Geometry of illumination, luminance contrast, and gloss perception

Frédéric B. Leloup,[1,2,*] Michael R. Pointer,[3] Philip Dutré,[2] and Peter Hanselaer[1]

[1]*Light and Lighting Laboratory, Catholic University College Sint-Lieven, Gebroeders Desmetstraat 1, B-9000 Gent, Belgium*
[2]*Computer Graphics Group, Department of Computer Science, K.U. Leuven, Celestijnenlaan 200A, B-3001 Leuven, Belgium*
[3]*Department of Color Science, University of Leeds, Leeds LS2 9JT, UK*
**Corresponding author: frederic.leloup@kahosl.be*

Received February 16, 2010; revised July 1, 2010; accepted July 26, 2010; posted July 26, 2010 (Doc. ID 124066); published August 18, 2010

The influence of both the geometry of illumination and luminance contrast on gloss perception has been examined using the method of paired comparison. Six achromatic glass samples having different lightness were illuminated by two light sources. Only one of these light sources was visible in reflection by the observer. By separate adjustment of the intensity of both light sources, the luminance of both the reflected image and the adjacent off-specular surroundings could be individually varied. It was found that visual gloss appraisal did not correlate with instrumentally measured specular gloss; however, psychometric contrast seemed to be a much better correlate. It has become clear that not only the sample surface characteristics determine gloss perception: the illumination geometry could be an even more important factor. © 2010 Optical Society of America

OCIS codes: 100.2960, 290.1483, 330.5020, 330.5510.

1. INTRODUCTION

The description of the appearance of objects and materials is a very difficult task to perform. Much work has been done to quantify color appearance, but there is still much to clarify about the perception of gloss. Physically, gloss results from directionally selective light scattering at the front surface of a material, with a preference toward the specular reflection direction. Instrumental gloss evaluation is generally performed with a standardized industrial device called a glossmeter [1,2]. This instrument compares the flux reflected from the surface to the reflected flux from a polished black glass with refractive index $n=1.567$ at 589.3 nm, measured in the same geometry at three predefined incident angles (20°, 60° and 85°). A gloss value of 100 gloss units is assigned to the black glass reference sample.

A major comprehensive study with respect to the relationship between the physical specular reflection and perceived gloss was conducted by Hunter and Harold [3]. Hunter and Harold defined six visual criteria that could possibly be used to rank gloss: specular gloss, sheen, contrast gloss, absence-of-bloom gloss, distinctness-of-image (DOI) gloss, and surface-uniformity gloss.

Based on this finding, Billmeyer and O'Donnell investigated the perceptual gloss dimensions on custom-prepared, painted specimens [4]. While, in all their experiments, at least three different visual criteria could be simultaneously identified, Billmeyer and O'Donnell concluded that for their specimen sets and test conditions, gloss appeared to be a unidimensional phenomenon. In addition, Billmeyer and O'Donnell investigated correlations between their obtained visual data and instrumental gloss measurements, and concluded that simple linear relations did not exist.

Two other investigations confirmed the nonlinear relationship between gloss perception and gloss measurement. Obein *et al.* [5] used the maximum likelihood difference scaling procedure [6] to estimate the glossiness of a series of ten black, coated samples. Ji *et al.* [7] employed a magnitude estimation technique to scale the gloss of a set of 84 neutral and colored samples, comparing them with a neutral reference sample that was assigned a gloss value of 50 gloss units. Visual gloss sensitivity was higher at extreme values, i.e., for matte and high-gloss samples. For the latter, Obein *et al.* argued that observers were bringing into account variations in the distinctness of the reflected image of the illumination, expressed as DOI. American Society for Testing and Materials (ASTM) defined DOI as an aspect of gloss characterized by the sharpness of images of objects produced by reflection at the surface [8].

In accordance with Billmeyer and O'Donnell, Obein *et al.* also investigated the influence of the illumination direction. Gloss difference scales were obtained for two incident angles, i.e., 20° and 60°. Although the angular distribution of the reflected luminous flux varied considerably in both illumination conditions, observers seemed to compensate for these variations and scaled the stimuli in a remarkably similar way. Obein *et al.* defined this ability as "gloss constancy" [5].

Computer graphics offer interesting opportunities to tackle the problem of visual gloss perception. To mimic light scattering from the surface, a reflectance model representing the bidirectional reflectance distribution func-

tion (BRDF) [9] is attributed to the material. Depending on the application a variety of BRDF models exist, and in psychophysical experiments the physically based model described by Ward [10] is frequently used. Three parameters are introduced: the specular reflectance ρ_s, the diffuse reflectance ρ_d, and the spread of the specular lobe α.

Wills *et al.* [11] performed a psychophysical study in order to construct a perceptual embedding of available BRDF measurements and related their embedding to gloss dimensions and input parameters of BRDF models, including that of Ward. The perceptual embedding seemed to be most successful by combining two gloss components, i.e., diffuseness and contrast gloss. Compared to BRDF models, the most obvious correlation was obtained with the parameter defining the diffuse reflectance. A strong correlation was also observed with the "roughness" parameters defining the spread of the specular lobe, such as Ward's parameter α.

Another advantage of computer rendering is the easy way in which the luminance distribution in the illuminating scene can be adapted. A complex real-world illumination scene with direct and indirect indoor and even outdoor contributions can be simulated. Especially for high-gloss surfaces, luminance contrasts within a reflected image and off-specular sample luminances engendered by the illuminating scene could influence gloss estimation. This idea was confirmed by Fleming *et al.*, who demonstrated that observers estimate lightness and gloss more reliably under complex realistic illumination than under a simple artificial light source [12,13]. Further investigations on the influence of collimated and diffuse illumination on gloss perception were recently reported by te Pas and Pont [14].

