# PERCEPTUAL ILLUMINATION COMPONENTS: A NEW APPROACH TO EFFICIENT, HIGH-QUALITY GLOBAL ILLUMINATION RENDERING

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

William Adams Stokes

August 2004

## ABSTRACT

We introduce a new perceptual metric for efficient, high quality, global illumination rendering. The metric is based on a rendering-by-components framework in which the direct, and indirect diffuse, glossy, and specular light transport paths are separately computed and then composited to produce a high quality image. The metric predicts the perceptual importances of the computationally expensive indirect illumination components with respect to image quality. To develop the metric we conducted a series of psychophysical experiments in which we measured and modeled the perceptual importances of the components. An important property of this new metric is that it predicts component importances from inexpensive estimates of the reflectance properties of a scene, and therefore adds negligible overhead to the rendering process. This perceptual metric should enable the development of an important new class of efficient global-illumination rendering systems that can intelligently allocate limited computational resources, to provide high quality images at interactive rates.

# Biographical Sketch

William Adams Stokes was born in Chicago, Illinois on July 31st of 1980. He grew up on the south side of the city and from a young age showed great interest in art, science, and technology. Will began programming in Logo in grade school. In high school he spent countless hours teaching himself how to program Texas Instruments graphing calculators in BASIC and later Z80 assembly language, sharing games he produced with students around the world via the world wide web.

Will traveled to Ithaca, NY in the fall of 1998, to began his undergraduate degree in computer science at Cornell University. During his senior year he took Professor Donald Greenberg's introductory computer graphics course, and was reunited with his passion for graphics he had not explored since high school. This course inspired him pursue a Masters degree in the Program for Computer Graphics at Cornell University in the fall of 2002.

*To Mom and Dad*

# Acknowledgements

I would like to begin by thanking Donald Greenberg for providing me with the opportunity to extend my study of computer graphics at Cornell's Program for Computer Graphics. His advice, background, and vision, were instrumental throughout my exploration of advanced topics in computer graphics I had barely considered during my undergraduate education. You have been most responsible for kindling my interest in computer graphics and creating a most throughly enjoyable experience these last two years. Thanks Don.

My research would not have been as successful without the regular advice and help with Jim Ferwerda, Bruce Walter, and Kavita Bala. Learning from each of you with respect to perception, global illumination, and rendering, has been an incredibly rewarding experience.

Special thanks is due to Henry Letteron for being a great officemate and candy supplier, friend, and partner in several courses we took together during the last two years. I'm sure we will keep in touch. Special thanks is required for Sebastian Fernandez for many lively conversations, a recent series of ping pong challenges, and for tempting me to work on various projects not directly related or entirely unrelated to my thesis project during my stay at the PCG. I would like to thank Vikash Goel, who sparked my interest in computer graphics as a junior, and was instrumental in getting me to apply to (and

possibly getting into) the PCG. I owe being here to you. I also would like to acknowledge my fellow lab mates and lunch buddies Jacky Bibliowicz, Aaron Birkland, Mike Donikian, Jeramiah Fairbank, Ryan Ismert, Adam Kravetz, HoTngsong Li, John Mollis, and Jeremy Selan for giving life to the lab.

This section would be incomplete if I did not thank Steve Marschner, Ken Torrance, Steve Westin, Peggy Andersen, Martin Berggren, Hurf Sheldon, and Linda Stephenson. While not directly involved in my research, the lab would not be where it is today without your guidance and work keeping things running smoothly.

I would like to thank Jeremiah Fairbank for building the kitchen model, my father Houston H. Stokes for statistical advice throughout the modeling phase of the project, and all participants for taking the time to sort the many many tiles through multiple rounds of experimentation. This work was made possible by NSF grant ASI-0205438, and Intel Corporation for generously providing funding and equipment.

Finally I would like to thank my girlfriend Melissa Stickle and the remainder of my family for their ongoing love and support.

# Table of Contents


**1 Introduction** **1**

**2 Previous Work** **5**
2.1 Tone Reproduction in Rendering . . . . . 5
2.2 Visual System Metrics in Rendering . . . . . 9
2.3 Leveraging Spatiotemporal Effects . . . . . 16
2.4 Summary . . . . . 19

**3 Perceptual illumination components** **21**
3.1 Rendering by components . . . . . 21
3.2 Perceptual illumination components . . . . . 24

**4 Experiments** **26**
4.1 Stimuli . . . . . 26
4.2 Equalizing contrast and hue . . . . . 35
4.3 Procedure . . . . . 39
4.3.1 Results and Preliminary Analysis . . . . . 40
4.4 Determining the perceptual importances of the illumination components 42

**5 Formulating the metric** **45**
5.1 Defining the indicators . . . . . 46
5.2 Indicator response functions . . . . . 47
5.3 Modeling perceptual importance . . . . . 49
5.4 Selecting the appropriate model . . . . . 50
5.4.1 Formulating the perceptual metric . . . . . 55
5.5 Using the metric . . . . . 55
5.6 A perceptually-based component renderer. . . . . 58

