The Perception and Misperception of Specular Surface Reflectance

Phillip J. Marlow, Juno Kim, and Barton L. Anderson

School of Psychology, University of Sydney, NSW, 2006, Australia

Summary

The amount and spectral content of the light reflected by most natural surfaces depends on the structure of the light field, the observer’s viewing position, and 3D surface geometry, particularly for specular (glossy) surfaces. A growing body of data has demonstrated that perceived surface gloss can vary as a function of its 3D shape [1–5] and its illumination field [6–12], but there is currently no explanation for these effects. Here, we show that the perception of gloss can be understood as a direct consequence of image properties that covary with surface geometry and the illumination field. We show that different illumination fields can generate qualitatively different patterns of interaction between perceived gloss and 3D surface geometry. Despite the complexity and variability of these interactions, we demonstrate that the perception (and misperception) of gloss is well predicted by the way that each illumination field modulates the size, contrast, sharpness, and depth of specular reflections. Our results provide a coherent explanation of the effects of extrinsic scene variables on perceived gloss [1–13], and our methods suggest a general technique for assessing the role of specific image properties in modulating our visual experience of material properties.

Results and Discussion

One of the main functions of visual processing is to provide information about the intrinsic properties of objects, surfaces, and materials. This information must be derived from the retinal images that are formed by the way that light is reflected from the bodies and surfaces of objects. Whereas the majority of work on perceived reflectance has focused on surfaces with idealized isotropic (Lambertian) reflectance functions [14–16], most natural surfaces exhibit complex reflectance functions that arise, in part, from specular reflections [17]. The structure, spectral content, and amount of light reflected specularly depend on the positions of the light sources, 3D surface geometry, amount of surface gloss, and the observer’s viewing position. Perfectly smooth glossy surfaces behave as mirrors, which reflect images of their surrounding environment distorted by the geometry of the reflecting surface. If the visual system behaved ideally, it would estimate gloss and other intrinsic material properties independently of the image structure contributed by the illumination field and 3D surface geometry. However, a growing body of data has revealed significant departures from this ideal [1–13]. The perception of surface gloss can vary significantly as a function of both its 3D geometry [1–5] and the illumination field in which it is embedded [6–12], which have no bearing on the intrinsic reflectance properties of a surface. There is currently no coherent explanation of these effects.

We reasoned that if irrelevant scene variables modulate the perception of surface gloss, then they must alter the image cues the visual system uses to generate our experience of gloss. Previous work has shown that the perceived gloss of a surface can vary as a function of its 3D surface relief [1, 2] or as a function of the particular illumination field in which it is embedded [6–12]. These effects suggest that it might be possible to identify the information the visual system uses to compute gloss by evaluating how perceived gloss varies as a function of a surface’s illumination field and 3D shape. Our approach was to measure perceived gloss across a range of scenes and search for proximal image cues that covary with perceived gloss. The goal was to create a complex and varied data set that could be used to assess the predictive capacity of any hypothesized cues and thereby reduce the risk of observing spurious correlations based on the particular scene variables we examined.

We constructed planar surfaces with a fixed level of gloss and varying degrees of surface relief (bumpiness) embedded in a number of different natural illumination fields (Figure 1). We included two illumination fields in which the surface was viewed along the same visual axis as the direction of the primary light source(s) and four illumination fields where the primary light source(s) illuminated the surface obliquely. This manipulation was included because the distribution of specular highlights depends critically on the surface geometry, the positions of the primary light sources, and the viewing position. Observers were shown a pair of surfaces with different relief levels and selected the surface that appeared glossier (see Experimental Procedures). Trials were blocked such that each bump level was compared with every other in the same illumination field (i.e., observers never directly compared surfaces across different illumination fields). Surfaces were viewed either with (Experiment 1a) or without (Experiment 1b) stereoscopic depth information (binocular disparity). Disparity locates specular reflections of convex surface patches behind the reflecting surface, which provides an additional source of information for identifying the image structure generated by specular reflections [18]. The results of these experiments are shown in Figure 2. In each graph, the percentage of time that a given image was chosen as glossier than its comparison is plotted as a function of the surface’s bump level. The black curves depict the results from Experiment 1a (with disparity), and the red curves depict the results from Experiment 1b (without disparity). The four panels on the left depict the results from the oblique illuminations, whereas the two panels on the right correspond to the two frontal illuminations. If these surfaces were perceived veridically, observers should generate flat response functions, and each surface in a pair should have a 50% chance of being selected. The data strongly violate this prediction. Perceived gloss varied up to 80% as a function of 3D surface relief within an illumination field. Importantly, the particular form of the interaction between perceived gloss and surface relief differed qualitatively in different illumination fields, ranging from monotonically increasing, monotonically decreasing, or a strongly nonmonotonic dependence on physical bump level.

