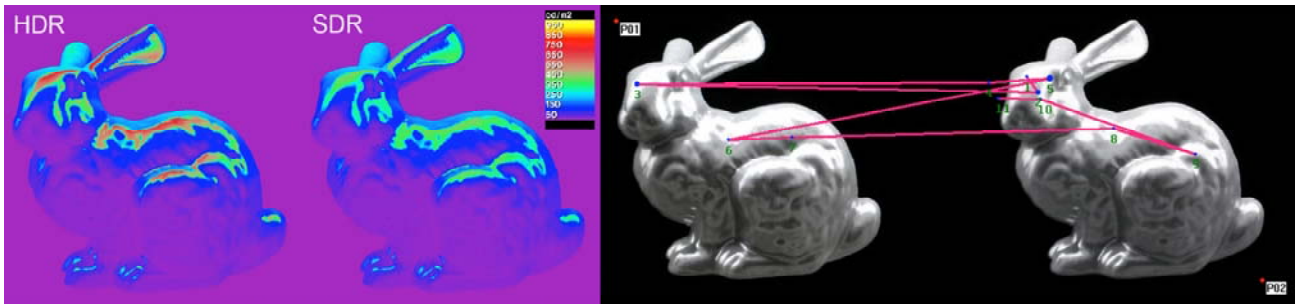
Gloss discrimination and eye movements

Jonathan B. Phillips, James A. Ferwerda, and Ann Nunziata
Munsell Color Science Laboratory, Chester F. Carlson Center for Imaging Science
Rochester Institute of Technology, Rochester, NY USA

Abstract

Human observers are able to make fine discriminations of surface gloss. What cues are they using to perform this task? In previous studies, we identified two reflection-related cues—the contrast of the reflected image (c, contrast gloss) and the sharpness of reflected image (d, distinctness-of-image gloss)--but these were for objects rendered in standard dynamic range (SDR) images with compressed highlights. In ongoing work, we are studying the effects of image dynamic range on perceived gloss, comparing high dynamic range (HDR) images with accurate reflections and SDR images with compressed reflections. In this paper, we first present the basic findings of this gloss discrimination study then present an analysis of eye movement recordings that show where observers were looking during the gloss discrimination task. The results indicate that: 1) image dynamic range has significant influence on perceived gloss, with surfaces presented in HDR images being seen as glossier and more discriminable than their SDR counterparts; 2) observers look at both light source highlights and environmental interreflections when judging gloss; and 3) both of these results are modulated by surface geometry and scene illumination.



(a) (b)

Figure 1*: (a) Luminance maps of a sample high dynamic range (HDR) and standard dynamic range (SDR) image pair of a bunny object. Note the high intensities in the HDR highlights and the highlight compression in the SDR image. (b) Eye movement scanpaths over the sample pair showing fixations made when discriminating gloss.*

Introduction

One of the defining characteristics of glossy surfaces is that they reflect images of their surroundings. High gloss surfaces produce sharp detailed reflections that clearly show all the features of the surround, while low gloss surfaces produce blurry reflections that only show bright "highlight" features. Due to the presence of light sources and shadows, the images reflected by glossy surfaces can be high dynamic range (HDR). However in conventional images of glossy objects, these high dynamic range reflection images are clipped or compressed through tone mapping so the images fit within the standard dynamic range (SDR) output of the display medium (see Figure 1). While the utility of conventional display systems demonstrates that the general characteristics of glossy surfaces are conveyed by tone-mapped SDR images, an open question is whether the tone mapping process distorts the apparent gloss of the imaged objects.

In this paper, we first supplement previous results from a psychophysical study that investigates the effects of image dynamic range on apparent surface gloss using a custom-built HDR display.[1] In the study, we present HDR and SDR images of glossy objects in pairs and ask subjects to choose the glossier object. We analyze the results of the study using Thurstonian scaling, and derive common scales of perceived gloss for the objects depicted in both the HDR and SDR images. To investigate the effects of geometric complexity, we use both simple and complex objects. To investigate the effects of environmental illumination, we use both a simple area light source and a captured, real-world illumination map. Finally, we present fixation analysis of a companion eye movement study that analyzes where subjects look during the gloss discrimination task. Eye tracking is used to determine the salient regions and to compare fixations with the luminance levels and positions on the rendered surfaces.

The main findings of the studies are: 1) image dynamic range has significant influence on perceived surface gloss, with surfaces presented in HDR images being seen as glossier and more discriminable than their SDR counterparts; 2) subjects look at both surface highlights and body reflections when judging surface gloss; and 3) both of these results are modulated by surface geometry and scene illumination.

Related work

Gloss perception

The earliest modern studies of gloss perception have been attributed to Ingersoll[2] who examined the appearance of glossy papers. In 1937, Hunter[3] observed at least six different visual phenomena related to apparent gloss. He defined these as:

specular gloss – perceived brightness associated with the specular reflection from a surface
contrast gloss – perceived relative brightness of specularly and diffusely reflecting areas
distinctness-of-image (DOI) gloss – perceived sharpness of images reflected in a surface
haze – perceived cloudiness in reflections near the specular direction
sheen – perceived shininess at grazing angles in otherwise matte surfaces
absence-of-texture gloss – perceived surface smoothness and uniformity

In 1937, Judd[4] formalized Hunter's observations by writing expressions that related them to the physical features of surface bi-directional reflectance distribution functions (BRDFs). Hunter and Judd's research established a conceptual framework that has dominated work in gloss perception to the present day.