**6 Conclusions/Future Work** **60**
6.1 Applying a Decision Theoretic . . . . . 61
6.2 Cost Quality Functions . . . . . 64
6.3 Additional Future Work . . . . . 65

**A Parameters used in the metric** **69**

**Bibliography** **70**

# List of Figures


1.1 Our perceptually based rendering by components approach to efficient, high quality, global illumination rendering. . . . . 2

3.1 Heckbert notation . . . . . 22
3.2 Examples of direct and indirect illumination components . . . . . 23

4.1 Office and kitchen scenes . . . . . 27
4.2 Viewpoints tested in the office scene . . . . . 29
4.3 Viewpoints tested in the kitchen scene . . . . . 30
4.4 Composites which make up a ranking set . . . . . 31
4.5 Compositing a global illumination image . . . . . 33
4.6 Differences between gold and silver standard images . . . . . 34
4.7 Accounted luminance . . . . . 35
4.8 Equalizing contrast . . . . . 36
4.9 Hue shifting the ambient term . . . . . 38
4.10 Ranking a set of images . . . . . 39
4.11 Linearized rank data . . . . . 41
4.12 Computed perceptual importances . . . . . 43

5.1 Indicators . . . . . 48
5.2 Indicator response functions . . . . . 49
5.3 Breaking up low and high gloss . . . . . 52
5.4 Computed and predicted $r^2$ values and perceptual importances . . . . . 54
5.5 Applying our perceptual metric . . . . . 56
5.6 A hypothetical perceptual component rendering system. . . . . 59

6.1 Possible optimal resource allocations . . . . . 61
6.2 A hill climbing approach to resource allocation. . . . . 63
6.3 Reaching the optimal resource allocation. . . . . 64
6.4 Cost quality functions . . . . . 66
6.5 Scaling cost quality functions using predicted perceptual importances . 67

# Chapter 1

# Introduction

Global illumination effects, while necessary for improved realism, are often omitted because of their high computational cost. Traditionally, global illumination simulations have only been available from computationally intensive off-line systems or low quality interactive alternatives. The goal of our research is to enable the production of high quality global illumination renderings at interactive rates by approaching the rendering problem from a perceptual standpoint.

We propose that an efficient approach to providing global illumination effects is to decompose the global illumination simulation into direct and indirect components. By quantifying the perceptual importance of each of these components, we can construct a perceptual quality metric that can be used to efficiently allocate computational resources and maximize image quality within system constraints. We ultimately envision an adaptive system which attempts to predict the relative importance of the pure indirect components, and adaptively allocates resources to their independent simulation before compositing with a direct illumination solution for display. Such a system is schematically shown in Figure 1.1. In this system, viewpoint and scene specific material statistics are obtained during the initial direct illumination simulation (step 1). These statistics are

