed, and the gloss underestimation for high gloss samples.

Perceptual Scale

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 [7].

In pre-testing samples were chosen by subdivision to form an approximately perceptually uniform scale with sub-JND intervals. For this reason, sample numbers (1-36) rather than Specular Gloss were used in the scaling analysis.

A gaussian distribution was approximated to the observer matching responses for each experiment, those were later used to

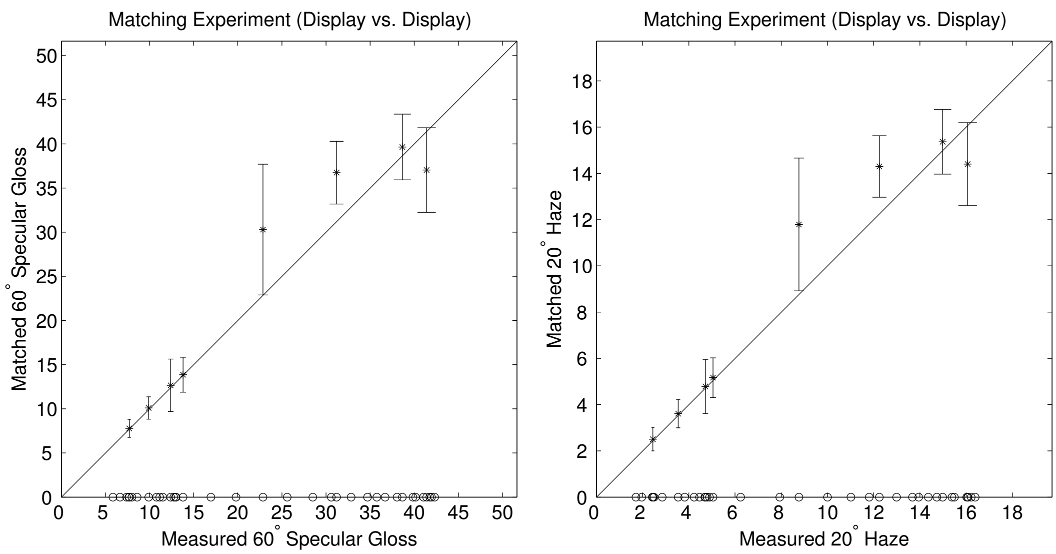


Figure 8. *Display vs. Display experiment results. Gloss matching performance relative to* 60° *Specular Gloss and* 20° *Haze. Sample distribution on the X axis.*

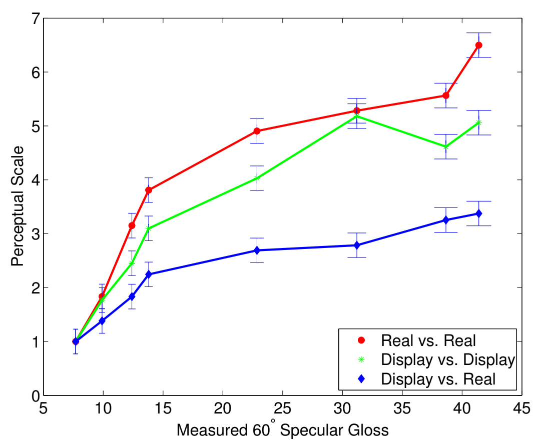


Figure 10. *Perceptual Scales obtained from the three different experiments.*

compute the probability matrix. For each experiment, every sample was compared against each other and the integration to the cross-over point of the distributions indicated the probability of a given sample to be selected as more glossy. While the probability of the other being selected was 1 minus that probability.

Then, the common Thurstone Case V evaluation was computed using the probability matrix obtained. The χ^2 test was performed and showed that the variance of the samples in each experiment was equal, thus being able to use the Case V.

Figure 10 shows the perceptual scales obtained for the different experiments. Significant differences in sensitivity between the different conditions are observed. The highest sensitivity was obtained in the Real vs. Real experiment, a slightly lower sensitivity was observed for the Display vs. Display experiment, and the lowest sensitivity was observed for the cross-media experiment.

Binocular vision, which provides binocular disparity, and the ability of having multiple viewing directions of the samples are probably the main reasons why the highest sensitivity was obtained in the Real vs. Real experiment. Probably, the lack of multiple viewing directions and lack of binocular disparity caused the reduction in sensitivity seen in the Display vs. Display experiment. Finally, the task to perform the experiment using different media decreased even more the sensitivity of the observers in the Display vs. Real experiment, which also had the viewing direction restricted and monocular vision was used.

Discussion

In this project, three different gloss discrimination experiments were performed. In the first one, the gloss matching performance with real samples was evaluated. The second experiment evaluated the matching ability when using synthetic images as representations of real objects. Finally, the last experiment evaluated the discrimination ability on the cross-media situation, where real objects were matched with synthetic representations.

The same trend seen in the Real vs. Real experiment was observed in the Display vs. Real experiment, where a high matching accuracy was obtained for low gloss samples and the gloss of mid and high gloss samples was underestimated. A more pronounced gloss underestimation for those samples was obtained in the cross-media study, as seen with the different slopes obtained.