In 1987, Billmeyer and O'Donnell[5] published an important paper that investigated the multidimensional nature of gloss perception. They collected ratings of the differences in apparent gloss between pairs of acrylic-painted panels with varying gloss levels viewed under a fluorescent desk lamp outfitted with a chicken-wire screen, then used multidimensional scaling techniques to discover the dimensionality of perceived gloss. For their experimental conditions, they found that gloss could be described by a single dimension. However, this work was significant because it was the first to study the multidimensional nature of gloss perception without preconceptions about how many or what the dimensions might be. In a 1986 report to the CIE, Christie[6] summarized the research findings on gloss perception up to that date. Since that time, McCamy[7,8] has published a pair of review papers on the gloss attributes of metallic surfaces and Sève[9] and Lozano[10] have outlined frameworks for describing gloss that seek to improve on Hunter's classifications. In the Imaging Science literature, there has been considerable interest in the effects of gloss on printed image quality with efforts to characterize artifacts like differential gloss, bronzing, and gloss mottle.[11-16]

Computer graphics studies of gloss perception

One of the challenges in conducting gloss perception research is producing and controlling the stimuli used in the experiments. Generating consistent physical samples is very difficult. Therefore, the development of physically-based computer graphics techniques that can produce and present radiometrically accurate images of complex scenes has been a boon to the psychophysical study of gloss perception. One of the earliest computer graphics studies was done by Nishida and Shinya[17] who rendered bumpy glossy surfaces using direct point lighting. They found that observers made consistent errors in matching gloss properties across different surface geometries and suggested that the results of their experiments could be explained with a simple image histogram matching strategy. Pellacini et al.[18] conducted a set of experiments inspired by Billmeyer and O'Donnell's multidimensional scaling studies, but with images of a glossy ball inside a checkerboard box with a ceiling-mounted area light source. For this stimulus set, they found that observers used two dimensions to judge gloss, "c" a measure related to the contrast of the image reflected by the surface, and "d" a measure related to the sharpness of the reflected image. Ferwerda et al.[19] extended this work to characterize multidimensional gloss differences. More recent work has examined the role of natural illumination patterns[20] and complex object geometry[21,22] on surface gloss perception.

Although the power of computer graphics has greatly facilitated the study of gloss perception, one of the caveats of all of these studies is that they use *images* of glossy surfaces as stimuli rather than the surfaces themselves. Conventional imaging systems do not have the dynamic range that actual surfaces can achieve, thus requiring tone mapping operators. Because the potentially high dynamic range reflections from glossy surfaces are compressed for display, we are studying the question of whether the gloss properties of the displayed surfaces are distorted and, therefore, not fully and accurately represented.

Eye movements

When observers are looking at complex imaging content such as pictures, eye tracking experiments reveal that scanpaths and fixations are dependent on the task given to the observer.[23,24] In a classic experiment by Yarbus[24] that used an image of a living room scene, observers made different patterns of eye movements depending on whether they were asked to remember the positions of objects in the scenes, judge the ages of the people in the scene, explain what was happening in the scene, or freely view the scene. This task dependence of eye movements means that eye movement studies of gloss discrimination tasks have the potential to reveal the image features that observers use to judge surface gloss.

Eye movement studies have also been shown to be an effective means for determining impact on image assessment related to dynamic range. Bloj et al.[25] have compared fixation records made to original and tone-mapped images and used the similarity of fixation patterns as a measure of the fidelity and non-destructiveness of different tone mapping operators.

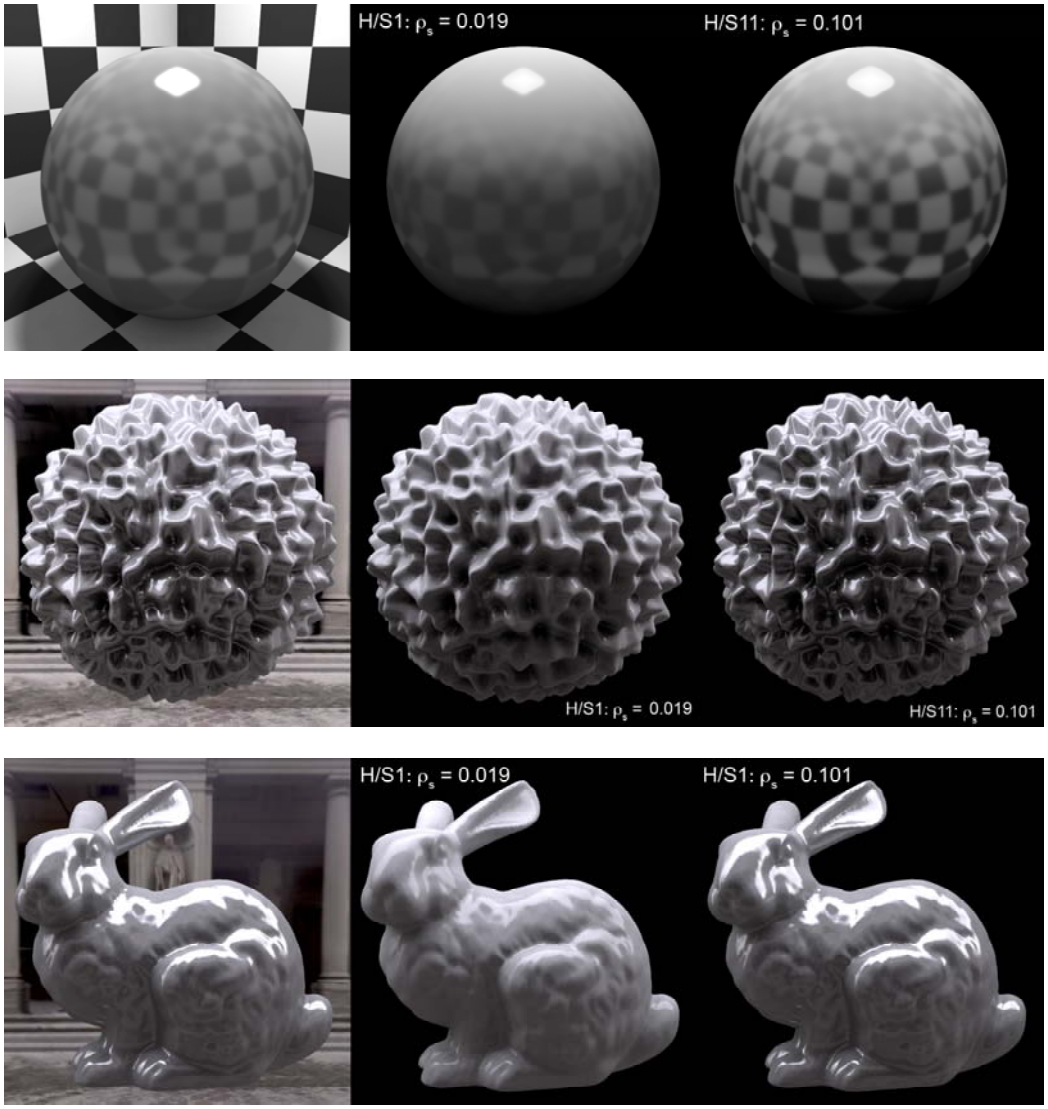
We have supplemented our previous experimentation[1] with additional scene content and more rigor. In our current research, we start with HDR images and then tone map to SDR. We display the HDR and SDR images using a custom-built HDR display coupled with eye tracking to determine the impact of dynamic range on the appearance of physically-based rendered surface gloss.