Furthermore, reports on the correspondence between gloss standards and components of BRDF models [15], on the multi-dimensionality of gloss perception [11,16,17], and on the interaction of surface gloss with 3D shape [18–20], color [21], and texture [22] have been published.

Computer-based research also suffers from some restrictions. The limited dynamic range of displays constitutes the main drawback. Although methods have been presented to map high-dynamic-range images on low-dynamic-range displays [23,24], digital imagery representations are not likely to allow for the investigation of any gloss attribute related to the absolute intensity of a highlight [17]. Furthermore, graphics studies are entirely based on a single BRDF model (Wills *et al.* excepted). For data-driven reflectance models, the set of parameters is particularly large and unintuitive [25]. Finally, displaying 3D simulations on flat displays is still problematic because of contradictions between pictorial cues and viewing geometry, screen self-luminosity, and other factors [26].

From this overview, it becomes evident that gloss perception of high-gloss surfaces is dominated by DOI on the one hand and the contrast between the virtual image of the illumination scene and its surroundings on the other. The surrounding surface luminance of the object is determined by the diffuse reflectance of the sample and the total illuminance originating from the entire illuminating scene. In this paper, an approach is proposed in that the different gloss-determining dimensions are isolated. The experimental goal is to find a new quantity that better corresponds to human visual gloss appraisal. Therefore, the influence of the luminance contrast between the virtual image of the illumination scene and the surroundings on perceived gloss is investigated. Visual gloss assessments of real achromatic glass samples, all having an identical DOI but a different diffuse back reflection, are described. To mimic complex real-world illumination conditions, an additional off-specular light source is introduced besides the generally adopted specular light source. Although this work is restricted to neutral samples with identical DOI, the use of a more complex illumination geometry offers the possibility to vary luminance contrast, a procedure that has not yet been applied to real samples.

2. EXPERIMENTAL SETUP

A. Samples

A set of six flat glass samples 10 cm × 10 cm and 3 mm thick was used. A mix of white and black matt paint was applied on the rear side of the samples. By varying the concentrations of both paints, approximately equal differences in lightness over the entire sample set could be obtained. This was verified by measuring the d:8° spectral reflectance of the front side with a Hunterlab UltraScan PRO sphere-based spectrophotometer. The sphere contains a specular exclusion port that can be opened during measurement. When the port is closed, the reflectance measurement includes the specular component of the reflected light (SPIN). When the port is opened, the specular component of the reflected light is excluded from measurement (SPEX). The difference between the two measurement results gives an indication of the glossiness of the measured sample.

All six samples were measured in both the SPIN and SPEX measurement modes. CIELAB colorimetric coordinates were calculated under CIE standard illuminant D65, in combination with the CIE 1964 standard colorimetric observer. SPEX mode results are represented in Table 1. Samples are numbered from 1 to 6 in decreasing order of lightness. Subtracting SPEX from SPIN values resulted in an identical specular reflectance value of 4% for each sample, originating from the front surface reflection of the glass at 8° angle of incidence.

Standard specular gloss measurements were performed at three basic geometries (20°, 60°, and 85° angle of incidence) using a Byk-Gardner micro-TRI-gloss-S glossmeter. Uniformity was checked by performing measurements at five different sample positions. The mean values and variances, expressed in specular gloss units (SGU), are also reported in Table 1. For each individual sample, gloss values increase with angle of incidence due to the increased Fresnel reflection. However, almost no differences are observed among samples within the same geometry, indicating that front surface reflection completely dominates the standard specular gloss measurement.

Spatial and spectral sample reflection characteristics were measured with a full three-dimensional BRDF instrument [27]. To mimic the geometrical conditions of the test booth (see below), the angle of incidence with respect to the sample normal was kept fixed at 60°. The viewing angle ranged from 0° to 85° in the opposite half-plane, 60°

Table 1. Colorimetric and Specular Gloss Characterization of the Six Glass Samples[a]

Sample Number	Sample Color	L*	a*	b*	Gloss 20° (SGU)	Gloss 60° (SGU)	Gloss 85° (SGU)
1	White	82.7	−1.76	1.19	86.6±0.3	93.4±0.2	99.4±0.1
2	Light gray	72.9	−2.25	−3.04	84.4±0.1	89.6±0.1	99.1±0.1
3	Light/Mid-gray	65.3	−2.41	−5.06	85.4±0.2	92.2±0.2	99.3±0.1
4	Mid-gray	46.1	−2.22	−6.35	86.9±0.1	92.5±0.1	99.3±0.2
5	Dark gray	22.8	−1.19	−6.00	83.6±0.1	88.1±0.1	98.9±0.1
6	Black	3.2	−0.05	0.08	85.2±1.0	91.2±0.6	99.2±0.3

[a]CIE L*a*b* values of samples obtained in SPEX mode and calculated under CIE illuminant D65 in combination with the CIE 1964 standard colorimetric observer. Average specular gloss values obtained in three basic geometries (20°, 60°, 85°), expressed in specular gloss units (SGU).

corresponding to the specular reflection direction (denoted −60°:60° geometry). In Fig. 1, BRDF values are shown for viewing angles ranging from 55° to 65°. Solid curves represent BRDF functions of the glass samples, while the dashed curve represents the BRDF function of the black glass reference sample of the specular glossmeter. All samples clearly show a similar specular lobe, corresponding to the instrument signature of the device and indicating an identical DOI of the samples without distortions in the reflected image [27]. At non-specular viewing angles, BRDF values tend to a constant aspecular value originating from the background reflection of the samples.