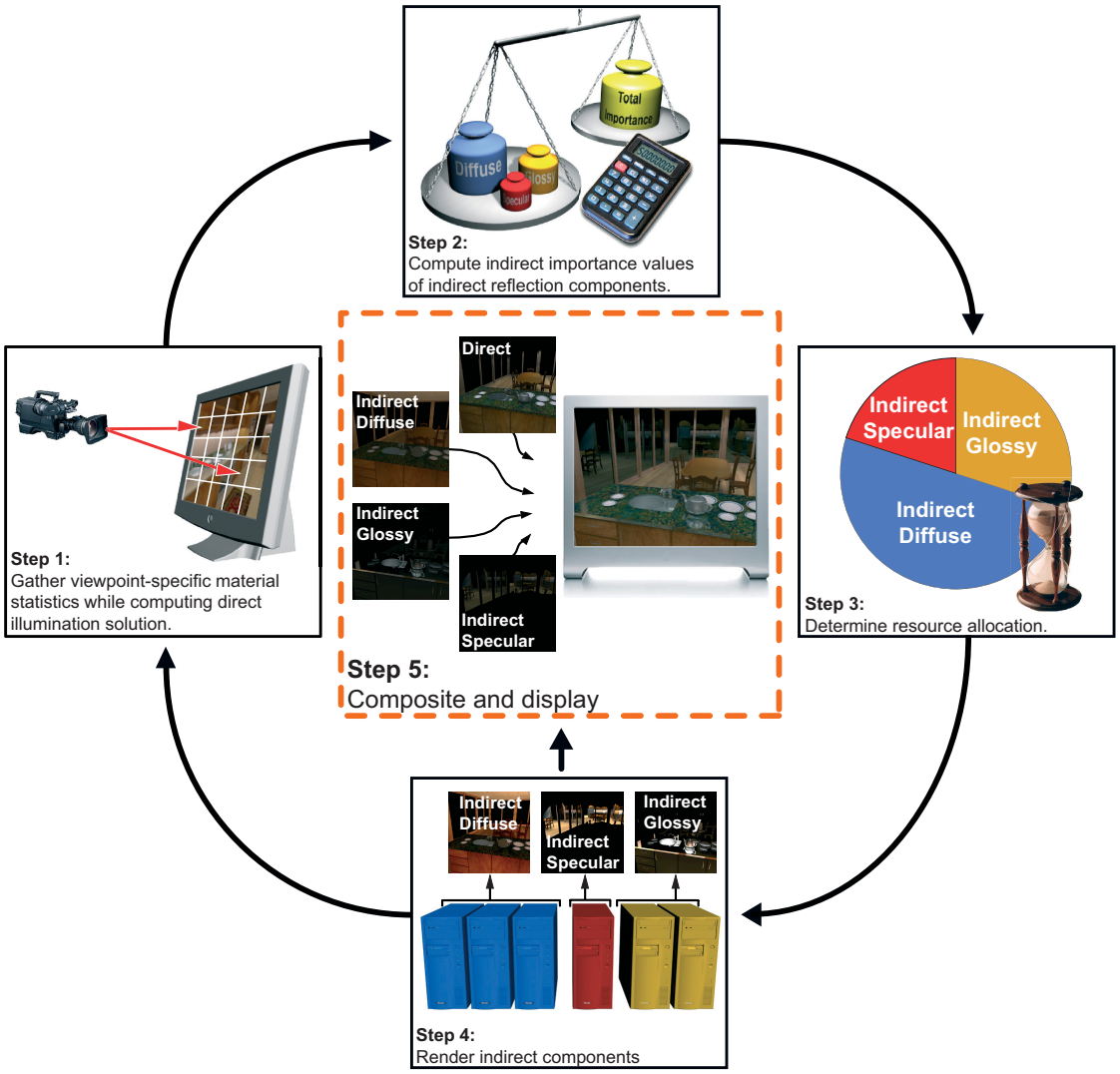


Figure 1.1: Using the perceptually based, adaptive rendering system system we propose, material statistics gathered during a direct illumination pass guide resource allocation for the independent approximation of indirect illumination components. A direct illumination solution in combination with approximated indirect illumination components are composited to produce an optimal global illumination image given an overall resource budget.

then filtered through a perceptual metric in order to predict the relative importance an actual observer would assign to each indirect component (step 2). The optimal allocation of computational resources is found using the relative perceptual importances and prior knowledge of how each indirect illumination approximation technique converges over time (step 3). Finally, the indirect components are computed using the computational resources assigned (step 4), and composited with the direct illumination solution before being displayed (step 5). This entire process would be repeated every frame in an interactive system, although evaluation of the perceptual metric and computational breakup optimization might benefit from being run asynchronously in order to minimize temporal artifacts which may arise.

In order to develop the required perceptual metric for a rendering by components framework, we first conducted a series of experiments in which subjects ranked the quality of global illumination images rendered using different combinations of direct and indirect components. Using the data provided by these experiments, we derived a ranking of the perceptual importance of each of the components. We then constructed a model that can predict these importance values by measuring the physical reflectance properties of the objects visible in a particular image. We formulated a perceptual metric that can be used to predict the visual quality of different global illumination renderings from knowledge of the physical scene characteristics and the illumination components used in the rendering. Finally, while we have yet to build a working implementation of our proposed rendering by components framework, we explored various approaches to performing triage in order to determine the optimal resource allocation for rendering the various indirect reflection components.

We believe this new metric will enable the development of an important new class of efficient global-illumination rendering systems that can intelligently allocate com-

putational resources between the components of the global illumination simulation to provide high quality renderings at interactive rates.

Chapter two reviews previous perceptually-based rendering techniques and chapter three introduces our novel rendering by components approach. Chapter four describes the perceptual experiments that were conducted and the process by which resulting data is used to compute image quality and indirect component importance values. In chapter five we discuss the process by which we developed a series of perceptual metrics capable of predicting indirect component importance values using material statistics. Finally, chapter six explores the steps that remain before the proposed rendering by components framework can be realized.

# Chapter 2

# Previous Work

This chapter reviews previous perceptually based attempts to accelerate rendering systems. Existing work in the field falls into three categories: (1) algorithms that leverage the compression of luminance values resulting from the application of tone mapping operators, (2) more sophisticated rendering techniques for static environments based on complex visual system metrics, and (3) a recent set of work that takes advantage of limitations of the human visual system with respect to motion when rendering static walkthrough sequences and dynamic environments. Each of these sections will be discussed in turn.

## 2.1 Tone Reproduction in Rendering

The human visual system is able to perceive luminance values between $10^{-5}$ to $10^{5}$ $\frac{cd}{m^2}$, while conventional CRT and LCD displays can only reproduce luminance values between approximately 1 and 100 $\frac{cd}{m^2}$. Tone reproduction operators have been developed to map real world luminance values onto display luminance values, taking into account the limits of the particular display device and the visual state of the scene and display

observers. This mapping is applied such that the visual experience of viewing the displayed image is as close as possible to the experience of a standard observer viewing the actual scene.

