Appears in Transactions on Graphics, vol. 28, no. 3, 2010 (SIGGRAPH 2010)

Effects of Global Illumination Approximations on Material Appearance

Jaroslav Křivánek*
Cornell University
Charles University in Prague

James A. Ferwerda†
Rochester Institute of Technology

Kavita Bala‡
Cornell University

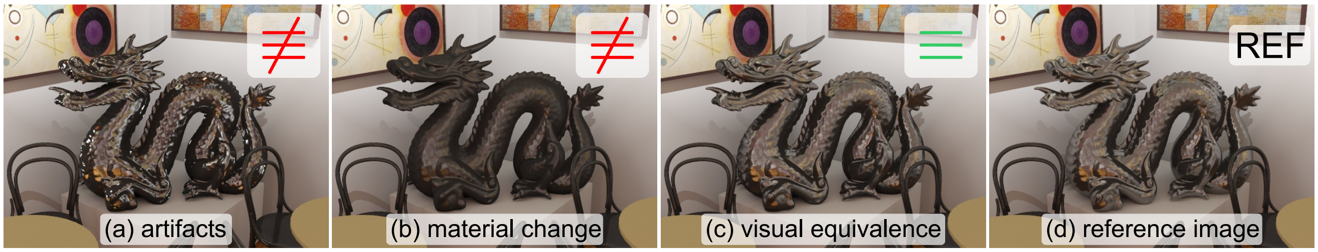

Figure 1: *Examples of inequivalent and equivalent VPL rendering. (a)-(b) are VPL renderings with 1k VPLs, and clamp levels C1 = 316 and C8 = 0.1, respectively, that are* not equivalent (≢) *to the reference (d) because they have image artifacts (a) or different perceived material appearance (b). (c) VPL rendering produces an image that is* visually equivalent (≡) *to the reference for 100k VPLs and clamp level C4 = 10, even though some reflections are lost where the Dragon is in contact with the pedestal and around its silhouette.*

Abstract

Rendering applications in design, manufacturing, ecommerce and other fields are used to simulate the appearance of objects and scenes. Fidelity with respect to appearance is often critical, and calculating global illumination (GI) is an important contributor to image fidelity; but it is expensive to compute. GI approximation methods, such as virtual point light (VPL) algorithms, are efficient, but they can induce image artifacts and distortions of object appearance. In this paper we systematically study the perceptual effects on image quality and material appearance of global illumination approximations made by VPL algorithms. In a series of psychophysical experiments we investigate the relationships between rendering parameters, object properties and image fidelity in a VPL renderer. Using the results of these experiments we analyze how VPL counts and energy clamping levels affect the visibility of image artifacts and distortions of material appearance, and show how object geometry and material properties modulate these effects. We find the ranges of these parameters that produce VPL renderings that are visually equivalent to reference renderings. Further we identify classes of shapes and materials that cannot be accurately rendered using VPL methods with limited resources. Using these findings we propose simple heuristics to guide visually equivalent and efficient rendering, and present a method for correcting energy losses in VPL renderings. This work provides a strong perceptual foundation for a popular and efficient class of GI algorithms.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture;

Keywords: Perception, material perception, visual equivalence, instant radiosity, virtual point light, VPL, global illumination.

*jaroslav@graphics.cornell.edu, † jaf@cis.rit.edu, ‡ kb@cs.cornell.edu

1 Introduction

Applications like industrial design, manufacturing, and ecommerce use computer graphics rendering to design and visualize their products. It is critical for these applications that the rendered images accurately convey the object's appearance, including its material and shape. Distortion of perceived appearance can result in design and manufacturing errors, commercial mistakes, and unacceptable economic costs in such applications.

Global illumination (GI) rendering algorithms simulate global light transport with varying degrees of accuracy depending on the approximations they make. Monte Carlo and Metropolis renderers [Dutré et al. 2006] are accurate and ubiquitous, but extremely slow. Faster GI algorithms, such as instant radiosity based methods [Keller 1997], radiosity, and precomputed radiance transfer [Sloan et al. 2002], often limit their support of materials (e.g., restricting them to near-diffuse) to gain performance. Further, approximations made for performance considerations can cause image artifacts such as splotchiness (e.g., when interpolating cached shading), bright spots (e.g., in instant radiosity), blurred reflections (e.g., when using low order spherical harmonics), and noise (e.g., when using too few samples). These algorithms employ various approaches to suppress artifacts, for example artificially restricting scenes to be diffuse or low gloss only; unrealistically raising the diffuse albedo of materials to decrease the gloss contrast; strategically lighting the scene to wash out artifacts; and eliminating classes of light paths like caustics, or glossy inter-reflections. But these adhoc choices are not appropriate for applications that must accurately simulate the appearance of real-world materials and illumination.

In this paper we systematically study the impact on image fidelity of approximations made by a class of rendering algorithms which have recently been gaining a lot of popularity: instant radiosity based approaches [Keller 1997; Walter et al. 2005; Walter et al. 2006; Laine et al. 2007; Hašan et al. 2007; Ritschel et al. 2008; Hašan et al. 2009]. These algorithms simulate global illumination by tracing light particles into the scene and creating *virtual point lights* (VPLs) on the surfaces intersected; these VPLs then illuminate the scene. The number of VPLs is a user-controlled parameter that lets the user make a performance-quality tradeoff. More VPLs are slower to render, but fewer VPLs cause image artifacts (see Figure 1). Another user-controlled parameter (*clamping*) is used to eliminate these artifacts by clamping the corresponding light energy. But clamping eliminates short-range and glossy reflections, and causes overall image darkening. This in turn can negatively impact the fidelity of material appearance.

We chose to study VPL rendering algorithms for the following reasons: (1) they are popular, because they are amenable to both GPU and CPU implementation and therefore, can be used in interactive applications; (2) a broad range of algorithms that build on VPL rendering have been developed; (3) they provide a continuous method of control to improve quality by increasing the VPL count and to trade material appearance fidelity for artifacts by adjusting the clamping parameter.

We investigate how VPL count and clamping affect image quality and object appearance. We ask two questions:

(1) When are images rendered by VPL algorithms artifact-free?

(2) When is the material appearance of objects in VPL renderings the same as in reference renderings?

The VPL rendering is *visually equivalent* to a reference if it is both artifact-free and preserves material appearance. Figure 1 shows examples of equivalence and inequivalence.

We ran psychophysical experiments to investigate equivalence for a range of geometries, materials, illumination conditions, VPL counts, and clamping values. Our studies indicate that:

(1) VPL count has a big impact on equivalence; more VPLs increase equivalence. For glossy objects, very large VPL counts are needed for any equivalence at all.

(2) Geometrically complex objects are more forgiving to the VPL algorithm. Simpler glossy objects need more VPLs for equivalence, and so counter-intuitively, are more expensive to render artifact-free, and with accurate material appearance.

(3) Metals are often too challenging for VPL algorithms. Dielectrics, and more-diffuse materials are more forgiving than smooth metals. Both metals and dielectrics demonstrate increasing equivalence with greater roughness.

(4) Our findings on whether accurate direct illumination can mitigate the distortions to material appearance from indirect illumination approximations are inconclusive.

We validated our results by confirming our findings on geometries, materials, and VPL settings that did not appear in our main study. The results indicate that the trends we have found are robust across a range of real-world object geometries and material properties. We also apply our findings to scene design by proposing simple heuristics to guide the selection of user-controlled rendering parameters in VPL algorithms and demonstrate improved efficiency in rendering. Further, we develop a simple energy normalization method that mitigates distortions of material appearance in VPL renderings.