B. Test Booth

A picture of the test booth designed to perform the visual assessments is presented in Fig. 2. A white fluorescent lamp was covered by a diffuser. A black mask with an aperture of 25 cm by 4 cm resulted in a well-defined aperture with uniform luminance. According to the work of Fleming *et al.*, this rectangular light source generates images that appear more similar to images under "real-world" illumination than do other artificial light sources [13]. The light source is positioned at a distance of 60 cm from the sample, with an incidence angle of 60° toward the sample normal. This luminaire will be referred to as the *specular light source*.

A second fluorescent lamp is positioned perpendicular to the sample, again at a distance of 60 cm. This luminaire will be referred to as the *background light source*. A baffle between the two light sources prevents mutual illumination. The luminance of both dimmable light sources can be controlled.

Two samples are placed side-by-side on a sample holder and observed in the specular reflection direction of the specular light source. The observer's head is fixed by a chin rest, that guarantees both a well-defined viewing direction and a viewing distance of 100 cm. The field of view covered by the sample is approximately 4.4°. Illumination and viewing distances have been chosen such that the sample surface and the reflected image of the specular light source are both within the depth of focus of the eye.

C. 2D Luminance Meter

Luminance measurements of the samples were performed from the observer's viewing position with a 2D luminance camera (MURATest by Eldim) [28]. Measurements were analyzed by calculating the average luminance from the image of the specular light source, denoted by the image luminance L_i, and by calculating the average sample luminance from both regions next to this image, denoted by the background luminance L_b. A picture of the light/mid-gray sample (sample 3) under four illumination conditions is presented in Fig. 3. The image of the specular light source and the sample surroundings can be clearly discerned. The image luminance L_i includes both the specular front reflection and the diffuse reflection component of the rear side, while the background luminance L_b originates only from the diffuse rear-side reflection of the sample.

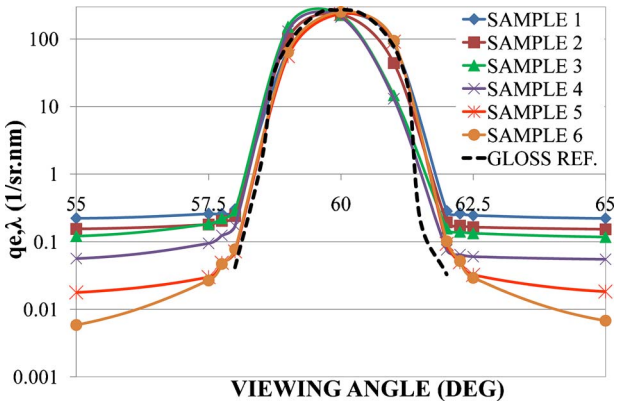

Fig. 1. (Color online) BRDF functions at an angle of incidence of −60° and at wavelength 589.3 nm. The viewing angle ranges from 55° to 65°. Solid curves represent the results obtained from the six test samples, while the dashed curve represents the black glass reference sample.

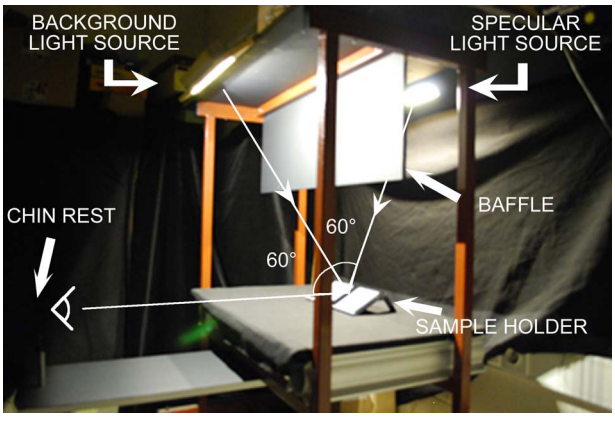

Fig. 2. (Color online) Side view of the test booth with specular and background light source.

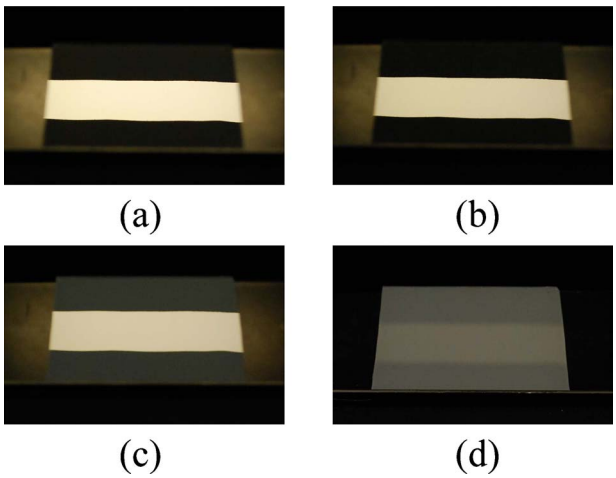

Fig. 3. (Color online) Pictures of sample 3 (light/mid-gray sample) under test conditions (a) A, (b) B, (c) C, and (d) D. All pictures were taken with the same exposure settings. The image of the specular light source and the off-specular surroundings can be clearly discerned.

3. PSYCHOPHYSICAL EXPERIMENTS

A. Method of Paired Comparison

The method of paired comparison, as described by Scheffé [29], was used to evaluate the effect of luminance contrasts on gloss appraisal. Paired comparison was preferred above other methods such as rating or ranking. The validity and reliability of rating data requires a very large number of trials and trained participants [30]. Pairwise comparisons may prove to be a more realistic approach to range alternatives than direct ranking [31]. Furthermore, paired comparison allows for the evaluation of transitivity, i.e., the intra- and inter-observer consistency.