Various tone reproduction operators, which strive to match overall brightness [TR93] or apparent contrast and visibility [War94], have been developed. Considerable research on tone-mapping operators exists but goes beyond the scope of this thesis. Ramasubramanian [Ram00] provides an excellent summary of such work. Adaptive rendering systems benefit greatly by using tone-mapped screen luminance values instead of raw radiance values for refinement criteria, since large variance in raw samples is often visually undetectable once tone mapped due to the inherent dynamic range compression that takes place. Using tone-mapped luminance values allows adaptive techniques to avoid wasting computational effort in refining such visually undetectable noise.

Tamstorf and Jensen [TJ97] incorporated a simple tone reproduction operator into an adaptive sampling technique, which used confidence intervals [Pur87] in order to remove nonuniformly distributed noise resulting from Monte Carlo sampling. By continuing sampling until displayed luminance values are within a given tolerance of an estimate, they were able to reduce noise and achieve homogeneous image quality. The authors presented this acceleration technique as an alternative to computationally inexpensive filtering techniques because filtering invariably introduces unpredictable errors in the form of blurring. Unfortunately, the drawback of any approach based on tone-mapping operators is that results should not be used with display types which have a different dynamic range.

Gibson and Hubbold [GH97] incorporated Tumblin and Rushmeier's tone-mapping operator [TR93] into their progressive radiosity algorithm. They developed a method that estimates ambient illumination of a scene in order to establish a view-independent

real-world adaptation level necessary for applying the tone-mapping operator. Path radiance values are first converted to displayed color values. A nonlinear transformation function then converts these values from a quantitatively linear scale (CIE XYZ) to the perceptually linear domain (CIE LUV). At this point they are able to use a perceptually based, just noticeable difference (JND) threshold to terminate patch refinement by computing the Euclidean distance between colors. Similar applications of the tone-mapping operator and JND thresholds are used to optimize form-factor computation and *a posteriori* mesh optimization portions of their radiosity algorithm.

Gibson and Hubbold [GH00] improved on their previous work [GH97] by developing a voxel-based, uniform grid structure for accelerating visibility testing during radiosity simulations. They applied a perceptual technique that identified energy transfers not adversely effected by this efficient but inexact alternative to traditionally computationally expensive form-factor computation. They found that 98% of energy transfers performed during the course of a radiosity simulation do not require accurate visibility testing, resulting in a 50 times speedup without a reduction in image quality.

Hedley et. al. [HWP97] applied a similar tone reproduction metric to that used by Gibson and Hubbold [GH97] in order to form a perceptual basis for thresholding when discontinuity meshing during the adaptive subdivision stage of their radiosity algorithm. Using this technique, they were able to remove subdivisions in areas of low illumination, while leaving most subdivisions in brighter regions where shadow boundaries are more noticeable. This produces more compact meshes necessary for interactive radiosity and dynamic mesh updating techniques, while limiting the introduction of visible artifacts.

Walter et. al. [WHSG97] took advantage of the nonlinear response of the human visual systems (HVS) to luminance changes in the mesh decimation phase of their view-independent density estimation technique. By applying a simple tone-mapping operator

in order to obtain display luminance values, a conservative initial mesh is progressively decimated by removing vertices provided that the resulting change in luminance is below a perceptually based threshold. Walter et. al. formalize this step in a manner that allows the reuse of previous geometric simplification techniques. Reduced meshes maintain as much accuracy and detail as possible, while being small enough to be displayed at interactive rates.

Prikryl and Purgathofer [PP99] developed a perceptually based termination heuristic for accelerating stochastic radiosity algorithms. The authors were not satisfied with the technique presented by Myszkowski [Mys98], which used a representative set of viewpoints, and the Daly [Dal93] visual difference predictor (VDP), to establish a termination criterion for radiosity techniques. VDP's encapsulate advanced models of the human visual system, and will be discussed in Section 2.2. They argue it is difficult to construct small, representative sets that minimize VDP computational overhead, while containing enough representative viewpoints necessary for accurate evaluation of termination criteria. They developed a true view-independent technique for terminating radiosity simulations, using Gibson and Hubbold's [GH97] method for estimating ambient scene illumination in conjunction with the tone-mapping operators of Tumblin and Rushmeier [TR93] or Ward [War94].

Prikryl and Purgathofer present a heuristic that predicts the iteration at which maximum-displayed luminance change falls below a perceptually based threshold, using interpolated sample data. While this technique works fairly well, they found pixel computations with high initial variance were terminated too quickly, while their approximation heuristic tended to overestimate the proper stopping iteration later in radiosity simulations. They suggested various solutions to these problems which take into account variance in obtained data. By avoiding application of more sophisticated VDP's, their

technique only leverages limitations of the HVS with respect to the contrast sensitivity function (CSF).