Experiments

To investigate the effects of image dynamic range on apparent gloss and to discover the relationships between eye movements and gloss perception, we conducted a gloss scaling experiment. The stimuli, apparatus, and procedure are described in the following sections.

Stimuli

The stimuli used in the experiment consisted of computer graphics images of glossy objects with different material properties rendered in different lighting environments. The values for each of these dimensions are described below.



(a) ***(b)*** ***(c)***

Figure 2*: Column (a) ball/box, blob/uffizi, and bunny/uffizi stimuli used in the experiments. The backgrounds were eliminated in the actual stimulus images, columns (b,c). The (b,c) columns show the gloss ranges used in the experiments for the (top) least/most glossy balls, (middle) least/most glossy blobs, and (bottom) least/most glossy bunnies. Note that the gloss differences are significantly compressed in these printed versions.*

Materials: Gloss is an attribute of many surfaces, including wood, plastics, papers, and paints. Eventually we would like to study each of these materials. However, in these experiments we used measured data on the reflectance properties (BRDFs) of achromatic latex paints. The BRDFs were modeled using the Ward[26] light reflection model:

$$\rho(\theta_i,\phi_i,\theta_o,\phi_o) = \frac{\rho_d}{\pi} + \rho_s \cdot \frac{\exp\left[-\tan^2\delta/\alpha^2\right]}{4\pi\alpha^2\sqrt{\cos\theta_i\cos\theta_o}} \quad (1)$$

where $\rho(\theta_i,\phi_i,\theta_o,\phi_o)$ is the surface BRDF, θ_i,ϕ_i, and θ_o,ϕ_o are spherical coordinates for the incoming and outgoing directions, δ is the half-angle between them, and α is the standard deviation (RMS) of the surface slope.

The three reflectance parameters of the Ward model are ρ_d (the diffuse reflectance), ρ_s (the energy of the specular lobe), and α (the spread of the specular lobe). We fixed ρ_d at 0.19 (mid-gray) and set α at 0.04 (small spread) to optimize visible gloss differences. To change surface gloss, we varied ρ_s across the range indicated in Table I. The range was selected to produce significant visible differences in apparent gloss from end to end. Representations of the endpoint images for each range are shown in Figure 2 (note that the visible differences are significantly compressed in the printed images). The step sizes were selected to be small enough to produce confusion between adjacent steps, which is necessary for the Thurstonian scaling analysis we used.

Table I. The ρ_s values for the HDR (H) and SDR (S) experimental images. Identification used in the paper is noted.

Specular energy (ρ_s)										
H1/S1	H2/S2	H3/S3	H4/S4	H5/S5	H6/S6	H7/S7	H8/S8	H9/S9	H10/S10	H11/S11
0.019	0.026	0.033	0.041	0.048	0.056	0.065	0.073	0.082	0.091	0.101

Geometry: Recent studies point to the importance of mesoscale surface texture in the perception of material properties.[21] To investigate this issue we studied three object geometries, a smooth sphere (ball), a bumpy sphere (blob), and a 3D laser scan of a ceramic rabbit (bunny).[27]

Illumination: Recent studies have also demonstrated the importance of real-world illumination for the accurate perception material properties.[28,29] To investigate this issue, we rendered images of scenes using two illumination environments, a simple square area source and Debevec's "Uffizi" HDR illumination map that captures the illumination field outside the Uffizi Museum in Florence, Italy.[30]

Rendering: The combinations of geometry and illumination yielded the three scene sets: ball/box, blob/Uffizi, and bunny/Uffizi shown in Figure 2. The scenes were rendered using the Radiance rendering system.[31] Fore and aft planes of the viewing camera were set to clip the surrounding environment. Thus, the illuminated objects appeared solely against a black background. Image size was 600x600 pixels and the images were saved as linear floating point high dynamic range HDR images. Eleven images of each scene were rendered using the parameters listed in Table I. The HDR images were scaled so the maximum luminance in each image set was 760 cd/m^2.

Tone mapping: To evaluate the effects of image dynamic range on apparent surface gloss, the rendered HDR image sets were duplicated and a sigmoidal tone mapping operator was applied to the duplicate sets to create corresponding standard dynamic range (SDR) images.[32] The aim of the basic tone mapping operator was to compress the HDR images to mimic the 160:1 measured dynamic range of a standard Apple 30" display (6x6 checkerboard VESA method[33]). The HDR and SDR image sets had the same mean luminance but differed in their overall dynamic ranges. Both image sets were mapped linearly to the display.

Apparatus

Eye tracker: An Applied Science Laboratories (ASL) Model 6000 desktop eye tracker with a Desktop Optics D6, Version 2.7 optics module and camera (120 Hz) was utilized in the experiment. The equipment was controlled with ASL Eye-Trac 6 User Interface Program software and a nine-point calibration was generated for each observer. Prior to the experiment, a nine-point calibration to a 25 × 10 degree onscreen grid target was generated for each observer. The accuracy of the calibration was verified with a 25-point target, indicating <0.9 degrees accuracy within the nine-point region and <1.6 degrees in the near surround. ASL Eyenal analysis software (version 2.106) was used for evaluating the results with default settings used for the fixation calculations.