Samples were presented in sets of two alternatives (i,j) to the observers, who answered the question "Rate the glossiness of the right sample i as compared to the left sample j, by use of the following preference scale; (3) i is much more glossy than j; (2) i is more glossy than j; (1) i is slightly more glossy than j; (0) i and j are of equal glossiness; (−1) i is slightly less glossy than j; (−2) i is less glossy than j; (−3) i is much less glossy than j." No recommendations were made about that visual cue observers had to rely on.

Six samples ($n=6$) and fifteen sample pairs ($M=15$) were presented for each illumination set up. All pairs were judged by ten observers. In total, eighteen male and sixteen female naïve observers participated in the experiments. All had normal or corrected to normal vision. No time restrictions were imposed during assessments, but the duration of each assessment was limited to approximately 20 min. to avoid observers losing attention.

To account for potential within-pair order effects, optimal order designs as proposed by Ross were applied [32,33]. These include the following properties: (i) all stimuli are compared with one another; (ii) every stimulus appears as often on the left as on the right of a pair; (iii) the order of pairs is such that the time gap between two consecutive presentations of the same stimulus is maximal.

From the recorded preference scores x_{ijk} of stimulus i over stimulus j evaluated by observer k, the mean preference $\hat{u}_{ij}$ for stimulus i over stimulus j was calculated as

$$\hat{u}_{ij} = 1/r \sum_{k=1}^{r} x_{ijk}, \tag{1}$$

with r half of the total number of observers for each specific illumination condition. As half of the observers rated the glossiness of stimulus i over stimulus j, the second half rated the glossiness of stimulus j over stimulus i. From both mean preferences $\hat{u}_{ij}$ and $\hat{u}_{ji}$, the average preference $\hat{\pi}_{ij}$ of stimulus i over stimulus j was calculated as

$$\hat{\pi}_{ij} = 1/2(\hat{u}_{ij} - \hat{u}_{ji}). \tag{2}$$

Finally, the mean glossiness estimate $\hat{\alpha}_i$ of stimulus i was computed as

$$\hat{\alpha}_i = 1/n \sum_{j=1}^{n} \hat{\pi}_{ij}. \tag{3}$$

As stimulus i was not rated against itself, $\hat{\pi}_{ii}=0$.

In order to evaluate the statistical significance of $\hat{\alpha}_i$, a confidence coefficient of 95% was employed to calculate the corresponding statistic $Y_{0.05}$ [29]. $Y_{0.05}$ is defined as

$$Y_{0.05} = q_{0.95}(S_e/4M(r^2 - r)n)^{1/2}. \tag{4}$$

$q_{0.95}$ is the upper 5% point of the Studentized range q for n variates and $2M(r-1)$ degrees of freedom, and can be deduced from tables [34]. S_e is the error sum of squares, defined as

$$S_e = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{r} (x_{ijk} - \hat{u}_{ij}). \tag{5}$$

Differences between two glossiness estimates $\hat{\alpha}_i$ and $\hat{\alpha}_j$ are statistically relevant at the 95% confidence level if they differ by more than the value of $Y_{0.05}$.

B. Analysis of Observer Consistencies

The method proposed by Kendall and Babington Smith [35,36] was used, respectively, to check the intra- and inter-observer consistency. Intra-observer consistency was examined by calculating the number of cyclic triads d occurring in an observed configuration of preferences. From this number, the coefficient of consistency ζ was calculated, which for an even number of stimuli n is defined as

$$\zeta = 1 - [24d/(n^3 - 4n)]. \tag{6}$$

Observers' results were rejected if their coefficient of consistency ζ was inferior to 0.75, which corresponds to an occurrence of more than two cyclic triads in their preference matrix ($n=6$). In turn, this limiting value was deduced from a 95% probability that there would occur more than two cyclic triads if decisions were made at random [36].

Investigation of the inter-observer agreement was considered by calculating the coefficient of agreement u, defined as

$$u = \frac{2\Sigma}{\binom{2r}{2} \cdot \binom{n}{2}} - 1. \quad (7)$$

Σ represents the sum of the number of agreements between all $\binom{2r}{2}$ possible pairs of observers, $\binom{n}{2}$ being the number of possible sample pairs [36]. In the case of complete agreement among observers, $u=1$.

4. RESULTS AND DISCUSSION

Gloss appraisal of the six glass samples was investigated under four illumination conditions, characterized by the sample illuminance originating from both the specular and the background light source (see Table 2).

In test condition A, only the specular light source was used to illuminate the samples. This condition corresponds to the illumination geometry used in all previous gloss experiments on real samples. Image and background luminances of the six samples, respectively L_i and L_b, are presented in Table 3. The luminance contribution of the specular front reflection can be calculated as the difference between these luminances. As expected, this value is almost invariant across all samples and on average has a value of 884 cd m^{-2}. Compared with the luminance of the light source, this value corresponds to a specular reflectance of 9%, in agreement with calculated specular reflectance values from both the BRDF measurements and the theoretical Fresnel equations at an incidence angle of 60°. The sample illuminance originating from the specular light source E_s was measured to be 250 lux (see Table 2). The values of L_b were verified by calculating the product of E_s and the off-specular BRDF values (−60°:58° geometry).