Our work is the first that leverages the visual equivalence concept to systematically study the visual fidelity of a practical rendering application. This work provides a strong perceptual foundation for VPL methods, a popular and efficient class of global illumination algorithms. It draws attention to the limitations of VPL methods in their ability to correctly simulate lighting with limited resources. As such, it can have an impact on the use of VPL algorithms in practical computer-aided appearance design applications and on the development of new VPL-based rendering systems.

2 Related Work

We review related work in VPL rendering, perceptually-based rendering, material appearance, and visual equivalence.

VPL rendering. The original VPL rendering algorithm, *instant radiosity* [Keller 1997], was one of the first GPU-accelerated global illumination (GI) methods. Since then, VPL rendering has been the basis of a number of interactive GI algorithms [Wald et al. 2002; Segovia et al. 2007; Laine et al. 2007]. Ritschel et al. [2008] and Dong et al. [2009] accelerate VPL rendering by making approximations to visibility, and validate this approach with a perceptual study [Yu et al. 2009]. These interactive VPL algorithms typically use a relatively low number of VPLs, up to several thousand. Matrix row-column sampling [Hašan et al. 2007] improves image quality by selecting the VPL set from an initial pool of around hundred thousand lights. Lightcuts [Walter et al. 2005; Walter et al. 2006] focuses on high-fidelity, though at the price of interactivity; this scalable algorithm renders images using millions of VPLs. *Clamping* is used by all VPL algorithms to suppress artifacts. Kollig and Keller [2004] compensate clamping energy loss using path tracing. Hašan et al. [2009] introduce a new type of virtual light for which no clamping is required.

Perceptually-based rendering. The goal of perceptually-based rendering is to increase rendering efficiency while preserving image fidelity by taking advantage of the limits of human vision. One approach has been to use models of visual contrast coding such as the Visible Differences Predictor [Daly 1993] to guide the rendering process [Myszkowski 2002; Bartz et al. 2008]. Higher level characteristics of visual processing such as attention, salience, change blindness, sensitivity to natural image statistics, and task dependence of perception have also been exploited [O'Sullivan et al. 2004; Bartz et al. 2008]. The perception of illumination and its role in high-fidelity graphics has also been investigated. Stokes et al. [2004] and Debattista et al. [2005] have developed GI rendering algorithms based on the perception of illumination components.

Material appearance. The focus of this paper is on visually accurate rendering of material appearance, and this has also been an active area of research within the field of perceptually-based rendering. Pellacini et al. [2000] and Westlund and Meyer [2001] developed psychophysical models of gloss perception. Fleming et al. [2003; 2005] established that natural illumination statistics play an important role in the faithful rendering of material appearance. Khan et al. [2006] have leveraged these findings to develop tools for interactive material editing. Vangorp et al. [2007; 2008] show that object shape can affect gloss perception, and that objects with moderate surface undulations provide more stable perceptions of gloss than simple objects like spheres.

Visual equivalence. A line of work particularly relevant to ours is Ramanarayanan et al.'s [2007; 2008] research on visual equivalence. They conducted a series of studies that investigate how blurring and warping transformations on illumination maps affect the appearance of rendered objects. They found that for several transformations, objects rendered with transformed maps appeared the same as objects rendered with reference maps despite the fact that the images were visibly different. They termed these images: *visually equivalent*. They also found that visual equivalence was modulated by the geometry and material properties of the object.

The concept of visual equivalence provides a strong foundation for our work. VPL rendering algorithms approximate the illumination in a scene as a set of point lights. This can be seen as a complex transformation on the actual illumination field. Thus, similar to Ramanarayan et al., we are seeking to determine when images rendered with VPL illumination approximations are visually equivalent to reference renderings, and to understand how both rendering parameters and object properties affect this relationship.

3 Problem Statement

Our goal is to find when VPL rendering can produce artifact-free images that accurately represent object appearance. We first explain why artifacts arise in VPL rendering. Then we describe how VPL rendering parameters affect both image artifacts and appearance. Finally we list the trends we want to investigate by systematically studying VPL rendering.

3.1 Review of VPL Rendering

VPL algorithms simulate global illumination in two passes. In the first pass, particles are traced from the light sources, creating *virtual point lights* (VPLs) at the particle-surface interactions. In the second pass the VPLs illuminate the scene. Radiance $L(x)$ at point x is computed as a sum of contributions from all VPLs (or from a carefully selected subset of VPLs [Walter et al. 2005]):

$$L(x) = \sum_{k=1}^{\#\text{VPLs}} \frac{M_k(x)}{r_k^2(x)} \cos_k I_k, \tag{1}$$

where $M_k(x)$ is the material term at x, $r_k(x)$ is the distance from the surface point x to the VPL, $\cos_k$ is the cosine term at the VPL and I_k is the VPL intensity.

The premise of this approach is that all the VPLs together approximate illumination smoothly, therefore, the contribution of any one VPL to a surface point should not be unduly important. However, this assumption breaks when: (1) the VPL is near the surface point, where the inverse-squared-distance term $1/r_k^2(x)$ is large; and (2) on glossy materials, where a large BRDF value may greatly amplify the contributions of some individual VPLs. Large contributions from individual VPLs cause visual artifacts in the form of bright spots, as shown in the left column of Figure 2. The artifacts become more apparent as the number of VPLs decreases.

All VPL algorithms suppress these artifacts by imposing an upper bound on the VPL contribution, a technique known as *clamping*. Similar to [Walter et al. 2006], we clamp the product of the material term and the inverse-squared-distance: $L(x) = \sum_{k=1}^{\#\text{VPLs}} \min\{c, M_k(x)/r_k^2(x)\} \cos_k I_k$. (Notice the inverse dependence between the value of the clamping constant c and the impact of clamping: a smaller value of c corresponds to a higher clamping level.) Figure 2 shows how clamping eliminates artifacts. Clamping suppresses artifacts at the cost of energy loss, which is most pronounced for short-range reflections and sharp glossy reflections—gloss is affected to a much larger extent than diffuse color, as seen in the bottom row.

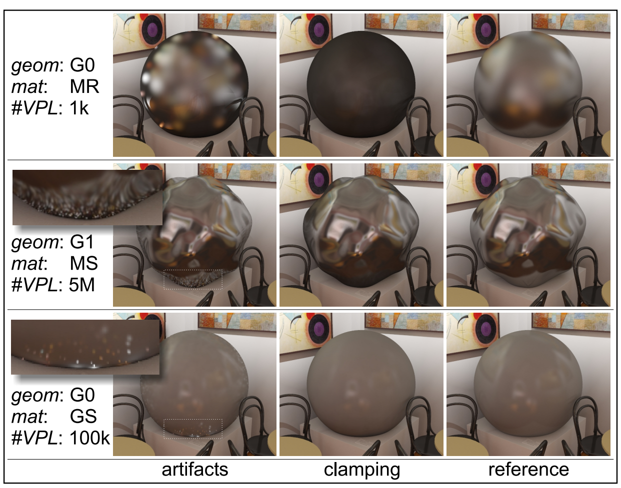

Figure 2: *Examples of artifacts in VPL renderings (left column) and their suppression by clamping (middle column).* Top row*: Clamping suppresses artifacts due to low VPL count at the cost of removing most of the indirect illumination.* Middle row*: Clamping eliminates the short-range and glossy reflections of the pedestal.* Bottom row*: Much of the reflected environment is missing from the clamped image, while the overall diffuse reflectance is mostly preserved. Identifiers for geometries and materials are defined in Section 4.1.*