Display: The images were shown on a custom-made high dynamic range display (see Figure 3) built from components of a 30" Apple Cinema Display with 2560 x 1600 addressable pixels and a pair of Planar PR5022 DLP projectors. The commercial backlight was removed from the LCD and substituted with backlighting from the two projectors which were rotated ±90° and tiled behind the LCD. The rotated, tiled projectors provided a backlight resolution of approximately 1500 x 935 addressable pixels. Images from projectors and on the LCD display were aligned geometrically and corrected colorimetrically using custom camera-based calibration software.[34] The maximum luminance of the display was 760 cd/m^2 with minimum luminance of 0.018 cd/m^2 for a small black center region surrounded by a completely white field. This translates to an approximate contrast ratio of 40,000:1 for the case of an image with maximum flare, but would increase for less-extreme cases. The display was driven by a Mac Pro 3 with a Quad-Core Intel Xeon processor and a Dual NVIDIA GeForce 8800 GT graphic card. Matlab software (version R2007b) and the Psychophysics toolbox[35] (version 3.0.8) were used to control stimulus presentation. Remote Desktop software on a MacBook Pro was used to coordinate the eye tracker and HDR display systems.

***Figure 3**: Experimental set up: (left) front view of HDR display showing LCD panel and adjacent eye tracker system, (right) rear view of HDR display showing tiled DLP projectors.*

Procedure

Method: The 11 HDR and 11 SDR image sets of the ball/box, blob/Uffizi, and bunny/Uffizi scenes were presented in succession to each observer using a paired-comparison method. On each trial, two of the glossy objects listed in Table I were randomly paired and the observer was asked to identify which object looked glossier (shinier or more reflective). For each scene, 253 image pairs were presented which represents a full factorial of the image set, including self-pairs. Thus, over the course of the experiment, each HDR image in the set was compared to every other HDR and SDR image and each SDR image was compared to every other HDR and SDR image as well as each image to itself.

Observers: Twenty-three observers participated in the experiment (11 female, 12 male). The group included both imaging and non-imaging faculty, staff, and students. Observers ranged in age from 17 to 47 and had normal color vision and acuity. Viewing distance was fixed at 25 inches with use of a headrest. Stimuli subtended approximately 13° of visual angle per object in the scene. Based on the quality of the 9-point eye tracking calibration grid, the eye movement for 12 observers was selected for scanpath and fixation analysis.

Results

In the following sections, we first present the results of the gloss scaling study then present the results of the eye movement analysis.

Effects of image dynamic range on apparent gloss

Thurstonian scaling methods were used to derive scales of apparent gloss for the objects represented by the HDR and SDR images.[36] In Thurstonian scaling, the variance in the paired comparison judgments is used to calculate response distributions for each object. Overlaps in the distributions are then used as a measure of perceived distances between the objects.

The derived scales are summarized in Figures 4 through 6. Each graph shows the perceived gloss scale as a function of specular energy, ρ_s. The least glossy object has been normalized to have a value of 1. In each figure, the blue diamond glyphs represent the objects depicted by SDR images and the red square glyphs represent objects depicted by HDR images.

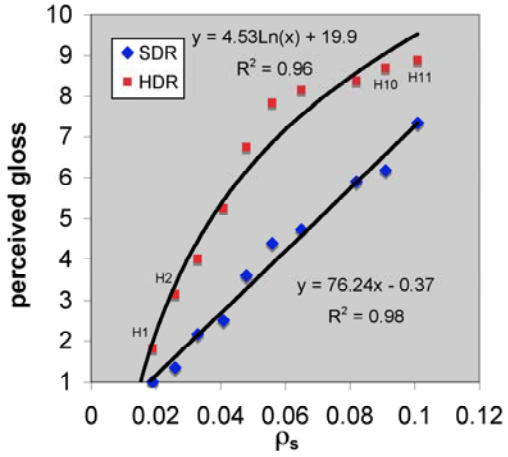


***Figure 4**: Scaling results for ball/box scene. Note that test objects with ρ_s of 0.073 were removed from the analysis because rendering errors were found in their images post experiment.*

Ball/box scene: Figure 4 shows the gloss scales derived for the ball/box scene. There are several trends to notice. First, note that the objects shown in the HDR images were all seen as glossier than the corresponding objects shown in the SDR images. This suggests that under these conditions (simple geometry, simple high contrast illumination), the brightness of the single specular highlight is an important cue to gloss. Next, observe that the overall range of the HDR object scale is larger (Δ7.1) than the corresponding SDR object scale (Δ6.3). This suggests that presentation in an HDR image makes the gloss differences between the objects more salient. Note that the perceived gloss response for the SDR images follows a linear fit (R^2 = 0.98). However, the scale for the HDR images appears to be compressive (natural log fit, R^2 = 0.96). The compressive behavior of the HDR gloss scale at the high end suggests that there may be limits to the effectiveness of specular brightness as a gloss cue. One possible explanation could be Weber's Law constraints on luminance JNDs. The rendered luminance differences between the highlights of the lowest gloss H1 and H2 objects is 24% (239 vs. 296 cd/m^2), but the difference for the highest gloss H10 and H11 objects is only 7.3% (708 vs. 760 cd/m^2). Thus, the brightest highlights may be less discriminable, which could lead to compression of the scale range.