Results of paired comparisons in test condition A are gathered in Table 3. The average gloss estimates $\hat{\alpha}_i$, the statistic $Y_{0.05}$, and both the average coefficient of consistency ζ and coefficient of agreement u are presented. Although the specular glossmeter predicts the same gloss for all samples (see Table 1), visual gloss appraisal decreases with increasing lightness of the samples. Contrast thus seems to be a dominant factor for gloss perception. A closer look at the glossiness estimates of the lightest samples 1, 2, and 3 in combination with the value of $Y_{0.05}$ (0.334), however, reveals that no significant differences are demonstrated between sample pairs (1, 2) and (2, 3). Further analysis of these data showed that two subgroups of observers could be identified. One group indicated visual gloss differences between all samples, while the second group rated samples 1 to 3 as of equal glossiness. All second-group observers indicated that the intense image of the specular light source dominated their gloss perception, thereby masking the variation in contrast. This inconsistency among observers offers an explanation for the lower coefficient of agreement u.

Since contrast between the image luminance L_i and the background luminance L_b could be the dominant factor, the psychometric contrast C as defined by CIE [37] was calculated as

$$C = \frac{|L_i - L_b|}{L_b}. \quad (8)$$

Log C values are listed in Table 3. It should be noticed that the numerator represents the luminance corresponding to the external front reflection, which is equal for all samples. As a consequence, contrast variations among the six samples can totally be attributed to the variation of the background luminance L_b.

In Fig. 4, the average gloss estimates $\hat{\alpha}_i$ are plotted against log C, and a good linear correlation is obtained. The results are in accordance with the results obtained by Aida on paper samples [38], although Aida adopted another definition of contrast, i.e., $C=L_i/L_b$. However, for large values of the image luminance L_i, calculated values from the two contrast definitions become almost identical.

Visual scaling of gloss, using images on a computer display, was also performed by Ferwerda and colleagues, who rendered composite images of a sphere enclosed in a checkerboard box and illuminated by an overhead area light source. From their experiments, Ferwerda and colleagues identified two perceptually meaningful gloss dimensions: contrast gloss *CG* and DOI. From magnitude estimation experiments on both individual dimensions, Ferwerda and colleagues introduced the following relationship for *CG* [16,17]:

$$CG = \left(\rho_s + \frac{\rho_d}{2}\right)^{1/3} - \left(\frac{\rho_d}{2}\right)^{1/3}, \quad (9)$$

with ρ_s and ρ_d respectively the specular and diffuse sample reflectance. Contrast gloss *CG* as defined by Ferwerda and colleagues was also calculated for our samples,

Table 2. Description of Each Test Condition, Together with the Measured Total Sample Illuminance E_t (lux)[a]

Test Condition	Description	E_s (lux)	E_b (lux)	E_t (lux)
A	Specular light source: high Background light source: off	250	—	250
B	Specular light source: medium Background light source: off	130	—	130
C	Specular light source: medium Background light source: low	130	250	380
D	Specular light source: low Background light source: high	15	1985	2000

[a]The contribution originating from both the specular and the background light source, E_s and E_b, respectively, are indicated.

Table 3. Measurement Results for Test Condition A[a]

Sample Number	L_i (cd m^{-2})	L_b (cd m^{-2})	$L_i - L_b$ (cd m^2)	$\hat{\alpha}_i$	Log C	Contrast Gloss CG
1	930	46.9	883	−0.883	1.275	0.042
2	906	35.1	871	−0.600	1.395	0.050
3	925	27.7	897	−0.517	1.510	0.058
4	917	17.1	900	−0.033	1.721	0.086
5	868	8.4	859	0.633	2.010	0.143
6	899	6.7	892	1.400	2.127	0.226

Additional Paired Comparison Results

$Y_{0.05}$	ζ	u	range
0.334	0.89	0.28	2.283

[a]Image luminance L_i and background luminance L_b, gloss estimates $\hat{\alpha}_i$, log C, and contrast gloss CG are presented for all six samples. Additionally, the calculated statistic $Y_{0.05}$, the average coefficient of consistency ζ, the coefficient of agreement u, and the variation of $\hat{\alpha}_i$ are listed.

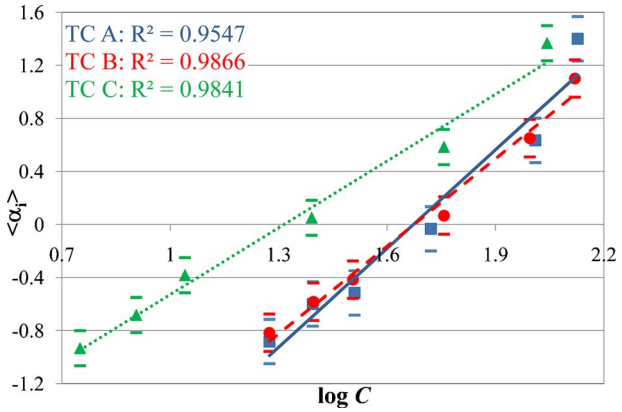

Fig. 4. (Color online) Gloss estimates $\hat{\alpha}_i$ plotted against calculated log C values for six glass samples. The squares and solid line, respectively, represent the results and linear fit in test conditions A (TC A). The circles and dashed line apply to test condition B (TC B). The triangles and dotted line apply to test condition C (TC C). Log C values increase with decreasing sample lightness. Error bars represent the calculated statistic $Y_{0.05}$. The coefficient of determination (R^2) indicates the correlation between the two variables.