*Correspondence: barton.anderson@sydney.edu.au
We observed that the interactions between surface relief and the illumination field led to systematic changes in four cues that have been shown to influence perceived gloss [18–23]. Specifically, we noted that the specular reflections systematically varied in contrast, sharpness, depth, and coverage. Contrast refers to the perceived "intensity" of specular reflections relative to the diffuse reflectance of a surface, sharpness refers to the relative spread of specular reflections (i.e., whether specular reflections generate sharp or blurred edges), depth refers to the perceived stereoscopic depth of the specular reflections, and coverage refers to the proportion of a surface that appears to be covered by specular reflections. Different studies have separately identified contrast [19–22], sharpness [20–22], depth [13, 18], and coverage [19, 23] as proximal cues for perceived gloss. We therefore tested whether the combined influence of these cues can account for the dependency of perceived gloss on shape and illumination field and thereby provide an explanation for the illusory variations in perceived gloss found in the large and growing body of work in gloss perception [1–13].

The goal of Experiment 2 was to test whether the complex interactions in our gloss data (Figure 2) are predicted by the way the four cues covaried with relief level in each illumination field. One problem in testing this hypothesis is that there is currently no model capable of computing these dimensions of specular reflections directly from images. Such computations require a method of segmenting specular reflections from other image structure (such as the shading from the surface's diffuse reflectance component) and then scaling each cue in a manner that quantitatively captures how each cue is psychophysically experienced. We reasoned that it should be possible to determine effects of these cues on perceived gloss by psychophysically measuring how each cue varies for all of our stimuli. If these dimensions of specular reflections are responsible for observers' gloss judgments, then the specific way that they covary with the illumination field and 3D surface shape should predict gloss judgments.

We chose three illumination fields from Experiment 1 that generated different patterns of perceived gloss by surface relief interactions (monotonically increasing, monotonically decreasing, or highly nonmonotonic). In Experiment 2, four groups of independent observers judged each of the four gloss cues, while a fifth group judged perceived gloss both with and without binocular disparity (see Experimental Procedures). Whereas observers in Experiment 1 only compared stimuli within an illumination field, observers in Experiment 2 compared all stimuli with each other stimulus in the set, which
The results of Experiment 2 are shown in Figure 3. The rightmost panels depict gloss judgments as a function of surface relief for the three different illumination fields viewed with disparity (top) and without disparity (bottom). The five panels on the left depict the four cues judged by each set of observers, while the fifth panel shows the histogram skew of each image, which is not a property observers can estimate directly but has been previously proposed to contribute to perceived gloss [24]. We tested whether gloss judgments obtained with and without binocular disparity could be modeled as a simple weighted (linear) combination of these five cues. The optimum combination was found by testing all weight combinations in steps of 0.5%, where the cue weights were constrained to sum to 100%. The weights for the no-disparity condition were created by assigning disparity a weight of zero and proportionally rescaling the remaining weights by dividing by the sum of the nonzero cue weights. The ratio of the weights for contrast, sharpness, coverage, and skew was therefore identical for the disparity and the no-disparity conditions for each weight combination tested. The weight combination that accounted for the most variance in perceived gloss was determined by summing the residual variance (least-squares) for both the disparity and the no-disparity conditions. The dashed curves in the two rightmost panels represent the best-fitting weighted average of these cues, which account for 94% of the variance in perceived gloss. Skew did not contribute to the fits and therefore received zero weight. The center of the figure shows the weights assigned to the other four cues to fit the gloss data for the disparity condition (upper right graph). The no-disparity gloss condition was fit by assigning zero weight to the disparity cue and normalizing the weights for coverage, contrast, and sharpness by the sum of these three weights. The resulting weights for the no-disparity condition were 48% coverage, 29% sharpness, and 23% contrast. Error bars represent SEM.