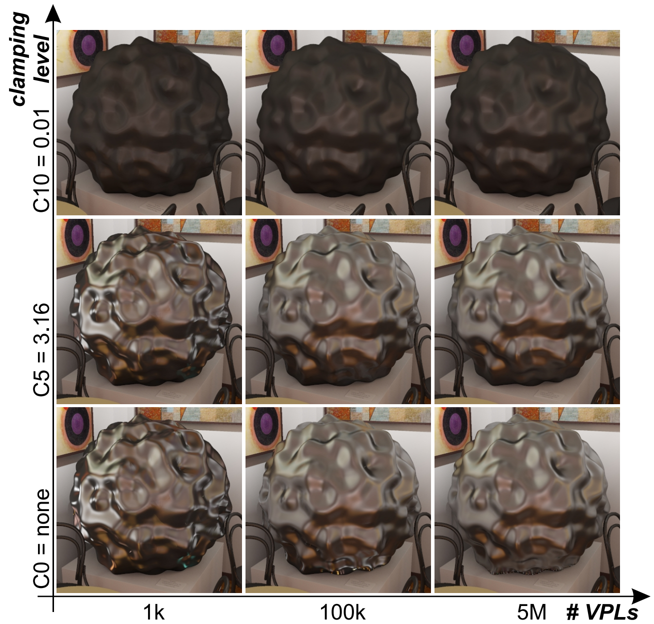

Figure 3: *Clamping level vs. VPL count for geometry G2, material MR (rough metal).* Bottom row*: No clamping, see visible artifacts at the bottom of the object.* Middle row*: Mostly artifact-free, at the expense of losing reflections of the pedestal and some overall darkening of the object. The 1k VPL rendering still has visible artifacts.* Top row*: Extreme clamping suppresses artifacts, but the object appears very dark. Our study identifies the 5M VPL/no clamp image (lower right) as equivalent to the reference.*

To summarize, VPL-based rendering algorithms have two important user-controlled parameters: the number of VPLs, and the clamping level. The setting of these parameters can have three different visual effects: (1) Artifacts arise when VPL count is low. Clamping results in energy loss that exhibits itself as (2) the loss of short-range and glossy reflections, and (3) overall object darkening. Our goal is to study the perceptual impact of these three effects with respect to different geometry, material, and illumination.

3.2 Visual Equivalence for VPL Rendering

We study how VPL approximations impact image quality and material appearance for a range of geometries, materials and illuminations, and for various VPL counts and clamping levels. Figure 3 shows the space of clamping levels vs. VPL counts for one choice of geometry, material, and illumination. Figure 4 sketches a graph of behavior we might expect for a given geometry, material, and illumination. Note that the true slope of these curves could differ.

Artifact-free images. As VPL counts increase the level of clamping necessary to produce artifact-free images decreases. The lower curve in Figure 4 defines an *artifact visibility threshold*. Above the curve are artifact-free images, and below are images with artifacts.

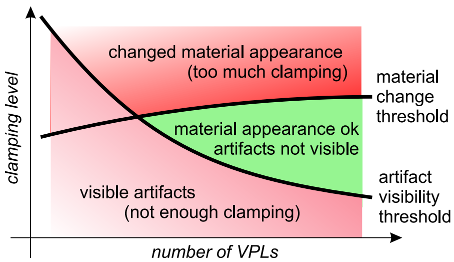

Figure 4: *Visual equivalence (green region) in the space spanned by the VPL count (x-axis) and clamping parameters (y-axis, no clamping at the bottom).*

Accurate material appearance. Similarly, the upper curve, the *material change threshold*, delineates the part of the space where material appearance is accurate (below the curve), from where it is changed (above the curve, with aggressive clamping).

Visual equivalence. Between these two curves is a region (shown in green) where:

(1) images are artifact-free; and,

(2) material appearance is the same as in the reference.

Images in this region are *visually equivalent* to the reference renderings. Our goal is to identify these green regions of visual equivalence. To do so we systematically study different combinations of geometry, material and illumination, find the artifact and material curves, and use them to find the equivalent region.

In considering equivalence in VPL rendering, we arrived at a set of questions to study:

(1) What is the performance vs. fidelity tradeoff imposed by VPL count selection? Are there materials or geometries for which it is possible to get visually equivalent images with the small VPL counts needed for interactive applications? How high should the VPL count be for "difficult" materials and geometries?

(2) How does shape complexity affect equivalence? As objects get more complex, are their images more, or less equivalent?

(3) How does material affect equivalence? Which materials are more "difficult" to render using VPL-based algorithms?

(4) Accurate direct illumination could mitigate the perceptual effects of indirect illumination approximations made by VPL rendering. Is this the case?

4 Experiments

To investigate how object and rendering parameters affect object appearance, we ran two psychophysical experiments. One measured the artifact visibility threshold and the other measured the material change threshold. The goals of these experiments were to identify combinations of VPL counts and clamping levels that produce images that are visually equivalent to reference images, and to investigate how these combinations are affected by geometry, material and illumination.

4.1 Stimuli

To conduct the experiments we first created a scene model and generated a set of stimulus images (see Figure 5). The scene was of an art gallery café with tables and chairs surrounding a sculpture (the object we studied) on a low pedestal, and various objects at a range of distances from the sculpture. Lighting included both spot and area fixtures (see supplementary material). This particular scene was created so that (1) the abstract objects we chose to study

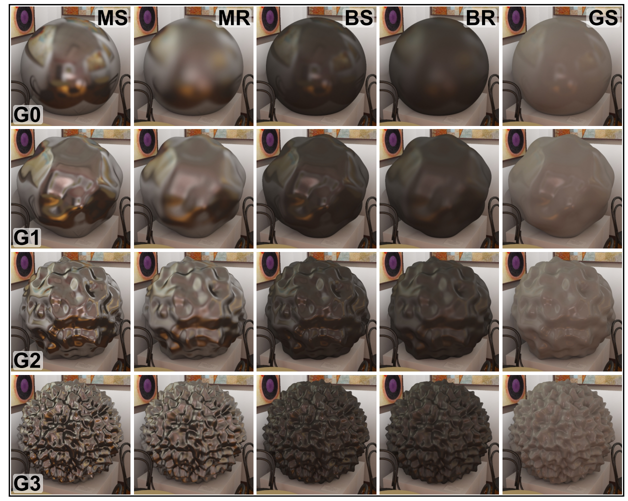

Figure 5: *Reference renderings of the stimulus objects with indirect-only illumination.*

could be seen in a plausibly realistic context; (2) the reflections in the object surfaces came from light sources and other objects at a range of different distances; and (3) both indirect-only and direct-and-indirect illumination could be used.

Geometry. The test objects were those used by [Ramanarayanan et al. 2007]. Object G0 was a sphere; objects G1-G3 were spherical blobs of varying "bumpiness" (see Figure 5). We chose these geometries because they have been used in prior work, reflect a large portion of the surrounding scene, and are approximately equally spaced perceptually [Ramanarayanan et al. 2007], thus allowing systematic investigation of surface complexity.

Materials. Materials were represented by the Ward-Dür BRDF model [Dür 2005], defined by three parameters: diffuse reflectance ρ_d, specular reflectance ρ_s, and surface roughness α. We chose this model, rather than the original Ward model [Ward 1992], because of its better fidelity in modeling real-world materials [Ngan et al. 2005]. Columns of Figure 5 show the materials used in the experiments. The material set consisted of a pair of smooth and rough achromatic metals MS (ρ_d, ρ_s, α) = (0.03, 0.22, 0.05), and MR (0.03, 0.22, 0.15) and a set of three dielectrics: smooth and rough blacks BS (0.03, 0.033, 0.05), BR (0.03, 0.033, 0.15) and a smooth gray GS (0.19, 0.033, 0.05). We choose this set of materials because (1) they provided rigorous test cases for the rendering algorithms (all else being equal, surfaces with low ρ_d will have the highest contrast reflections); (2) metals ($\rho_s = 0.22$) and dielectrics ($\rho_s = 0.033$) are important material classes to study; (3) by varying roughness, and diffuse and specular reflectance, we can systematically study the effects of these properties.