## 2.2 Visual System Metrics in Rendering

The human visual system is an incredibly complex system exhibiting many nuances that can be leveraged for accelerating rendering algorithms. Several perceptual metrics have been developed that strive to predict HVS response to stimuli. Ahumada [Ahu93] and Ramasubramanian [Ram00] provide excellent reviews and reference lists of the computational image quality metrics developed by the vision science and computer graphics fields. Nonlinear response, contrast sensitivity, disparate chromatic and achromatic spatial acuity, masking effects, and many other effects have been incorporated in various combinations into these models. Initially developed in order to determine if two images are perceptually indistinguishable, these models have grown to provide local discrepancy information and have been applied to guide adaptive sampling algorithms, in addition to many other rendering techniques.

A number of techniques have been developed which accelerate still image generation through the application of a visual metric. Unfortunately, far fewer techniques for accelerating view independent algorithms have been studied.

Mitchell [Mit87] developed an adaptive anti-aliasing technique, which guides nonuniform sampling techniques based on how the HVS perceives noise as a function of contrast and color. An initial coarse, nonuniform sampling of the image plane provides the data necessary to determine if further sampling is necessary at each pixel, based on contrast variation. Regions of the image that require additional sampling are isolated through the evaluation of a contrast sensitivity function and the use of separate

thresholds for each color channel. Red, green, and blue contrast levels are weighted independently in a crude approximation of the HVS's wavelength-specific contrast sensitivity functions. This technique successfully isolates aliasing noise to regions of an image that contain higher spatial frequencies where noise is less visible.

Meyer and Liu [ML92] developed a spatial subdivision method for accelerating ray tracing, based on the limited color spatial acuity of the HVS. By working in the $AC_1C_2$ color space and using a K-d structure to store sample pixel data, they were able to reduce color calculations by limiting the depth in the tree at which each channel is sampled, based on channel-specific acuity of the HVS. Consequently, achromatic image detail was computed at a higher resolution than that used for the computation of either chromatic channel. Psychophysical testing demonstrated that the resulting reductions in the number of rays used to compute chromatic channels do not have as severe an effect on image quality as do similar reductions in the number of rays used to compute the achromatic channel. Unfortunately, adaptive rendering techniques based on this method only afford modest improvements in speed through reduction of color computation.

Bolin and Meyer [BM95] developed a ray-tracing algorithm that exploits limitations of the HVS in the intensity, spatial, and temporal domains, by synthesizing images directly into the frequency domain. Their three-stage vision model first simulates the spectral sensitivities of the cones in the human eye. Contrast sensitivity is emulated by scaling response values using an amplitude nonlinearity function. Cone signals are subsequently transformed to an opponent color space. Finally, a spatial filter is applied to the achromatic and chromatic channels.

The developed HVS model was used to guide sampling by selecting the next image block to be sampled, using a quad tree subdivision method and controlling the spawning of indirect rays that intersect other objects in the environment. Variation in luminance

results in higher sampling rates in regions with lower average intensity, thus taking advantage of the HVS's nonlinear contrast sensitivity. Monochromatic differences receive higher sampling rates than chromatic differences, resulting from differences in the cutoff thresholds of the perceptual spatial frequency response for these channels. Finally, because a spatial frequency representation of the rendered image is directly computed from samples taken, visual masking effects are partially leveraged by reducing the number of subsequent rays spawned in regions where higher spatial frequency detail is observed.

Gaddipatti et. al. [GMY97] suggested using a wavelet-based perceptual metric to compare intermediate images in the termination evaluation portion of adaptive sampling, image synthesis algorithms. They also incorporated the frequency response portion of the HVS through the use of a contrast sensitivity function (CSF). This differed from previous work, which used mean squared error (MSE) techniques, since their method took into consideration variation in images at specific locations, orientations, and scales.

Compared images are split into subimages, using a dyadic subdivision of the images' frequency spectrum. A simple structure detection scheme identifies scale-coherent structures (edges, corners, ridges, etc.). Coefficients of the wavelet transformation are modulated, using the CSF in order to take into account frequency response to the HVS. Mean squared distances between modulated coefficients of the compared images are combined, providing a single scalar value as a measure of the differences of the images. This approach ignores masking effects, and the single scalar measure of distance prevents its application to any nonuniform adaptive sampling technique.

Neumann et. al. [NMP98] developed an alternative approach to comparing images which is more sensitive to differences in color, while avoiding the costs associated with Fourier space transformations and the limitations of wavelets associated with previous image comparison techniques. In their approach, average CIE LUV color differences of

quasirandomly chosen rectangles in input images are used to compute a single perceptual distance measure. Rectangle size and number are chosen based on view distance. Errors are weighted using rectangle size and the contrast sensitivity function. Overall, this was an ad hoc technique, which used very simplistic approximations of a limited number of HVS characteristics.

Bolin and Meyer [BM98, BM99b] developed an adaptive sampling algorithm that integrates an improved version of the Sarnoff Visual Discrimination Model (VDM) [Lub95]. During image refinement, boundary images are constructed, using the current image approximation and an estimate of its error. These upper and lower boundary images are fed into their improved vision model to obtain a visual difference map, which in turn is used to determine the next sample location.