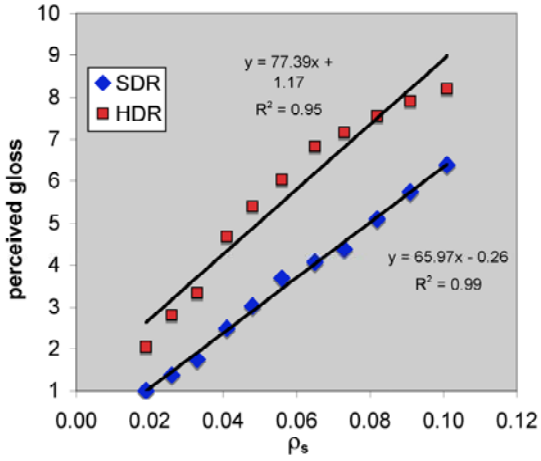


***Figure 5**: Scaling results for blob/Uffizi scene.*

Blob/Uffizi scene: Figure 5 shows the gloss scales derived for the blob/Uffizi scene. Our intent in testing this scene was to investigate how complex surface geometry and natural illumination interact with image dynamic range to affect the apparent gloss of depicted objects. Note that many of the trends seen in the ball/box data are also seen here. First, the objects in the HDR images are always seen as glossier than their corresponding SDR counterparts. Second, the slope of the regression line through the HDR data is higher than the line through the SDR data, suggesting that the objects in the HDR images are more discriminable. However an interesting difference between the ball/box and blob/Uffizi HDR datasets is that the blob/Uffizi data does not show as strong a compressive trend at the high end of the gloss scale. A possible explanation is that the complex object geometry breaks up the highlight reflection and intersperses them with off-specular shaded and shadowed regions, which may make it more difficult to use highlight brightness as a cue to gloss.

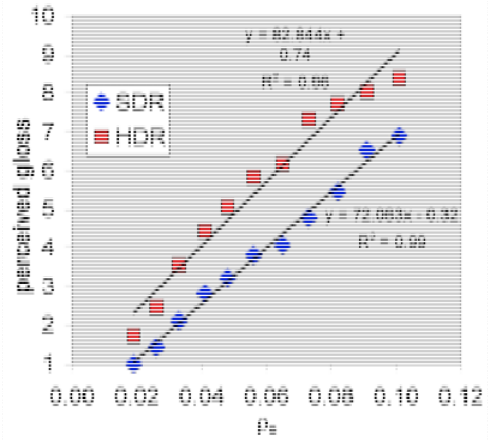


***Figure 6**: Scaling results for bunny/Uffizi scene.*

Bunny/Uffizi scene: Figure 6 shows the gloss scales derived for the bunny/Uffizi scene. As with the blob/Uffizi scene our intent was to investigate how interactions between complex geometry, natural illumination, and image dynamic range affect the perception of glossy rendered objects, but here the object had recognizable (and arguably more "natural") shape. The results are quite similar to those discussed for the blob/Uffizi scene and similar explanations apply.

In summary, the HDR/SDR image gloss scaling experiment presents three main findings: 1) that image dynamic range does affect the apparent gloss of surfaces depicted in the images, and that glossy objects shown in HDR images are perceived to have higher gloss than objects shown in SDR images; 2) that gloss differences are more discriminable in HDR images than in SDR images; and 3) that surface geometry and environmental illumination modulate these effects.

Eye tracking

On the basis of the gloss scaling study, we now know that observers are able to discriminate differences between the glossy objects and between the HDR and SDR renderings of these objects. Since we recorded the observers' eye movements during the gloss discrimination task, we can now analyze where observers were looking during the task to make inferences about the image features they are using to make the discrimination. The results of scanpath and fixation analyses are presented below.

Scanpath analysis: Figures 7 through 10 show scanpaths recorded for the ball/box, blob/Uffizi, and bunny/Uffizi object pairs. The figures present the scanpaths recorded from individual subjects, but the patterns shown are representative of the overall group. Several results can be observed. First, because of randomization in the presentation, on any particular trial subjects may have seen HDR/HDR, SDR/SDR, HDR/SDR, or SDR/HDR image pairs but the scanpath patterns did not appear to vary with image dynamic range. Second, to perform the discrimination task, observers scanned back and forth between the objects, comparing similar regions in each object. Third, we observed that many subjects showed a particular strategy for performing the discrimination task in which they looked at the object on the left, then quickly scanned over to a similar area on the right, returning to a different area on the left and then repeating the sequence. Thus, the object on the left appears to have been regarded as the "reference" with the object on the right serving as the "comparison/test." Finally, the number of scans/fixations appeared to vary with object complexity. The simple ball object had the lowest number of scans per trial, the complex blob object had the most, and scan counts on the bunny object were intermediate. This may indicate differences in the level of difficulty subjects had performing the gloss discrimination task.