Illumination. To study how illumination affects equivalence, we rendered scenes where the objects had both indirect-only and direct-and-indirect illumination. To achieve this the scene model included ceiling mounted track lights (spot lights) and a central area fixture. For the indirect-only renderings track-lights, that all faced away from the test object, illuminated the scene. For the direct-and-indirect renderings we added an area light that directly illuminated the object and created pronounced highlights on its surface. See supplementary material for details.

VPL counts and clamping levels. To study the effects of VPL counts on object appearance we created three image sets rendered with 1,000 (1k), 100,000 (100k), and 5,000,000 (5M) VPLs. We chose these values because they span the range of current VPL rendering applications: 1k VPLs are common in interactive applications [Ritschel et al. 2008], 100k VPLs correspond to fast global illumination previewing [Hašan et al. 2007], while 5M VPLs are currently used only in scalable VPL rendering methods aimed at high-fidelity rendering [Walter et al. 2005]. To study the effects of clamping we applied 11 different clamping levels to each of the VPL settings: no clamping (C0), and ten half log unit steps from least clamping (C1 = 316) to most clamping (C10 = 0.01).

Rendering and display. The images were rendered at the resolution of 560 × 420 using a custom-built physically-based renderer. VPL renderings of the test objects were produced using the Lightcuts algorithm [Walter et al. 2005] with the error ratio set to 0.75% so that the images were indistinguishable from brute-force rendering using all VPLs. Therefore, our study and its results are not affected by the use of Lightcuts in any way and apply to all VPL-based rendering algorithms that utilize clamping. Path tracing was used to render the backgrounds of the test images as well as the entire reference image set. Therefore, any artifacts or material changes due to VPL rendering could only appear on the test object, but not in the background. Two full image sets taken from slightly different camera positions (15 deg. rotation around a vertical axis) were rendered. The resulting HDR images were all tone-mapped for display using Reinhard et al.'s photographic tone mapping operator [2002] (see supplementary material).

4.2 Procedure

Using the stimuli defined above we ran two psychophysical experiments. The first experiment measured thresholds for seeing artifacts. The second experiment measured thresholds for perceiving material changes induced by clamping. Both experiments used a fully randomized two-alternative-forced-choice (2AFC) design, and were delivered using the interface shown in Figure 6 on high-quality LCD monitors under standard office lighting conditions. The monitors were set to their factory sRGB settings but no other calibrations were performed. At a standard 18-inch viewing distance each image in the interface subtended approximately 16 × 12 degrees of visual angle. Twelve subjects participated in the artifacts experiment and 14 participated in the material experiment. All were university students or employees, some with knowledge of imaging and graphics, but none were familiar with the specific goals of the experiments. The subject pool was roughly split between males and females and ages ranged from 20 to 65. All subjects had normal acuity and color vision. No time constraints were imposed but each experiment took about 20 minutes to complete.

Artifact experiment. The goal of this experiment was to find thresholds for seeing artifacts in the VPL renderings as a function of clamping level and object properties. Subjects were shown pairs of images from the stimulus set (the upper image in Figure 6 was not

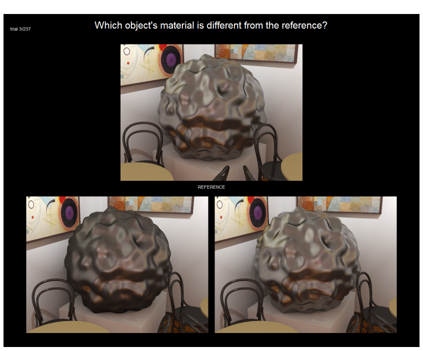

Figure 6: *Interface used in the Material experiment. In the Artifact experiment only the lower pair of images was presented.*

shown) and were asked to "Select the image that has the artifacts." One image was a reference rendering and the other was rendered with a particular combination of VPLs and clamping. The objects in the images had the same geometry and illumination but different materials to avoid biases from correlations between clamping level and material lightness. To make the task object-focused rather than image-focused, the reference and test images always had different camera viewpoints. Prior to testing, subjects were familiarized with the range of objects and artifacts they would be seeing as described in the supplementary material.

In the design phase of the experiment it was determined that examining the full test set (1,320 images) would take too much time and waste effort. Since the goal was to find artifact *thresholds*, for each object and VPL count, four of the eleven clamping levels spanning each supposed artifact threshold were selected for testing. This resulted in a final test set of 480 images (4 geometries × 5 materials × 2 illuminations × 3 VPL counts × 4 clamping levels).

We considered using a computational metric such as a visible difference predictor [Daly 1993] to predict the visibility of artifacts. However, such a measure is too restrictive and would not allow leveraging of the equivalence concept. Thus, a psychophysical experiment was necessary to map the space and to take advantage of the additional latitude provided.

Material experiment. The goal of this experiment was to determine when excessive clamping or an insufficient number of VPLs change the perceived material properties in VPL-rendered images. On each trial subjects were shown three images (see Figure 6), an upper reference image, and a lower test pair. Subjects were asked "Which object's material is different from the reference?" The objects in the reference image and test pair had the same shape, material, and illumination, but different camera viewpoints. One of the test images was a particular VPL count/clamping level combination from the test image set, the other was a reference rendering (but using the alternate viewpoint). As before, this approach was used to make the task material-focused rather than image-focused. To limit the size of the experiment, the artifact thresholds were used to prune the full 1,320 test image set so that only images subjects judged as artifact-free were included for testing. Details on the pruning procedure are given in the supplementary material.

4.3 Data Analysis

To establish the artifact visibility and material change thresholds, the individual subject responses were combined and averaged to calculate the proportion of subjects who saw a particular VPL count/clamping level combination for a particular object as either containing artifacts or showing a change in material appearance. Because we sampled around the thresholds we could have taken these proportions and fit them with psychometric functions and reported seemingly precise threshold values for artifacts and material change. However we felt that this modeling exercise would be going beyond the precision in the data and possibly lead to overfitting and misleading predictions. Therefore instead, we made simple cuts through the response data that correspond to the conventional 75% two-alternative-forced-choice threshold value (50% being chance performance, i.e., pure guessing). Due to quantization based on the number of subjects in the experiments (12 and 14 respectively in the artifact and material studies) the artifact and material thresholds correspond to 75% and 78.6% respectively. Based on this analysis the combined results of the experiments are shown in Figure 7 (a). Due to the clamping level subsampling used in the material experiment, threshold uncertainty is on the order of one clamping level. The green "visual equivalence" regions are conservative with respect to this uncertainty.