Bolin and Meyer's adapted Sarnoff VDM was improved in order to run efficiently, thus allowing it to be used to direct sampling and terminate image refinement once no visible errors are present. The vision model was also extended to handle the significant effects of chromatic abberation and variations in achromatic and chromatic contrast sensitivity as a function of spatial frequency. Unfortunately, this adaptive rendering technique incurs substantial overhead, due to evaluation of their image quality model after each sample is taken.

More recently, Bolin and Meyer [BM99a] presented an enhanced version of their adaptive sampling algorithm [BM98, BM99b], which addresses the computational expense involved by limiting the number of times the visual metric is applied. By performing a preliminary low-density sampling of an image prior to initiating their adaptive sampling algorithm, considerable computational overhead is avoided when large sample variance is expected at initial low sampling rates. Further computational savings are achieved during the iterative refinement process by sampling all regions of intermediate

images which contain noise above a user specified threshold before reapplying the visual error metric. Unfortunately, these advances do not provide the order of magnitude improvements necessary for use in an interactive setting.

Ramasubramanian et. al. [RPG99] developed a new framework for perceptually based rendering. Their method reduces the computational costs associated with previous techniques that require the evaluation of expensive VDP's, by precomputing the expensive, spatially dependent components of such models. Ramasubramanian et. al. suggest direct illumination solutions capture most of the higher spatial frequency and contrast content in a scene, thus validating the use of a composite of a direct illumination solution with an ambient image (in order to incorporate spatial frequency in regions not receiving direct illumination) so as to construct their initial elevation factor map. During indirect illumination computation, an inexpensive luminance-dependent threshold map is generated, using a composite of the current indirect illumination approximation and the precomputed direct illumination solution. Elevating this threshold map by the precomputed elevation factor map provides a set of luminance contrast thresholds that are then used to steer subsequent sampling.

Walter et. al. [WPG02] presented a technique for accelerating any algorithm that uses per-pixel error thresholds. While the technique described by Ramasubramanian et. al. [RPG99] is able to substantially accelerate rendering global illumination solutions, their approach requires a preprocessing step that prohibits its use for time-critical direct illumination and interactive environments. Walter et. al. present an alternative method for generating elevation factors based on the spatial frequencies and base contrasts present in material textures, by preprocessing all scene textures using an inexpensive, discrete cosine transform (DCT) method based on the JPEG standard. They show how such elevation factors can easily be stored and looked up using existing texture

mip-map techniques. These values can be combined with luminance masking based on Weber's law, in order to improve the performance of adaptive shadow testing, irradiance caching, and path-tracing algorithms. Possibly the most interesting feature of this approach is its freedom from rendering an initial image, allowing its immediate application to dynamic environments.

Myszkowski [Mys98] studied the applicability of the Daly VDP [Dal93] in deciding on a stopping condition for progressive radiosity and Monte Carlo techniques, in addition to driving adaptive mesh subdivision in radiosity computations. He studied multiple scenes in order to determine the optimal spacing of intermediate images for comparison in his path-tracing termination technique. If a view-independent solution is desired, a set of representative views is used. Myszkowski used the Daly VDP, despite the fact that more advanced vision models exist [FPSG97, GMY97], since such advanced models return a single scalar value as a measure of the differences between two images. In contrast, the Daly VDP returns a perceptual difference map, providing local predictive ability required by adaptive refinement techniques.

Many previous adaptive refinement techniques rely on energy-based error metrics, which do not necessarily correspond to visible improvements in image quality. By applying the Daly VDP to intermediate images, Myszkowski was able to leverage Weber's law-like amplitude compression, an advanced contrast sensitivity function model, and masking effects, all which limit visual acuity, in order to terminate the applied simulation technique. Unfortunately, his brute force implementation of the Daly VDP, which required evaluation at every iteration, incurs substantial computational expense.

Recently, a few authors have developed novel applications of visual system metrics that do not precisely fall into the categories of adaptive sampling or view-independent algorithms.

Volevich et. al. [VMKK00] explored a view-independent technique for accelerating the convergence of progressive global illumination solutions. While previous perceptually based work has sought to directly integrate VDP's into the refinement stage of image generation, Volevich et. al. state that such integration incurs considerable computational expense, since the selected perceptual metric must be computed during each iteration, negating any performance improvements such perceptually aware techniques provide. They suggest an alternative approach, which uses a VDP to select the order in which, and the related switchover points between, a pool of global illumination techniques which are applied when rendering a global illumination solution. They applied the technique to a series of environments in order to determine the order and switchover points for applying deterministic direct lighting (DDL), hierarchal progressive radiosity (HPR), and density estimation photo tracing (DEPT) methods, using the Daly VDP [Dal93] during an initial preprocessing step.

Dumont et. al. [DPF03] developed a perception-based decision theoretic methodology for realistic rendering at interactive rates. While they applaud previous perception-based techniques, the authors suggest that the demands of interactive rendering prohibit the evaluation of expensive VDP models in order to remove *all* noticeable noise in rendered images. In an interactive setting, they suggest one cannot strive to produce an image that is visually indistinguishable from the gold standard; rather, one must attempt to render the highest quality image within the given constraints.