The results of Experiment 2 are shown in Figure 3. The rightmost panels depict gloss judgments as a function of surface relief for the three different illumination fields viewed with disparity (top) and without disparity (bottom). The five panels on the left depict the four cues judged by each set of observers, while the fifth panel shows the histogram skew of each image, which is not a property observers can estimate directly but has been previously proposed to contribute to perceived gloss [24]. We tested whether gloss judgments obtained with and without binocular disparity could be modeled as a simple weighted (linear) combination of these five cues. The optimum combination was found by testing all weight combinations in steps of 0.5%, where the cue weights were constrained to sum to 100%. The weights for the no-disparity condition were created by assigning disparity a weight of zero and proportionally rescaling the remaining weights by dividing by the sum of the nonzero cue weights. The ratio of the weights for contrast, sharpness, coverage, and skew was therefore identical for the disparity and the no-disparity conditions for each weight combination tested. The weight combination that accounted for the most variance in perceived gloss was determined by summing the residual variance (least-squares) for both the disparity and the no-disparity conditions. The dashed curves in the two rightmost panels represent the best-fitting weighted average of these cues, which account for 94% of the variance in perceived gloss. Skew did not contribute to the fits and therefore received zero weight. The center of the figure shows the weights assigned to the other four cues to fit the gloss data for the disparity condition (upper right graph). The no-disparity gloss condition was fit by assigning zero weight to the disparity cue and normalizing the weights for coverage, contrast, and sharpness by the sum of these three weights. The resulting weights for the no-disparity condition were 48% coverage, 29% sharpness, and 23% contrast. Error bars represent SEM.
To further explore the capacity of these cues to predict perceived gloss, we performed two additional experiments in which we factorially varied gloss level and relief height. In Experiment 3, different groups of observers performed the same tasks as Experiment 2 (using nonstereoscopic stimuli). Experiment 4 was a control experiment to test whether observers’ judgments of the individual gloss cues were contaminated by the perception of gloss. Although the different patterns of data exhibited by each cue makes this interpretation unlikely, Experiment 4 required observers to judge the coverage, sharpness, and contrast of the luminance minima in the photographic negatives of the gloss stimuli, which do not elicit a percept of gloss (see Supplemental Information available online). We used the illumination field containing the single extended light source so that the specular highlights in the original stimuli became luminance minima in the inverted stimuli. The results of these experiments support the results of Experiments 1 and 2. The same cues account for 97% of the variance in Experiment 3 (Figure S1), and 95% of the variance for the cues measured in the photographic negatives (Experiment 4, Figure S1). In keeping with a growing body of data on cue combination [25–29], the weights assigned to each cue differed between experiments, which likely reflect their relative utility in differentiating the stimuli being compared (see Supplemental Results). Observers mainly relied on coverage to judge gloss for images that mainly differed in coverage (Figure 3), whereas they relied more on sharpness and contrast for stimuli that produced larger variations in those two cues (Figure S1).

Our results suggest that the four cues we evaluated mediate illusory variations in perceived gloss, but it does not specify how these cues can be computed directly from images. A full solution to this problem requires a method for segmenting the image structure contributed by specular reflections from all other sources of image variation. Previous work has shown that the identification of specular highlights depends on their compatibility with the diffuse shading profile [18, 19, 30], but there is currently no model that can extract specular structure directly from images for arbitrarily complex illumination fields and shapes. Although a full solution to this modeling task is beyond the scope of this paper, we have developed a model that can compute these cues for the simplest illumination field (the extended light source used in Experiment 3), which are all highly correlated with the psychophysical judgments of each of these gloss cues (see Supplemental Information). We are currently engaged in extending this work to arbitrary combinations of gloss, 3D shape, and illumination fields.