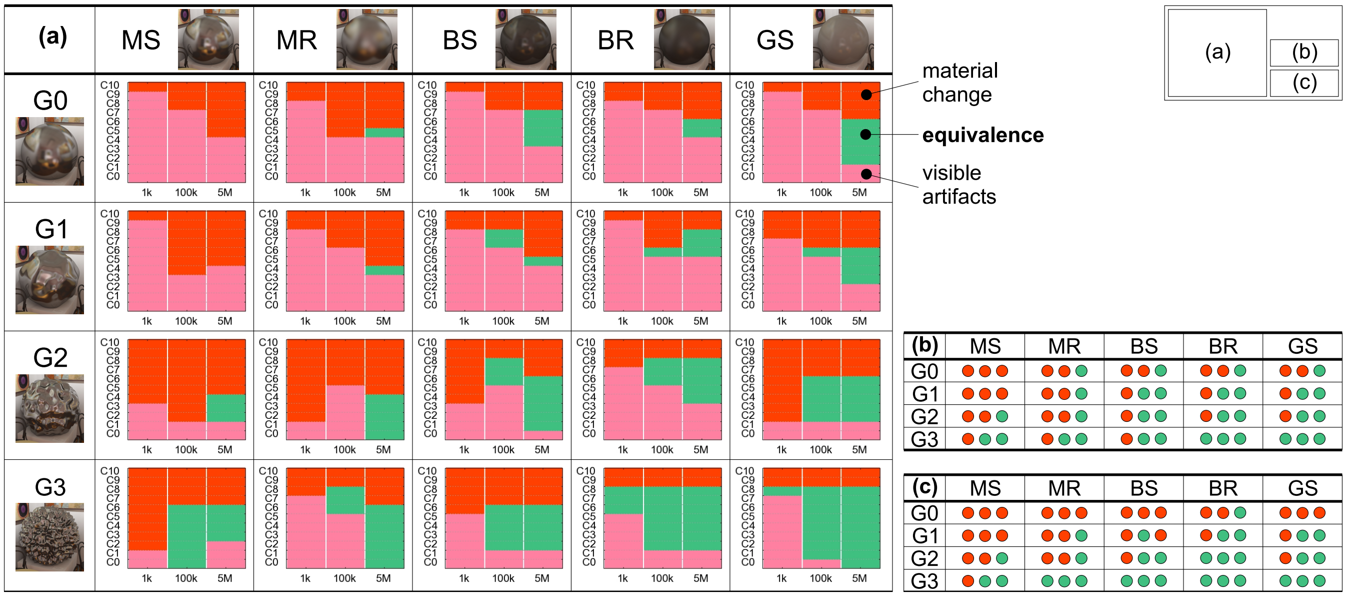

Figure 7: *(a) Results of the artifact and material experiments with* indirect-only *illumination (the corresponding tables for* direct-and-indirect *illumination are given in the supplementary material). Rows correspond to shapes G0–G3 (top to bottom). Columns correspond to different materials. Each graph represents the results for one object (geometry-material combination). Within each graph, each column shows the results for one VPL count. The lower pink region in each column indicates "bad" combinations where the objects were seen to have artifacts, the upper red region indicates "bad" combinations where the objects' material looked different from the reference, and the green region indicates "good" combinations where the VPL rendered objects were seen as visually equivalent to the reference renderings. In interpreting each graph, keep in mind the idealized behavior from Figure 4. (b) Equivalence-exists (EE) table for indirect-only illumination. (c) Equivalence-exists (EE) table for direct-and-indirect illumination.*

5 Discussion of Results

We now discuss our results and identify trends in equivalence based on the number of VPLs, geometry, material, and illumination.

5.1 Effect of VPL Count on Equivalence

First, we study how equivalence changes with the VPL count. To simplify exposition, we introduce a measure for each VPL count in each graph to characterize if *equivalence exists* (EE). The EE value is a boolean that digests the information for all clamping levels corresponding to a single VPL count. If there exists *some* clamp value for which the equivalence data is green, EE for that VPL count is set to green; else, it is set to red. Note that the EE measure ignores the *number* of equivalent clamp values. Figure 7 (b) shows the EE table corresponding to Figure 7 (a). Now we can easily study trends in equivalence with increasing VPL counts. For each triple, going from the smallest VPL count (first element) to the largest (third element), we see monotonic transitions from red (if it exists) to green (if it exists). Thus, for each geometry and material, *as the VPL counts increase, equivalence increases*.

Further we note that the VPL counts needed for equivalence for our glossy shapes are often quite large. For the smooth metal (MS), our most demanding material, there is no equivalence for G0 and G1 even with 5M VPLs. For all objects, 1k VPLs is insufficient for equivalence, except for G3 for the two materials with least visible gloss (the black rough BR, and gray smooth GS dielectrics).

5.2 Effect of Shape on Equivalence

To understand trends in shape complexity, consider each VPL count across all geometries (a column) in the tables in Figure 7. Moving down towards increasingly complex shapes (G0 to G3), equivalence increases. This indicates that *geometrically complex glossy shapes are more forgiving of illumination errors than simple ones*.[1] The practical implication of this finding is that rendering simpler glossy objects (for example, near-planar surfaces like tables, chairs, walls), requires *larger* VPL counts for equivalence than when rendering complex shapes. This finding is counter-intuitive from a rendering point of view, because simpler objects are typically cheaper to render. But in VPL rendering, to get equivalent images simpler glossy objects must be rendered with *higher* VPL counts; thus, at a greater rendering cost than complex glossy objects.

Ramanarayanan et al. [2007] observed a similar effect of shape complexity on equivalence for globally consistent illumination transformations. However, since clamping affects mostly local inter-reflections, it was unclear at the outset if the same effect would hold. Our studies show that equivalence is in fact increased for complex geometries even for this novel type of illumination transformation.

5.3 Effect of Material on Equivalence

We consider the two classes of materials, metals and dielectrics, separately to understand equivalence with material variation.

Metals. For our most unforgiving material, the smooth metal (MS), no equivalence exists for the simpler geometries (G0, G1), but the most complex geometry (G3) exhibits equivalence. Moving to the more forgiving rough metal (MR), G0 and G1 can now be rendered equivalently, but only at the highest VPL count, 5M.

[1]Note that this effect is not to be confused with contrast masking due to surface complexity (of which there might be some), but is more likely due to our inability to comprehend the causes of illumination features in high complexity objects.

Dielectrics. From the metal (MS), we step along specular reflectance (ρ_s) to BS (the smooth, black dielectric). Overall, there is an increase in equivalence. G3 has similar equivalence, with a slight drop for 100k VPLs (see below for an explanation). Trends within the dielectrics are similar to the metals: starting at BS (the smooth, black dielectric), a step in α to BR (the rough, black dielectric) confirms the trend that rougher materials have more equivalence. In fact this is the rare example where, for G3, 1k VPLs suffice.

Finally, we consider changing diffuse reflectance. Starting at BS (smooth, black dielectric) we take a step in diffuse reflectance ρ_d to the gray dielectric, GS. This material shows greater equivalence because the gray material has lower contrast reflections. Thus, the inaccuracies introduced by clamping or other VPL approximations are less perceptually salient. This result agrees with the findings of Pellacini et al. [2000] regarding the effects of surface albedo on perceived gloss and gloss discrimination.

Summary. The metals are generally unforgiving with not much equivalence except at very high VPL counts and for the most geometrically complex object. There is a small increase in equivalence when going from the smooth to rough metal. The dielectrics behave similarly to metals (rougher is more equivalent), but are generally more forgiving than the metals. Furthermore, lighter dielectrics (higher diffuse albedos) have greater equivalence.

5.4 Effect of Illumination on Equivalence

Unlike indirect illumination, VPL techniques do not in any way approximate direct illumination which introduces accurate surface highlights that could provide cues for material perception. Can accurate direct illumination mitigate the effects of clamping on indirect illumination?