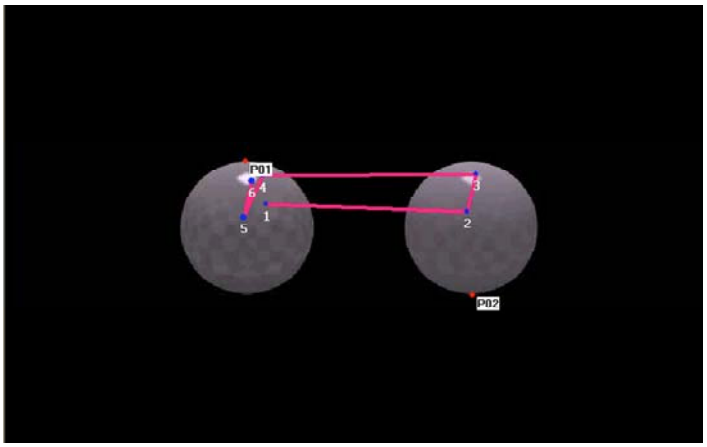


***Figure 7**: Scanpath for ball/box objects, Observer G.*

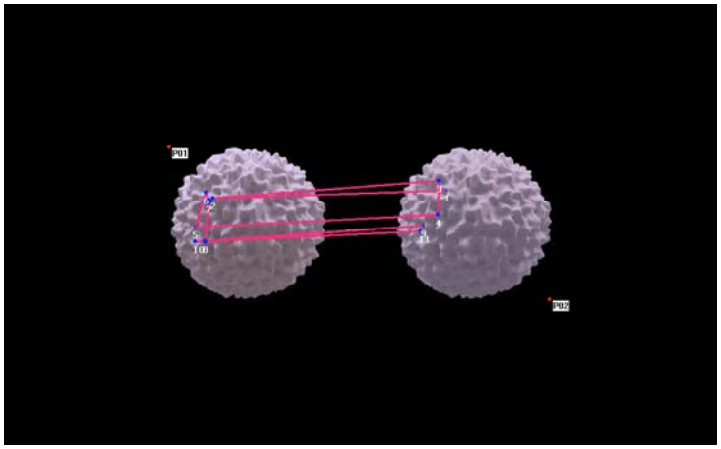


***Figure 8**: Scanpath for blob/Uffizi objects, Observer Q.*

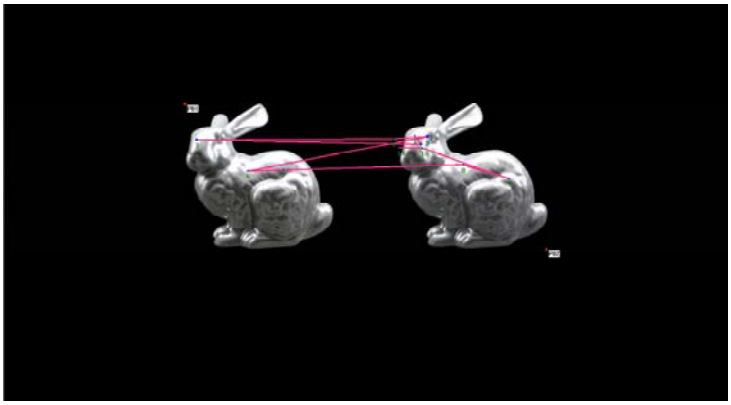


***Figure 9**: Scanpath for bunny/Uffizi objects, Observer R.*

Fixation analysis: Figures 10 through 14 show the results of fixation analysis performed on the scanpath data. To perform the fixation analysis, fixation records for each subject for each trial were calculated using the ASL software. The raw x,y coordinate data from the eye tracking hardware were clustered to identify fixations. Software thresholds were set so that a minimum of 6 data points of 100 ms within a 1.5 degree region were necessary to constitute a fixation. To combine fixation records across subjects, homography matrices based on the 9-point calibration targets measured for each subject were used to calculate projective transformations that registered the data to a common reference frame. Following this process, the fixation records for the 12 subjects were combined. Ten-pixel by ten-pixel binning of

this data was used to the plots shown in the figures. Due to limitations with synchronization of the HDR display and the eye tracking system, the fixation records were only resolvable by object type. Therefore, the distributions illustrated in the figures represent the fixation patterns to the left and right objects consolidated across material properties and HDR/SDR image type. Several interesting findings emerge.

Ball/box scene: Figure 10 shows the fixation distributions for the ball/box objects. The distributions are similar for the left and right ball; are centered on each ball (which may have been the easiest region to fixate given the headrest constraint); and are distributed longitudinally on the surface. The highest levels of fixation occur at the lower border of the specular highlight suggesting that subjects may be using information given by the brightness of the highlight, and the contrast and sharpness of the edge between the highlight and non-highlight region as information for gloss. The fixation distribution also spreads downward into the region below the highlight that is reflecting the checkerboard surround. Fixations within this region suggest that subjects are also using the contrast and sharpness of these interreflections as gloss cues. In fact, there were individual differences in the patterns of fixation recorded, and Figure 11 shows the fixation record for one subject who focused entirely on interreflections and ignored the specular highlights.

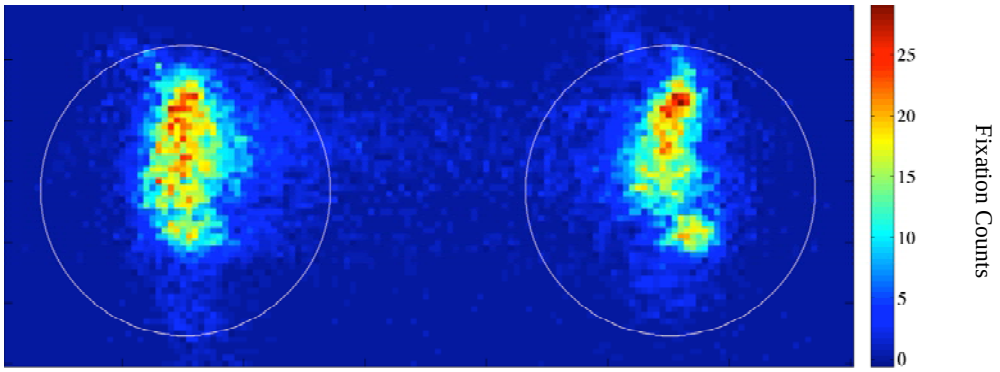


Figure 10: *Combined ball/box fixation count distributions for 12 subjects. The color scale indicates fixation counts within each 10 × 10 pixel region. The outlines indicate the boundaries of each object.*

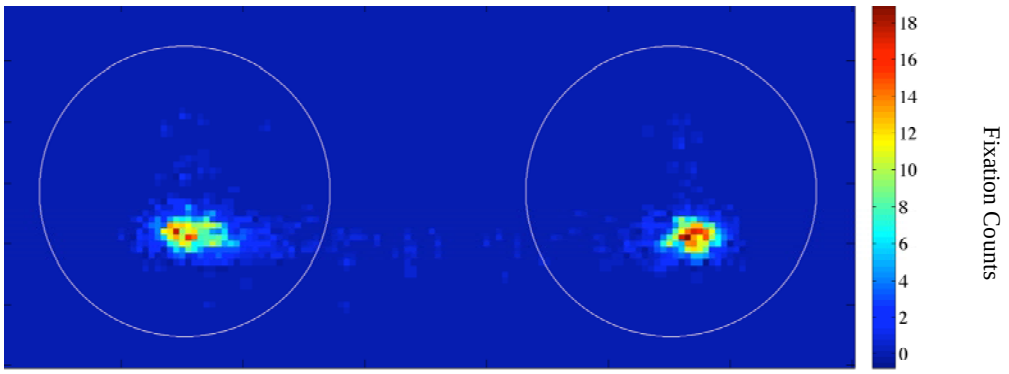


Figure 11: *Ball/box fixation count distributions for Observer L. Note that the fixations are focused in an area corresponding to interreflections rather than specular highlights.*

Blob/Uffizi scene: The fixation distributions for the blob/Uffizi objects are shown in Figure 12. As with the results for the ball/box objects, the distributions for the left and right objects are very similar. Here, however, most fixations are focused within a particular region of the blob surface. Geometrically this region has an open pattern of convex nodes, and photometrically it is above the blob's equator so it shows light source highlights *and* shadows as well as interreflections. These properties produce a rich, high contrast pattern of reflections (higher than other regions). Our hypothesis (which needs further testing) is that this region is somehow optimally informative about surface gloss for this kind of object and is therefore diagnostic for the gloss discrimination task.