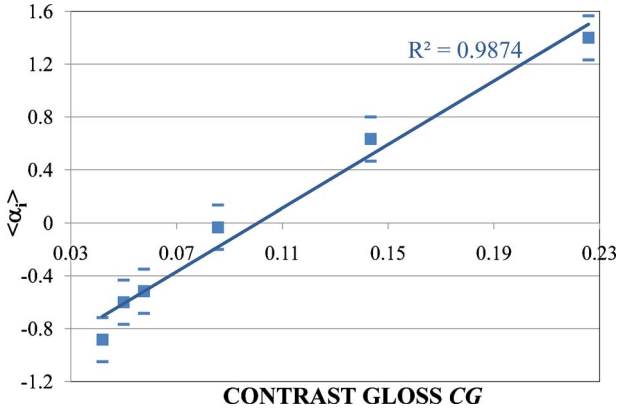

Fig. 5. (Color online) Gloss estimates $\hat{\alpha}_i$ plotted against calculated contrast gloss CG for six glass samples in test condition A. CG values increase with decreasing sample lightness. Error bars represent the calculated statistic $Y_{0.05}$. The coefficient of determination (R^2) indicates the correlation between the two variables.

for which ρ_d values were obtained from integrating sphere measurements in SPEX mode, while ρ_s was calculated to be 9% for all samples. Calculated CG values are presented in Table 3, while from Fig. 5, a linear correlation can be observed between visual gloss estimates $\hat{\alpha}_i$ and CG. Both log C and CG seem to be valuable candidates to describe visual gloss perception. However, the concept of contrast gloss as defined by Eq. (9) is relevant only if the samples are illuminated by the specular light source alone. The addition of the background light source will increase values of both L_i and L_b, while the values of ρ_s and ρ_d remain the same.

In test condition B, the luminance of the specular light source was decreased by 50%. This reduces both L_i and L_b by the same ratio, resulting in the same contrast values, as can be concluded from Table 4. Gloss estimates obtained by paired comparison reveal results similar to those in test condition A. Consequently, the plot of the average gloss estimates $\hat{\alpha}_i$ against log C (Fig. 4) remarkably resembles the plot for test condition A. This finding confirms that it is not the luminance of the specular image but the luminance contrast that is dominating the perception of gloss.

In test condition C, the specular source luminance was kept the same as in test condition B, but a background illumination was added. This increased the total sample il-

Table 4. Measurement Results for Test Condition B[a]

Sample Number	L_i (cd m^2)	L_b (cd m^2)	Log C	$\hat{\alpha}_i$
1	464	23.4	1.274	−0.817
2	450	17.4	1.396	−0.583
3	458	13.9	1.505	−0.417
4	496	8.5	1.758	0.067
5	433	4.3	1.995	0.650
6	450	3.4	2.120	1.100

Additional Paired Comparison Results

$Y_{0.05}$	ζ	u	range
0.282	0.81	0.20	1.917

[a]Image luminance L_i and background luminance L_b, log C, and gloss estimates $\hat{\alpha}_i$ are presented for all six samples. Additionally, the calculated statistic $Y_{0.05}$, the average coefficient of consistency ζ, the coefficient of agreement u, and the variation of $\hat{\alpha}_i$ are listed.

Table 5. Measurement Results for Test Condition C[a]

Sample Number	L_i (cd m^{-2})	L_b (cd m^{-2})	Log C	$\hat{\alpha}_i$
1	509	77.0	0.750	−0.933
2	487	53.9	0.906	−0.683
3	488	40.7	1.041	−0.383
4	506	19.7	1.391	0.050
5	437	7.5	1.757	0.583
6	451	4.1	2.043	1.367

Additional Paired Comparison Results

$Y_{0.05}$	ζ	u	range
0.260	0.86	0.48	2.300

[a]Image luminance L_i and background luminance L_b, log C, and gloss estimates $\hat{\alpha}_i$ are presented for all six samples. Additionally, the calculated statistic $Y_{0.05}$, the average coefficient of consistency ζ, the coefficient of agreement u, and the variation of $\hat{\alpha}_i$ are listed.

luminance E_t to 380 lux (see Table 2). The increase of the values of L_b was verified by calculating the product of the sample illuminance originating from the background light source E_b with the BRDF values at 0°:58° geometry. The luminances in the image and background zone, together with contrast values and gloss estimates, are gathered in Table 5. A plot of the gloss estimates $\hat{\alpha}_i$ against log C values is presented in Fig. 4. Compared with test condition B (Table 4), the addition of the background light source results in an increase of both L_i and L_b with the same amount for each individual sample. Consequently, the involved contrast values are lower and the range of contrast values has slightly increased. However, the range of estimated gloss values has increased as well, and the linear correlation between average gloss estimates $\hat{\alpha}_i$ and log C still holds. Taking into account the calculated statistic $Y_{0.05}$, only one sample pair, i.e., (1, 2), reveals no significant differences in gloss perception. In addition, the coefficient of agreement u has significantly increased. Lower contrast values seem to improve the sensitivity of perceiving gloss differences.

Finally, test condition D was evaluated. In this condition, the specular light source was dimmed further, while the background light source luminance was increased.

Table 6. Measurement Results for Test Condition D[a]

Sample Number	L_i (cd m^{-2})	L_b (cd m^{-2})	Log C	$\hat{\alpha}_i$
1	465	420	−0.970	−1.400
2	338	290	−0.781	−0.950
3	260	214	−0.668	−0.367
4	132	90.0	−0.331	0.242
5	62.0	23.0	0.229	0.958
6	45.0	3.8	1.035	1.517

Additional Paired Comparison Results

$Y_{0.05}$	ζ	u	range
0.177	0.94	0.60	2.917

[a]Image luminance L_i and background luminance L_b, log C, and gloss estimates $\hat{\alpha}_i$ are presented for all six samples. Additionally, the calculated statistic $Y_{0.05}$, the average coefficient of consistency ζ, the coefficient of agreement u, and the variation of $\hat{\alpha}_i$ are listed.