Our study included a full set of images rendered with direct-and-indirect illumination generating a result table similar to Figure 7 (a) (included in the supplementary material). We present the corresponding equivalence-exists (EE) table in Figure 7 (c), which shows the trends in equivalence. Quite surprisingly, we found that direct-and-indirect illumination behaved qualitatively similar to indirect-only illumination with some small changes in the G0 row (the differences are close to threshold), the G1 row (increase of equivalence for G1/BS is not monotonic with VPL count), and the G3 row (more equivalence with direct). The lack of a significant effect runs counter to expectations a rendering practitioner may have. Further exploration is needed to understand the effect of direct illumination on equivalence.

5.5 Summary

Geometry, material, illumination, and VPL counts all interact with each other in affecting equivalence. Our findings are:

(1) VPL count has a big impact on equivalence. For glossy objects, a high number of VPLs (in millions) are needed for artifact-free images that accurately render material appearance. 1k VPLs rarely produce equivalent images.

(2) Geometrically complex objects are more forgiving to the VPL algorithm, while geometrically simple objects require very large VPL counts to achieve equivalence. Thus, counter-intuitively, simpler glossy objects are more expensive to render artifact-free, and with accurate material appearance.

(3) Metals are unforgiving, and cannot necessarily be rendered accurately using VPLs. Dielectrics and more diffuse materials are all more forgiving than the smooth metal and have greater equivalence. Both metals and dielectrics demonstrate increasing equivalence with greater roughness.

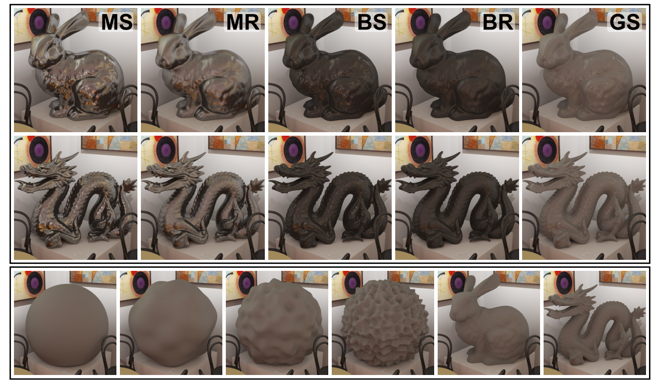

Figure 8: *Reference renderings of the images used in the validation study (cropped).* Top two rows*: Stanford Bunny and Dragon with the five materials from the main study.* Bottom row*: All six objects (G0–G3, Bunny, and Dragon) rendered with a purely Lambertian gray material.*

(4) Our findings on whether accurate direct illumination can mitigate the distortions to material appearance from indirect illumination approximations are inconclusive. Further systematic exploration is needed.

Our study explores the range of geometry, material, illumination, and VPL rendering parameters for which equivalent images can be achieved. Our study also points out the limitations of VPL methods in their ability to render equivalent images with limited resources. Performance considerations drive the selection of small VPL counts and aggressive clamping, but often sacrifice equivalence.

Outliers. While the above trends in equivalence for VPL count, geometry, material, and illumination broadly exist, there are a few outliers. The size of the subject pool, inherent noise in experiments that involve input from human subjects, the relatively coarse sampling of the clamping levels, are all reasons for outliers. However, in a couple of cases there is a more systematic deviation.

The transition from the smooth metal (MS) to the rough metal (MR) shows a drop in equivalence for 100k. There is also a decrease in equivalence for G3 in the transition from the smooth metal (MS) to the smooth, black dielectric (BS). We hypothesize that the effect may be due to the ambiguity between the artifacts and actual highlights on the surface in the cases when the two have similar shape. The dependence of artifact shape and size on VPL count, and their relation to artifact visibility warrant more exploration in the future. Examples of stimulus images corresponding to the outliers are shown in the supplementary material.

6 Validation

To confirm the trends observed in the main experiments, we ran a set of follow-up validation studies, where we generalized geometry, VPL count, and material (see Figure 8).

6.1 Stimuli and Procedure

Geometry. To generalize our geometries we rendered the Stanford Bunny and Dragon for all five materials from the main experiments. We tested images of each of these objects rendered with 100k and 5M VPLs, but excluded 1k VPLs since it offered no equivalence in the main experiment.

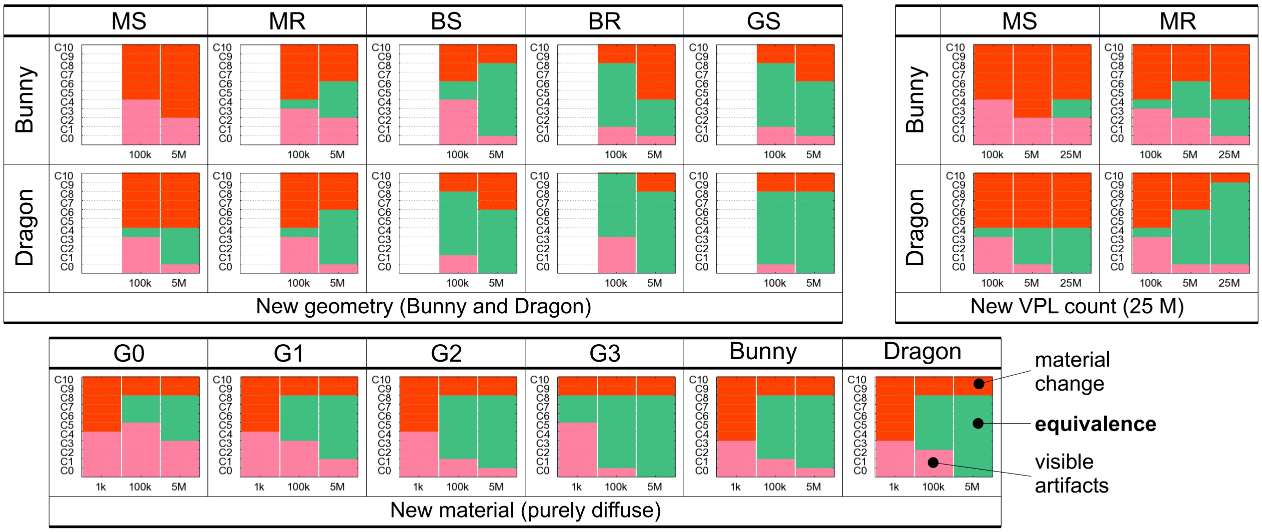

Figure 9: *Validation results for new geometry (Bunny and Dragon), new VPL count (25M), and new material (diffuse).*

VPL count. To confirm the trend that the range of equivalence across clamping levels increases with increasing numbers of VPLs, we tested metal versions (MS and MR) of the Bunny and Dragon objects rendered using 25,000,000 (25M) VPLs.

Material. To investigate whether low VPL counts ever yield equivalence we rendered all six objects (G0–G3, Bunny, and Dragon) with a purely Lambertian gray material ($\rho_d = 0.19$, $\rho_s = 0$), for each of the three VPL counts used in the main experiment.

The indirect-only illumination was used for the validation studies. We used the same experimental design and procedure as in the main experiments (with a slight difference in how images were selected for the material experiment. See the supplementary material.).

6.2 Validation Results

Figure 9 shows the validation result tables.

Geometry. The graphs for new geometry confirm the trends observed in the main experiments: equivalence increases with the number of VPLs; with increasing shape complexity; and along the contrast gloss and roughness material axes. The only exception is the transition from the black smooth (BS) to black rough (BR) material for the Bunny. A possible reason for this anomaly may be the coarse geometry of the Bunny. It is interesting to note that with some exceptions the results for the Bunny geometry fall roughly between the results for the G1 and G2 objects and the results for the Dragon geometry fall between those for the G2 and G3 objects. This complements the findings of Ramanarayanan et al. [2007] that the "bumpiness" of the Bunny and Dragon are G1.5 and G2.5 respectively.