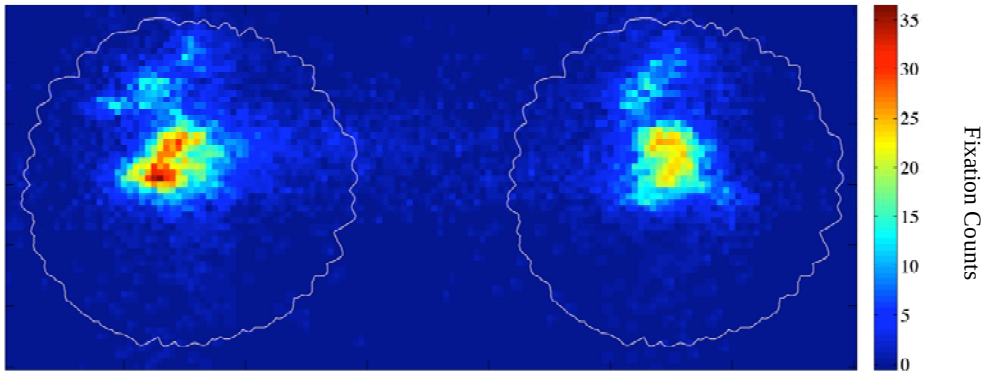


Figure 12: *Blob/Uffizi fixation count distributions for 12 subjects.*

Bunny/Uffizi scene: Figure 13 shows the fixation distributions for the bunny/Uffizi objects. Here, the regions with highest fixation counts were the specular reflections on the bunny's back, which again have both high luminance and high luminance contrast. To a lesser extent, other similar regions such as the face and ear were used, and again there were individual differences in the preferred region. For example, Figure 14 shows the fixation record for one observer who in addition to using the back highlights also focused on the grating-like pattern of specular reflections and shadows in the foot region.

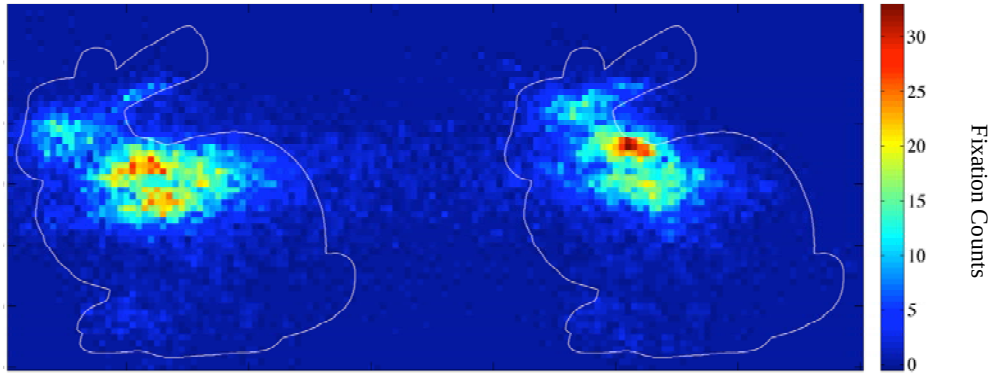


Figure 13: *Bunny/Uffizi fixation count distributions for 12 subjects.*

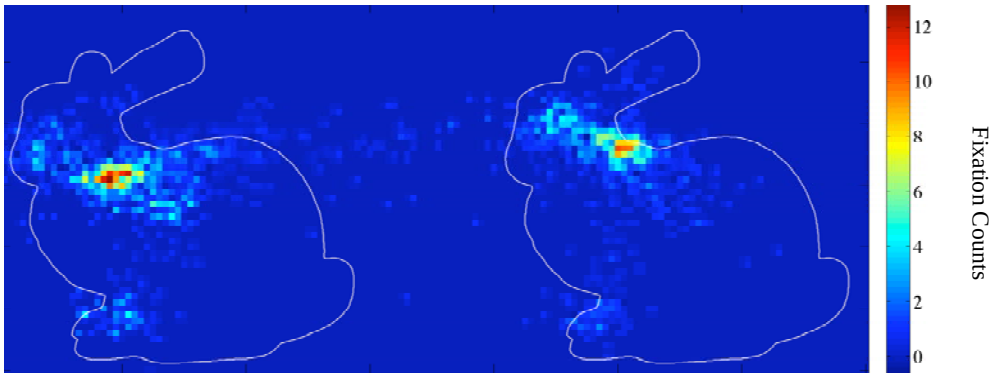


Figure 14: *Bunny/Uffizi fixation count distribution for Observer Q.*

Conclusions

In this paper, we presented results from a pair of psychophysical studies that investigate gloss perception. In the first part, we investigated the effects of image dynamic range on apparent gloss using a gloss discrimination task. In the second part, we recorded eye movements made during the gloss discrimination task and used scanpath and fixation analyses to reveal the image features subjects are

using to perform the task. The main findings are: 1) image dynamic range has significant influence on perceived surface gloss, with surfaces presented in HDR images being seen as glossier and more discriminable than their SDR counterparts; 2) subjects look at both surface highlights and body reflections when judging surface gloss; and 3) both of these results are modulated by surface geometry and scene illumination. The gloss scaling results suggest that the appearance of glossy objects is different in HDR and SDR images and that the HDR images, by correctly rendering highlight intensity, may provide higher-fidelity representations. This needs to be confirmed through comparison of real objects and HDR renderings in future work. The eye movement studies offer some initial insights into the image features observers use to perceive and discriminate surface gloss and suggest future studies to investigate the information content and limits of different features and their effects on gloss perception.