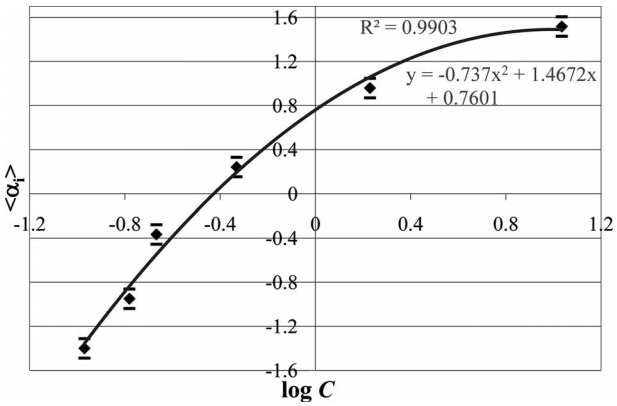

Fig. 6. Gloss estimates $\hat{\alpha}_i$ plotted against calculated log C values for the six glass samples in test condition D. Log C values increase with decreasing sample lightness. Error bars represent the calculated statistic $Y_{0.05}$. The coefficient of determination (R^2) indicates the correlation between the two variables.

This resulted in a total sample illuminance E_t of 2000 lux (Table 2). L_i and L_b values, log C, and average gloss estimates $\hat{\alpha}_i$ are reported in Table 6. The front reflection luminance contribution is reduced to only 44 cd m^{-2}. In this condition, the range of contrasts obtained over the six samples is higher, and test subjects now seem to observe much larger gloss differences between all samples. All gloss differences between consecutive samples are significant, with both a high coefficient of agreement u and average coefficient of consistency ζ.

As can be seen from Fig. 6, the linear correlation between gloss estimates $\hat{\alpha}_i$ and log C values no longer holds. The observers' gloss appreciation becomes very sensitive to contrast variation because the contrast approaches threshold values. In comparison to test conditions A to C, the contrast values are much lower, and consequently similar differences in log C have a larger impact on gloss discrimination. Clearly, the visual acuity to discern a reflected image is an additional visual cue. This explains both the increased coefficient of agreement u and the enhanced glossiness range.

Even for samples without distortions in the reflected image, gloss perception can be drastically influenced by varying the luminance distribution in the illumination scene. Only the light sources within the observer's specular field of view will contribute to the luminance of the reflected image. However, all light sources, even those positioned out of the specular field of view, will contribute to the luminance of the surface surroundings adjacent to the reflected image. Knowledge of the entire luminance distribution in the illumination scene and the full sample BRDF is required to calculate the contrast value, which is the main parameter determining gloss perception when DOI effects are not involved.

5. CONCLUSIONS AND FUTURE RESEARCH

Visual assessment of the appearance of objects and materials is a very complicated process, and four research areas have been suggested: color, gloss, texture, and translucency. Much work has been done to quantify color appearance, but there is still much to understand about

the perception of gloss, while the elaboration of visual correlates of texture and translucency is in its infancy. Until now, instrumental gloss evaluation has been restricted to the measurement of specular reflection characteristics of the sample. In this paper, it was shown that perceived glossiness also depends on the luminance distribution of the illumination scene. This result suggests the need for a new generation of gloss measurement instrumentation.

Gloss estimations were performed on six glass samples showing identical DOI without distortions in the reflected image. The reverse side of the samples was covered with a neutral paint, varying in lightness from white to black. Real complex illumination scenes were simulated using two light sources. The specular light source with an adjustable luminance created a virtual image, superimposed on a background luminance, that could also be varied independently using a second light source of that the reflected image was invisible. With this original experimental setup, contrasts between the reflected image and adjacent surroundings could be controlled. The method of paired comparison was adopted to obtain gloss estimates.

It was found that visual gloss assessment does not correlate with the specular gloss units as measured with commercial glossmeters. Previous investigations [5,7] already indicated a nonlinear relationship between gloss perception and specular glossmeter readings over a large range of specular gloss units. However, these results were obtained on samples having a different surface roughness, and consequently a different DOI and specular gloss value.

In this work, samples with identical DOI and specular gloss were investigated. It was shown that contrast between the virtual image of the illumination scene and its surroundings influences the gloss perception. With DOI kept invariant, it has become clear that psychometric luminance contrast seems to be a satisfactory correlate. Glossiness estimation becomes very sensitive to observed contrasts if contrast values approach threshold, as the potential discernibility of a reflected image becomes an additional visual cue to differentiate between samples. Luminance contrast is determined by the reflection characteristics of the sample. However, the sample illuminance and consequently the entire luminance distribution in the illumination scene around the sample strongly affect the gloss perception.

This study is only a first step in the development of a gloss perception correlate. Absolute magnitude scaling experiments in randomly generated illumination settings of the specular and background light source will be performed to work toward a new quantity to characterize gloss. In addition, the impact of DOI, the influence of the color of the sample, and the effect of the non-uniformity of the image light source must be included. Finally, investigations using real complex illumination and environment scenes combined with real 3D samples could be performed, in analogy to computer-based research.