VPL count. Results for 25M VPLs confirm the trend of increasing equivalence with the number of VPLs. We now get equivalence for the smooth metal (MS) Bunny where none exists for 100k and 5M VPLs. This shows that some challenging geometry/material combinations *can* be rendered with equivalence using VPL methods.

Material. Results for the purely diffuse material confirm the trends of increasing equivalence with the number of VPLs and with shape complexity. To validate the trend along the material axis, we compared this diffuse material with the results for the most "forgiving" material from the main experiments (the gray smooth dielectric, GS) which has the same diffuse reflectance. Equivalence for the diffuse material is significantly greater in many cases for both 100k and 5M VPLs. 100k VPLs are now sufficient to obtain equivalence for G0. Quite surprisingly, *1k VPLs are still unable to produce equivalent images for most shapes* (the only exception being G3). Looking at the images shows that the most likely reason is that some visible shadow discontinuities caused by individual VPLs are perceived as artifacts. The Dragon with 100k VPLs is the only exception to the trend of increasing equivalence in the transition from GS to the purely diffuse material (see end of Section 5).

Summary. In summary we have validated the findings of our main experiments with respect to the effects of object geometry, VPL count, and material properties on visual equivalence. This indicates that the experiments are probing a robust phenomenon that should serve well as a basis for efficient and effective perceptually-based rendering algorithms.

7 Applications

This section presents two practical applications of our findings: a set of heuristics to guide efficient and equivalent VPL rendering, and a luminance normalization technique for improving equivalence of VPL-rendered objects.

7.1 Heuristics for Efficient Equivalent Rendering

VPL algorithms typically provide rendering solutions an order of magnitude faster than Monte Carlo methods—a difference that can impact the feasibility of a project. However, there are limitations to their applicability. Algorithms for fast previewing that use low number of VPLs are limited to rough, near-diffuse materials. Modern industrial design and architectural applications, however, extensively use high-gloss materials such as plastics and finished metals. We show that millions of VPLs are required to render such materials faithfully. Therefore, scalable VPL rendering methods with sub-linear cost are a necessity. While there is some research in this area [Walter et al. 2005], more exploration is needed.

Even with millions of VPLs, there is still a limit on the range of materials that VPL methods can render. Our main finding is that this range strongly depends on the object shape. Our results can serve as a guideline to practitioners with respect to what VPL methods can achieve, although the exact limits will depend on the scene size and configuration, arrangement of light sources, and viewpoint.

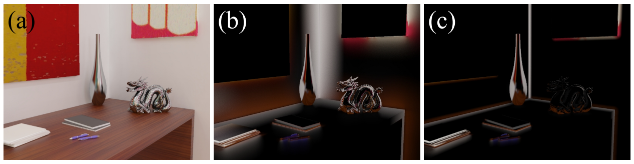

Figure 10: *Optimizing energy compensation by per-object clamping. Image (a) was rendered with Lightcuts using Kollig and Keller's [2004] method to compensate for the clamped energy. Images (b) and (c) show the compensated energy for uniform and per-object clamping, respectively. (b) If a uniform clamping level is selected that is high enough to suppress artifacts in the entire scene, much energy must be recovered by the slow compensation method. (c) By setting the clamping level on a per-object basis, less energy is lost by clamping, and the compensation is sped up by a factor of 2, while still producing an equivalent image.*

We can use our findings to drive efficient VPL rendering while producing equivalent images. One possible approach is in the context of the energy compensation algorithm by Kollig and Keller [2004]. Their method achieves high image fidelity with VPLs by using path tracing to compensate for the clamping energy loss. Unfortunately, this path tracing is often a performance bottleneck. If all objects are rendered with the same clamping value, very conservative clamping settings may have to be used, which delegates most of the rendering effort to the slow path tracing compensation. Our results show that less aggressive clamping suffices to achieve equivalence for geometrically complex and more diffuse objects. Thus, a recommended clamping strategy is to apply the minimum possible *clamping on a per-object basis*. By following this strategy, we were able to speed up the path tracing compensation by a factor of 2 (see Figure 10). Building an automatic model for choosing the clamping constant is left for future work.

7.2 Luminance Normalization

Materials in VPL renderings often appear different from reference renderings due to the energy loss caused by clamping. Based on this observation, we propose and validate a luminance normalization technique to improve equivalence in these cases. The idea is to recover the loss in surface lightness by rescaling the luminance of objects in the VPL renderings.

In practice we render the object with clamping as usual, but also generate an auxiliary image with no clamping (at negligible additional cost). We then rescale the luminance of the object such that its average luminance matches the average luminance of the auxiliary, no-clamp image. Doing so recovers the light energy lost due to clamping and effectively redistributes it over the entire object surface. Note that the no-clamping image can be used for the normalization because VPL rendering with no clamping provides an unbiased estimator of pixel luminance. Since the normalization only needs the average luminance over all object pixels, the estimate has low enough variance for our purposes. Figure 11 shows the effect of normalization for an example object.

We validated the luminance normalization technique for material MR (rough metal) with all the geometries and VPL values used in the main study; Figure 12 shows the results. Normalization introduces equivalence for G1 with 100k VPLs and increases the size of the equivalence region for G2 with 5M VPLs. Thus, it can mitigate some of the consequences of energy loss due to clamping, but only to a limited degree, because the problem with clamping is not just "darkening" (which normalization can remedy) but rather a loss of contrast (which it cannot).

(a) original (b) normalized (c) reference

Figure 11: *Luminance normalization technique. (a) VPL rendering before normalization. (b) After normalization. (c) Reference.*

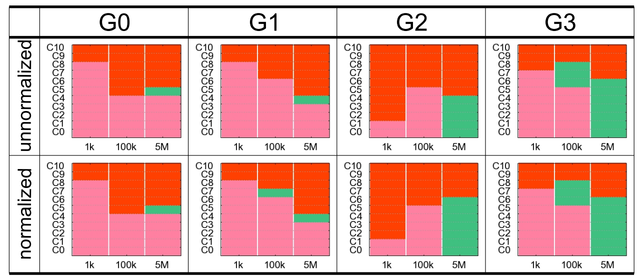

Figure 12: *Results of the luminance normalization validation for material MR (rough metal).* Top row*: Graphs for unnormalized images.* Bottom row*: Graphs for normalized images.*

8 Conclusions and Future Work

In this paper we systematically study the perceptual effects of VPL-based global illumination algorithms on image quality and material appearance. In a series of psychophysical experiments we study how VPL counts and clamping levels affect visual equivalence, and show how equivalence is modulated by object geometry and material properties. We validate our findings on shapes, materials, and illumination that were not in our main study. Further, we propose simple heuristics to guide rendering, and we present a method for correcting energy losses in VPL renderings to increase equivalence.

By explicitly and systematically studying the impact of rendering approximations on appearance this paper takes some initial steps towards providing a strong perceptual foundation for VPL methods, a popular and efficient class of global illumination rendering algorithms. The paper also draws attention to the limitations of VPL methods in their ability to correctly simulate lighting conditions with limited resources.

While we have identified high-level trends, our work has some limitations, and many avenues of future work can be explored. To generalize our results over a wide range of scenes we need perceptual models, but our current data is too coarse to create such models. Cole et al. [2009] demonstrate a promising approach using Amazon Mechanical Turk to collect psychophysical data; it remains to be seen if experiments like ours, that require careful judgment of image quality, can be run robustly in such settings. Further, validation with respect to illumination is important, but it is difficult to control scene lighting to systematically change indirect illumination. Another issue worth studying is the ambiguity between artifacts and highlights. On the application side, we want to use our perceptual models to develop scalable, interactive VPL-based algorithms that can automatically finetune the performance vs. fidelity tradeoff for applications.