They outline a perceptually based decision theoretic framework, which forms the backbone of a texture cache management scheme necessary for producing global illumination effects at interactive rates. Optimal texture resolutions are determined through a triage technique that minimizes cost while maximizing utility. Since they use extensive hardware techniques, such as environment maps, to provide approximations of specular

and glossy reflections, texture memory and texture processing quickly become limited commodities, further justifying their cache management scheme.

It should be noted that while the foundation of the perceptual basis of their decision methodology considers luminance sensitivity, contrast sensitivity, and visual masking effects, they currently have not leveraged temporal effects on contrast sensitivity inherent in any interactive environment.

## 2.3 Leveraging Spatiotemporal Effects

More recently, a great deal of research has attempted to incorporate temporal effects on HVS acuity into rendering systems. The HVS experiences decreased contrast sensitivity with increasing motion; however, the eye exhibits a miraculous tracking ability, which is able to significantly increase the acuity of objects of attention. Perceptually based algorithms, which leverage these effects, can significantly accelerate the rendering process required for accurate simulations in interactive and real-time settings.

Myszkowski et. al. [MRT99, MK00] were the first to develop and apply a spatiotemporal perceptual metric for the purpose of accelerating the rendering of antialiased walkthrough sequences. They developed a hybrid rendering technique, which used expensive raytracing techniques to render key frames, and inexpensive image-based rendering (IBR) techniques to construct in-between frames. Using their approach, a fixed set of key frames are optimally placed in order to minimize the number of pixels in in-between frames that cannot be properly derived from key frame data due to visibility problems. Next, an animation quality metric (AQM) is used to guide in-between frame computation. Their AQM is an extended version of Erickson's static image quality metric [RE98], which incorporates the spatiovelocity CSF.

Myszkowski et. al. apply models of the HVS's increased sensitivity to distortions in image regions with low image flow (IF) velocities, using ray-tracing techniques to compute pixel values of in-between frames when IF velocity values fall below a specified threshold. Ray-tracing techniques are also applied in order to compute strong specular effects and glossy reflectance when the AQM indicates more accurate computation is required. Like most IBR techniques, ray-tracing is used to fill in remaining holes in in-between frames. Finally, a novel perception-based antialiasing technique was explored that, using information about the IF, allows traditionally super-sampled keyframe pixels to be replaced with raw pixels derived using IBR or single sample ray-tracing techniques.

Myszkowski et. al. [MTAS01] developed a perceptually guided global illumination technique for rendering dynamic environments. Their approach concentrated on taking advantage of the generally high temporal coherence in the indirect component. Sparse samples of indirect illumination across a number of local frames are combined with a dense sampling of direct illumination and specular effects at each frame to produce high-quality interactive animation sequences.

The approach of Myszkowski et. al. works as follows. First, the appropriate number of photons for each frame in the current frame segment (collection of frames from which indirect illumination photo samples can be used) is estimated. Photon sampling is then conducted for all frames. The maximum segment size (number of frames forward and backwards in time) is refined, based on temporal variations in indirect lighting using energy-based criteria. The average number of photons per frame is adjusted, based on the response of an AQM previously presented by Myszkowski et. al. [MRT99, MK00]. Finally, spatiotemporal reconstruction and spatial filtering are employed to construct an indirect illumination image. Direct illumination and specular effects are computed

at each pixel, using a standard ray tracer, and combined with a constructed indirect illumination solution to produce a single animation frame.

Haber et. al. [HMYS01] also took advantage of spatiotemporal effects on perceptual importance in their interactive walkthough system. Indirect diffuse illumination is first computed during a preprocessing step. As a user walks about an environment, sampling of direct illumination and other view-dependent lighting effects is guided using an extended version of the attention-driven perceptual metric developed by Itti et. al. [IKN98]. Cached samples are splatted on the screen using a technique similar to the render cache developed by Walter et. al. [WDP99]. In addition, the vision metric is applied to determine regions of interest that require further sampling. Because indirect diffuse illumination is computed during an initial preprocessing step, this approach is limited to static environments. However, unlike the technique developed by Yee et. al. [YPG01] mentioned below, since camera position is interactively controlled by the user, they argue they are unlikely to suffer from degraded quality resulting from repeated viewing experiences.

Yee et. al. [YPG01] developed a method for accelerating global illumination computation in prerendered animations by taking advantage of spatiotemporal limitations of the HVS. Their approach leverages decreased contrast sensitivity as image velocity increases, while adapting to the compensation for such loss of sensitivity due to motion through the smooth pursuit tracking capabilities of the eye.

When applying the technique of Yee et. al., an error tolerance map, based on the velocity-dependent contrast sensitivity function, is constructed using a rapid image estimate of the scene. This map is augmented by the viewers predicted visual attention, using a saliency map to compute an Aleph map. The Aleph map is a complete estimate of the spatiotemporal sensitivity of the HVS, and is used to terminate the global

illumination rendering technique used to compute displayed images.