REFERENCES

1. American Society for Testing and Materials Test Method D523, "Standard test method for specular gloss" (ASTM, 2008).
2. International Organization for Standardization Standard 2813, "Paints and varnishes—Determination of specular gloss of non-metallic paint films at 20 degrees, 60 degrees and 85 degrees" (ISO, 1994).
3. R. S. Hunter and R. W. Harold, *The Measurement of Appearance* (Wiley, 1987).
4. F. W. Billmeyer and F. X. D. O'Donnell, "Visual gloss scaling and multidimensional scaling analysis of painted specimens," Color Res. Appl. **12**, 315–326 (1987).
5. G. Obein, K. Knoblauch, and F. Viénot, "Difference scaling of gloss: nonlinearity, binocularity, and constancy," J. Vision **4**, 711–720 (2004).
6. L. T. Maloney and J. N. Yang, "Maximum likelihood difference scaling," J. Vision **3**, 573–585 (2003).
7. W. Ji, M. R. Pointer, R. M. Luo, and J. Dakin, "Gloss as an aspect of the measurement of appearance," J. Opt. Soc. Am. A **23**, 22–33 (2006).
8. American Society for Testing and Materials Standard E284, "Standard terminology of appearance" (ASTM, 2009).
9. F. E. Nicodemus, J. C. Richmond, and J. J. Hsia, "Geometric considerations and nomenclature for reflectance," Natl. Bur. Stand. Monogr. **160** (1977).
10. G. J. Ward, "Measuring and modeling anisotropic reflection," Comput. Graph. **26**, 265–272 (1992).
11. J. Wills, S. Agarwal, D. Kriegman, and S. Belongie, "Toward a perceptual space for gloss," ACM Trans. Graphics **28**, 103:1–103:15 (2009).
12. R. W. Fleming, R. O. Dror, and E. H. Adelson, "How do humans determine reflectance properties under unknown illumination?," in *Proceedings of the IEEE Workshop on Identifying Objects Across Variations in Lighting: Psychophysics & Computation* (IEEE, 2001), p. 347.
13. R. W. Fleming, R. O. Dror, and E. H. Adelson, "Real-world illumination and the perception of surface reflectance properties," J. Vision **3**, 347–368 (2003).
14. S. F. te Pas and S. C. Pont, "Both illumination and the material of context objects influence perceived glossiness," Perception **38**, ECVP Abstract Supplement, 97 (2009).
15. H. B. Westlund and G. W. Meyer, "Applying appearance standards to light reflection models," in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)* (ACM, 2001), pp. 501–510.
16. J. A. Ferwerda, F. Pellacini, and D. P. Greenberg, "A psychophysically-based model of surface gloss perception," Proc. SPIE **4299**, 291–301 (2001).
17. F. Pellacini, J. A. Ferwerda, and D. P. Greenberg, "Toward a psychophysically-based light reflection model for image synthesis," in *Proceedings of SIGGRAPH'00* (ACM, 2000), pp. 55–64.
18. R. W. Fleming, A. Torralba, and E. H. Adelson, "Specular reflection and the perception of shape," J. Vision **4**, 798–820 (2004).
19. J. F. Norman, J. T. Todd, and G. A. Orban, "Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information," Psychol. Sci. **15**, 565–570 (2004).
20. P. Vangorp, J. Laurijssen, and P. Dutré, "The influence of shape on the perception of material reflectance," ACM Trans. Graphics **26**, 77:1–77:9 (2007).
21. B. Xiao and D. H. Brainard, "Surface gloss and color perception of 3D objects," Visual Neurosci. **25**, 371–385 (2008).
22. Y.-H. Ho, M. S. Landy, and L. T. Maloney, "Conjoint measurement of gloss and surface texture," Psychol. Sci. **19**, 196–204 (2008).
23. J. Tumblin, J. K. Hodgins, and B. K. Guenter, "Two methods for display of high contrast images," ACM Trans. Graphics **18**, 56–94 (1999).
24. G. W. Larson, H. Rushmeier, and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," IEEE Trans. Vis. Comput. Graph. **3**, 291–306 (1997).
25. W. Matusik, H. Pfister, M. Brand, and L. McMillian, "A data-driven reflectance model," ACM Trans. Graphics **22**, 759–769 (2003).
26. Y. Ling and A. Hurlbert, "Color and size interactions in a real 3D object similarity task," J. Vision **4**, 721–734 (2004).

27. F. B. Leloup, S. Forment, P. Dutré, M. R. Pointer, and P. Hanselaer, "Design of an instrument for measuring the spectral bidirectional scatter distribution function," Appl. Opt. **47**, 5454–5467 (2008).
28. http://www.eldim.fr/products/uniformity-series/muratest-16m.
29. H. Scheffé, "An analysis of variance for paired comparisons," J. Am. Stat. Assoc. **47**, 381–400 (1952).
30. M. Kendall and J. D. Gibbons, *Rank Correlation Methods* (Griffin, 1975).
31. A. Agresti, *Categorical Data Analysis* (Wiley, 2002).
32. R. T. Ross, "Optimum orders for the presentation of pairs in the method of paired comparisons," J. Educ. Psychol. **25**, 375–382 (1934).
33. R. T. Ross, "Discussion: Optimal orders in the method of paired comparisons," J. Exp. Psychol. **25**, 414–424 (1939).
34. E. S. Pearson and H. O. Hartley, "Tables of the probability integral of the studentized range," Biometrika **33**, 89–99 (1943).
35. S. I. Gass, "Tournaments, transitivity and pairwise comparison matrices," J. Oper. Res. Soc. **49**, 616–624 (1998).
36. M. G. Kendall and B. Babington Smith, "On the method of paired comparisons," Biometrika **31**, 324–345 (1940).
37. Commission Internationale de l'Eclairage Technical Report 95:1992, "Contrast and Visibility" (CIE, 1992).
38. T. Aida, "Glossiness of colored papers and its application to specular glossiness measuring instruments," Syst. Comput. Japan **28**, 95–112 (1997).