Acknowledgments

This work was supported by NSF/CPA-0811680 award to James A. Ferwerda, and a grant from the Eastman Kodak Company. The authors thank Jeff Pelz for the loan of the eye tracking equipment and subsequent training and consultation. The authors also thank the observers who participated in the experiments.

References

[1] Phillips, J. B., Ferwerda, J. A., Luka, S., "Effects of Image Dynamic Range on Apparent Surface Gloss," Proc. CIC17 25, 193-197 (2009).
[2] Ingersoll, L.R., "A means to measure the glare of paper," Electr. World 63, 645-647 (1914).
[3] Hunter, R.S., "Methods of Determining Gloss," J. Res. NBS 18(77), 281 (1937).
[4] Judd, D.B., "Gloss and glossiness," Am. Dyest. Rep. 26, 234-235 (1937).
[5] Billmeyer, F.W. and O'Donnell, F.X.D., "Visual gloss scaling and multidimensional scaling analysis of painted specimens," Color Res. App. 12(6), 315-326 (1987).
[6] Christie, J.S., "Evaluation of the attribute of appearance called gloss," CIE Journal 5(2), 41-56 (1986).
[7] McCamy, C.S., "Observation and measurement of the appearance of metallic materials. Part I. Macro appearance," COLOR research and application 21, 292-304 (1996).
[8] McCamy, C.S., "Observation and measurement of the appearance of metallic materials. Part II. Micro appearance," COLOR research and application 23, 362-373 (1998).
[9] Sève, R., "Problems connected with the concept of gloss," Color Res. & Appl. 18, 241-252 (1993).
[10] Lozano, R.D., "A new approach to appearance characterization," Color Res. & Appl. 31, 164-167 (2006).
[11] Arney, J. S., Smith, C., MacDonald, C., "A Gonio-photometric Analysis of Ink Jet Bronzing," NIP24 24, 432-434(2008).
[12] Farnand, S., Wilosek, D., "Measurement of Differential Gloss Using a Micro-goniophotometer," NIP23 23, 368-372 (2007).
[13] Gatt, A., Westland, S., Bala, R., "Increasing the Dimensionality of Soft Proofing: Gloss and Material," CIC06 14, 78-83 (2006).
[14] Kam-Ng, M., Reed, P., "Diagnosis for Differential Gloss in Color Hard Copy Outputs," NIP20 20, 450-453 (2004).
[15] Mehta, P., Johnson, K., Wolin, D., "A New Method of Measuring Gloss Mottle and Micro-Gloss using a Line Scan CCD-Camera Based Imaging System," NIP17 17, 714-717 (2001).
[16] Ng, Y. S., Wang, J., "Gloss Uniformity Attributes for Reflection Images," NIP17 17, 718-722 (2001).
[17] Nishida, S. and Shinya, M., "Use of image-based information in judgements of surface reflectance properties," J. Opt. Soc. Am. 15(12), 2951-2965 (1998).
[18] Pellacini, F., Ferwerda, J. A., Greenberg, D. P., "Toward a psychophysically-based light reflection model for image synthesis," Proc. SIGGRAPH, 55-64 (2000).
[19] Ferwerda, J. A., Pellacini, F., Greenberg, D. P., "A psychophysically-based model of surface gloss perception," Proc. SPIE 4299, 291-301 (2001).
[20] Fleming, R. W., Dror, R. O., Adelson, E. H., "Real-world illumination and the perception of surface reflectance properties," Journal of Vision 3(5), 347-368 (2003).
[21] te Pas, S. F., Pont, S. C., "A comparison of material and illumination discrimination performance for real rough, real smooth and computer generated smooth spheres," Proc. APGV, 75-81 (2005).
[22] Vangorp, P., Laurijssen, J., and Dutré, P., "The influence of shape on the perception of material reflectance," Proc. SIGGRAPH 26, 77-85 (2007).
[23] Buswell G.T., [How People Look at Pictures], Chicago Press, Chicago,(1935).
[24] Yarbus, A. L., [Eye Movements and Vision], Plenum, New York, (1967).
[25] Bloj, M., Harding, G., Chalmers, A., "Exploring eye movements for tone mapped images," Proc. SPIE 7240, 1-11 (2009).
[26] Ward, G. J., "Measuring and modeling anisotropic reflection," Proc. SIGGRAPH, 265-272 (1992).
[27] Turk, G., Levoy, M., "Zippered Polygon Meshes from Range Images," Proc. SIGGRAPH, 311-318 (1994).
[28] Fleming, R. W., Torralba, A., Adelson, E. H., "Specular reflections and the perception of shape," Journal of Vision 4(9), 798-820 (2004).
[29] Fleming, R. W., Dror, R. O., Adelson, E. H., "Real-world illumination and the perception of surface reflectance properties," Journal of Vision 3(5), 347-368 (2003).
[30] Debevec, P.E., Malik, J., "Recovering High Dynamic Range Radiance Maps from Photographs," Proc. SIGGRAPH, 369-378 (1997).

[31] Ward, G., "The RADIANCE lighting simulation and rendering system," Proc. International Conference on Computer Graphics and Interactive Techniques, 459 - 472 (1994).
[32] Tumblin, J., Hodgins, J. K., Guenter, B. K., "Two Methods for Display of High Contrast Images," ACM Transactions on Graphics 18(1), 56-94 (1999).
[33] Video Electronics Standards Association (VESA) display measurement standard, "Flat Panel Display Measurements (FPDM) Standard, Version 2.0," (2001).
[34] Luka, S. and Ferwerda, J. A., "Colorimetric Image Splitting for High-Dynamic-Range Displays," Proc. SID 40(1), 1298-1301 (2009).
[35] Brainard, D. H., "The Psychophysics Toolbox," Spatial Vision 10(4), 433-436 (1997).
[36] Torgerson, W.S., [Theory and Methods of Scaling], Wiley, New York, (1958).