Similar accuracy was obtained for low gloss samples in the Display vs. Display experiment, but the observer responses for the mid and high gloss samples were the opposite than the ones seen in the other experiments, as gloss was overestimated.

The observers' sensitivity decrease as more constrains were set in the experimental design. The highest sensitivity was observed in the first experiment, where real samples were used. The sensitivity decreased on the matching experiment done with synthetic images due the lack of binocular disparity and fixed viewing direction. Finally, the cross-media study added more constrains like monocular vision and the challenge to deal with multiple media, obtaining the lowest observer sensitivity.

The results obtained in this experiment show that a small increase in gloss might be needed when synthetic images are used as representations of physical objects. This gloss increase refers to both, Specular Gloss and Haze. The increase of Specular Gloss

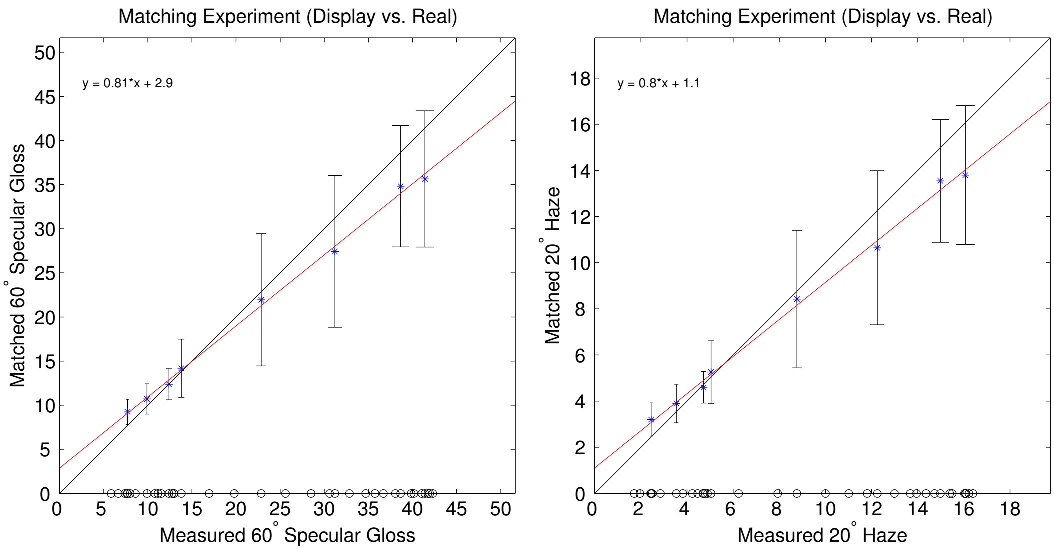


Figure 9. *Display vs. Real experiment results. Gloss matching performance relative to* 60° *Specular Gloss and* 20° *Haze. Sample distribution on the X axis.*

could probably be implemented by scaling the specular lobe, but there is no direct mapping for performing the same task with Haze. If using the remapped Ward BRDF model presented in [8], the distictness-of-image parameter might be influencing the shape of the specular lobe in a similar fashion as the Haze is doing in the samples of this study. An interesting venue of future work would be the creation of a perceptually based gloss space that could input both, measured data and analytical models, and could be used to describe and model the results presented in this paper.

The gloss range evaluated in this study was limited by the sample creation process used, which limited the highest gloss sample to be in the mid gloss region. As a future work, it would be interesting to evaluate the observer discrimination of high gloss materials, as the observer discrimination is probably going to be different than the one seen on the studied sample set.

Acknowledgments

We would like to thank Françoise Viénot from the Centre de Recherche sur la Conservation des Collections in Paris, France, for allowing us to use the Eldim EZContrast and assisting us during the measurement process. James Ferwerda was partially supported by NSF grant HCC-10644112.

References

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

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

[3] Ellen A. Day, Lawrence A. Taplin, and Roy S. Berns. Colorimetric characterization of a computer-controlled liquid crystal display. *Color Research and Application*, 29:365–373, 2004.

[4] Adria Fores, James Ferwerda, and Jinwei Gu. Toward a perceptually based metric for brdf modeling. In *Twentieth Color and Imaging Conference. Los Angeles, California, USA*, pages 142–148, November 2012.

[5] Hunter and Harold. *Measurement of Appearance*. Chapter 6 (Gloss), 2nd edition, 1987.

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

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

[8] 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.

[9] 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.

[10] Report. CIE Technical Committee TC1-27 Summary Report. Technical report, CIE, 2001.

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

[12] Michael M. Stark, James Arvo, and Brian Smits. Barycentric parameterizations for isotropic brdfs. *IEEE Transactions on Visualization and Computer Graphics*, 11(2):126–138, March 2005.

[13] Peter Vangorp and Philip Dutré. Shape-dependent gloss correction. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, APGV '08, pages 123–130, New York, NY, USA, 2008. ACM.

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

[15] Greg Ward and Elena Eydelberg-Vileshin. Picture perfect rgb rendering using spectral prefiltering and sharp color primaries. In *Proceedings of the 13th Eurographics workshop on Rendering*, EGRW '02, pages 117–124, Aire-la-Ville, Switzerland, Switzerland, 2002. Eurographics Association.

[16] 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.