Acknowledgements

Many thanks to friends and colleagues who participated in our experiments. This work was supported by the European Commission Marie Curie Fellowship PIOF-GA-2008-221716, NSF CAREER 0644175, NSF CPA 0811680, and grants from Intel Corporation and Microsoft Corporation. The Bunny and Dragon models are courtesy of the Stanford Computer Graphics Laboratory.

References

BARTZ, D., CUNNINGHAM, D., FISCHER, J., WALLRAVEN, C., AND BROWN, P. 2008. State-of-the-art of the role of perception for computer graphics. In *Proc. Eurographics 2008*, State of the Art Reports, 65–86.

COLE, F., SANIK, K., DECARLO, D., FINKELSTEIN, A., FUNKHOUSER, T., RUSINKIEWICZ, S., AND SINGH, M. 2009. How well do line drawings depict shape? *ACM Trans. Graph. 28*, 3, 28:1–28:9.

DALY, S. 1993. The visible differences predictor: an algorithm for the assessment of image fidelity. In *Digital Images and Human Vision*, A. B. Watson, Ed. MIT Press, 179–206.

DEBATTISTA, K., SUNDSTEDT, V., SANTOS, L. P., AND CHALMERS, A. 2005. Selective component-based rendering. In *GRAPHITE 2005*, 13–22.

DONG, Z., GROSCH, T., RITSCHEL, T., KAUTZ, J., AND SEIDEL, H.-P. 2009. Real-time indirect illumination with clustered visibility. In *Vision, Modeling, and Visualization Workshop 2009*.

DÜR, A. 2005. On the Ward model for global illumination. Unpublished manuscript, http://homepage.uibk.ac.at/˜c70240/publications.html.

DUTRÉ, P., BALA, K., AND BEKAERT, P. 2006. *Advanced Global Illumination, 2nd Edition*. A K Peters, Natick, MA.

FLEMING, R. W., AND BÜLTHOFF, H. H. 2005. Low-level image cues in the perception of translucent materials. *ACM Trans. Appl. Percept. 2*, 3, 346–382.

FLEMING, R. W., DROR, R. O., AND ADELSON, E. H. 2003. Real-world illumination and the perception of surface reflectance properties. *Journal of Vision 3*, 5, 347–368.

HAŠAN, M., PELLACINI, F., AND BALA, K. 2007. Matrix row-column sampling for the many-light problem. *ACM Trans. Graph. 26*, 3, 26:1–26:10.

HAŠAN, M., KŘIVÁNEK, J., WALTER, B., AND BALA, K. 2009. Virtual spherical lights for many-light rendering of glossy scenes. *ACM Trans. Graph. 28*, 5, 143:1–143:6.

KELLER, A. 1997. Instant radiosity. In *Proc. SIGGRAPH 97*, 49–56.

KHAN, E. A., REINHARD, E., FLEMING, R. W., AND BÜLTHOFF, H. H. 2006. Image-based material editing. *ACM Trans. Graph. 25*, 3, 654–663.

KOLLIG, T., AND KELLER, A. 2004. Illumination in the presence of weak singularities. In *Monte Carlo And Quasi-monte Carlo Methods*, 245–257.

LAINE, S., SARANSAARI, H., KONTKANEN, J., LEHTINEN, J., AND AILA, T. 2007. Incremental instant radiosity for real-time indirect illumination. In *Eurographics Symposium on Rendering*, 277–286.

MYSZKOWSKI, K. 2002. Perception-based global illumination, rendering, and animation techniques. In *Spring Conference on Computer Graphics 2002*, 13–24.

NGAN, A., DURAND, F., AND MATUSIK, W. 2005. Experimental analysis of BRDF models. In *Eurographics Symposium on Rendering*, 117–126.

O'SULLIVAN, C., HOWLETT, S., MORVAN, Y., MCDONNELL, R., AND O'CONOR, K. 2004. Perceptually adaptive graphics. In *Proc. Eurographics 2004*, State of the Art Reports, 141–164.

PELLACINI, F., FERWERDA, J. A., AND GREENBERG, D. P. 2000. Toward a psychophysically-based light reflection model for image synthesis. In *Proc. SIGGRAPH 2000*, 55–64.

RAMANARAYANAN, G., FERWERDA, J., WALTER, B., AND BALA, K. 2007. Visual equivalence: towards a new standard for image fidelity. *ACM Trans. Graph. 26*, 3, 76:1–76:11.

RAMANARAYANAN, G., BALA, K., AND FERWERDA, J. A. 2008. Perception of complex aggregates. *ACM Trans. Graph. 27*, 3, 60:1–60:10.

REINHARD, E., STARK, M., SHIRLEY, P., AND FERWERDA, J. 2002. Photographic tone reproduction for digital images. In *Proc. SIGGRAPH 2002*, 267–276.

RITSCHEL, T., GROSCH, T., KIM, M. H., SEIDEL, H.-P., DACHSBACHER, C., AND KAUTZ, J. 2008. Imperfect shadow maps for efficient computation of indirect illumination. *ACM Trans. Graph. 27*, 5, 129:1–129:8.

SEGOVIA, B., IEHL, J.-C., AND PÉROCHE, B. 2007. Metropolis instant radiosity. *Computer Graphics Forum 26*, 3, 425–434.

SLOAN, P.-P., KAUTZ, J., AND SNYDER, J. 2002. Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. In *Proc. SIGGRAPH 2002*, 527–536.

STOKES, W. A., FERWERDA, J. A., WALTER, B., AND GREENBERG, D. P. 2004. Perceptual illumination components: A new approach to efficient, high quality global illumination rendering. *ACM Trans. Graph. 23*, 3, 742–749.

VANGORP, P., AND DUTRÉ, P. 2008. Shape-dependent gloss correction. In *APGV '08: Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, 123–130.

VANGORP, P., LAURIJSSEN, J., AND DUTRÉ, P. 2007. The influence of shape on the perception of material reflectance. *ACM Trans. Graph. 26*, 3, 77:1–77:10.

WALD, I., KOLLIG, T., BENTHIN, C., KELLER, A., AND SLUSALLEK, P. 2002. Interactive global illumination using fast ray tracing. In *Eurographics Workshop on Rendering*, 15–24.

WALTER, B., FERNANDEZ, S., ARBREE, A., BALA, K., DONIKIAN, M., AND GREENBERG, D. P. 2005. Lightcuts: a scalable approach to illumination. *ACM Trans. Graph. 24*, 3, 1098–1107.

WALTER, B., ARBREE, A., BALA, K., AND GREENBERG, D. P. 2006. Multidimensional lightcuts. *ACM Trans. Graph. 25*, 3, 1081–1088.

WARD, G. J. 1992. Measuring and modeling anisotropic reflection. In *Proc. SIGGRAPH 92*, 265–272.

WESTLUND, H. B., AND MEYER, G. W. 2001. Applying appearance standards to light reflection models. In *Proc. SIGGRAPH 2001*, 501–510.

YU, I., COX, A., KIM, M. H., RITSCHEL, T., GROSCH, T., DACHSBACHER, C., AND KAUTZ, J. 2009. Perceptual influence of approximate visibility in indirect illumination. *ACM Trans. Appl. Percept. 6*, 4, 24:1–24:14.