The rapid advances in simulating the reflectance properties of complex 3D surfaces in real world illumination fields have provided new and powerful tools for evaluating the perception of surface and material properties. These graphical environments provide a way to manipulate the physical properties of the environment directly, which can then be used to evaluate how well our perceptual systems recover complex material and surface attributes. Such approaches typically focus on understanding the extent to which our experience of surface and material properties is invariant to changes in other scene dimensions (“constancy”). Although much progress has been made using such approaches [1–13], a primary goal of perceptual theory is to understand how the visual system generates our experience of material and surface properties, independently of whether that experience represents a veridical or fallacious depiction of the physical world. The experiments reported herein demonstrate that the perceived gloss of surfaces can be highly unstable and vary as a function of both its 3D surface geometry and the illumination field in which it is embedded. We reasoned that if ostensibly irrelevant scene variables modulate our perception of gloss, then these scene variables must modulate the cues the visual system uses to construct our experience of gloss. Our data reveal that the perceived gloss of our stimuli can be well explained by a linear combination of cues to gloss that covary with physical gloss, the illumination field, and 3D geometry. These results suggest that the visual system relies on a set of imperfect cues to construct our experience of gloss and potentially provides a unified explanation of the diverse body of research [1–13] that has revealed the instability of perceived gloss to changes in surface geometry and viewing conditions. Future work is needed to assess the generality of these results for different surface geometries and light fields and to elaborate models that compute these cues directly from images in a manner that captures how each of these dimensions of specular reflections is experienced psychophysically.

Experimental Procedures

The surfaces were composed of 400 ellipsoids pseudorandomly positioned in a 20 cm square grid. The x and y axis radii of the ellipses were 1 cm and their z axis radii were random values in different ranges for each bump level: 0–0.4 cm, 0–0.9 cm, 0–1.6 cm, 0–2.5 cm, 0–3.8 cm. The surfaces were set frontoparallel to the observer and were viewed from a distance of 70 cm. Photorealistic images of the surfaces in six illuminations were rendered (YafaRay 0.1.1) with two diffuse and specular interreflections. Observers viewed a pair of stimuli through a mirror stereoscope on each trial and judged which stimulus appeared glossiest in Experiments 1a and 1b. For Experiment 2, four attributes of the specular highlights were measured by instructing observers to judge which stimulus had greater luminance contrast of the highlights (i.e., the luminance difference between highlights and their surrounds), sharpness of the edges of the highlights, proportion of surface area covered by highlights, or amount of depth separating the highlights from the reflecting surface. Methods are described in more detail in the Supplemental Experimental Procedures. Observers were recruited with the approval of the University of Sydney’s Human Research Ethics Committee.

Supplemental Information

Supplemental Information includes one figure, one table, Supplemental Results, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2012.08.009.

Acknowledgments

This project was funded by Australian Research Council Grants awarded to B.L.A.

Received: June 6, 2012
Revised: July 15, 2012
Accepted: August 3, 2012
Published online: September 6, 2012

References

glossiness transfer function induced by illumination change and testing
its transitivity. J. Vis. 10, 8, 1–9.
nation and the perception of surface reflectance properties. J. Vis. 3,
347–368.
perceived surface reflectance across different illumination statistics.
ness under real-world illumination. J. Vis. 10, 5.
gloss: nonlinearity, binocularity, and constancy. J. Vis. 4, 711–720.
the perception of solid objects. Perception 35, 1331–1350.
nation discrimination performance for real rough, real smooth and
computer generated smooth spheres. Proceeding APGV ’05, 75–81.
motion, and color information improve gloss constancy performance.
J. Vis. 10, 7.
15. Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X.,
a psychophysically-based light reflection model for image synthesis.
ACM Trans. Graph. 34, 55–64.
Remote effects of highlights on gloss perception. Perception 34,
565–575.
24. Ernst, M.O., and Banks, M.S. (2002). Humans integrate visual and haptic
texture and disparity cues: optimal cue combination. J. Vis. 4, 967–992.
26. Todd, J.T., Christensen, J.C., and Guckes, K.M. (2010). Are discrimina-
tion thresholds a valid measure of variance for judgments of slant from
texture? J. Vis. 10, 20–, 1–18.
and orientation congruence in the perception of surface gloss. J. Vis.
11, 1–12.