While eye-tracking hardware exists, the authors opted to extend Itti et. al's [IKN98] computational model of stimulus-driven visual attention, by including a fourth conspicuity channel for motion when producing saliency. By opting to use a software-based approach, no extra work is necessary to allow simultaneous viewers of animated sequences. Unfortunately, they found visual attention was often poorly predicted, once viewers were shown the same animation sequence multiple times, resulting in visible artifacts in inaccurately predicted regions of interest.

## 2.4 Summary

This chapter reviewed previous applications of tone-mapping operators and perceptual metrics used to improve the performance of various rendering algorithms. These were categorized into systems that use tone-mapping operators, static environment visual metrics, and more advanced visual metrics that take into account temporal effects on visual acuity.

The first set of algorithms control image refinement using tone-mapped luminance values instead of sample radiance values. Display devices are unable to reproduce the full range of observable real world luminance values. Tone-mapping operators must be applied to sample data in order to match the subjective experience of viewing a real scene by mapping real world luminance values to the limited range of a particular display device. The compression of luminance variance that results is often not visually detectable. Rendering algorithms, which take advantage of this tone-mapping process, can accelerate the rendering process.

The second set of algorithms further accelerate the rendering process by applying

visual system metrics that incorporate nonlinear response, wavelength and frequency-specific contrast sensitivity functions and masking effects that nonuniformly limit visual acuity across the image plane.

Finally, a third set of algorithms applied more recently developed visual metrics that model temporal effects and tracking behavior on visual acuity. These algorithms strive to accelerate rendering systems in order to facilitate their application in dynamic environments and static walkthough sequences.

This thesis will take an entirely different approach to applying perceptual constraints to image generation by adopting a rendering-by-components approach to the global illumination problem. While some previous work has split the rendering process into direct and indirect components, this is the first work to split indirect illumination into three pure component channels, and study how the relative perceptual importance of each of the various indirect reflection components can be applied to rendering systems.

# Chapter 3

# Perceptual illumination components

In global illumination rendering, the radiance of a pixel in the rendered image can be determined by finding the intersection between the pixel's view ray and a surface, and calculating the sum of the energy emitted by the surface and the energy reflected in the direction of the view ray. This can be expressed as:

$$L^{out}(x) = \underbrace{L^{emission}}_{energy\ emitted} + \underbrace{\int_{\Omega}^{\substack{incoming\\directions}} L^{reflected}\,\partial\omega}_{energy\ reflected} \tag{3.1}$$

## 3.1 Rendering by components

Heckbert [Hec90] was first to analyze the subdivision of photon paths into components. Figure 3.1 illustrates the various paths by which photons can travel from a light (L) to the eye (E) by reflecting off diffuse (D) and specular (S) surfaces. Extension of this terminology with the inclusion of a fifth term (G) is necessary in order to incorporate the remaining set of glossy surfaces. All paths contributing a person's visual experience can be characterized by the regular expression $L(D|G|S)*E$, where |'s are logical OR's and * indicates zero or more reflections. By including an additional symbol ? signifying an optional reflection, ray tracing can be characterized as a subset of all possible photon

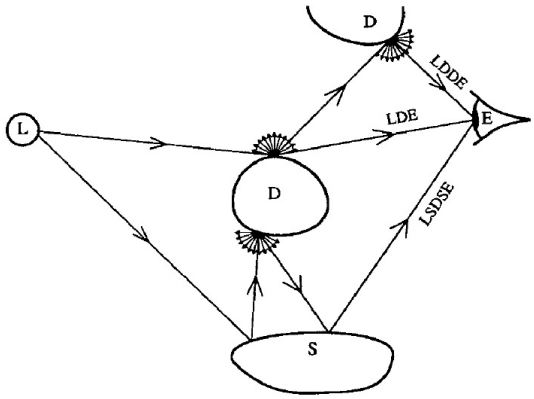


Figure 3.1: Selected photo paths from light (L) to eye (E) by way of diffuse (D) and specular (S) surfaces. Additional glossy (G) surfaces, while not shown, also exist.

paths which are matched by the expression $LD?S*E$. Similarly, radiosity simulations, which only capture indirect diffuse and direct diffuse illumination, can be expressed as $LD*E$. Finally, caustic paths, which are the set of paths which must reflect off one or more specular surfaces before reflecting off a diffuse surface toward the eye, are matched by the expression $LS+DE$ where + matches one or more reflections. Heckbert notation provides two insights. First, many interesting photon path sets can be expressed using regular expressions, and second, the set of all possible photon paths, which equate to a full global illumination solution, can be broken up into components and evaluated individually.

The starting point of our work is the insight that the global illumination simulation process can be successfully modeled by calculating light transport along direct and indirect illumination paths and combining the results. This allows the integral term in Equation 3.1 to be subdivided into direct and indirect components:

$$\int_{\Omega}^{\substack{incoming\\directions}} L^{reflected}\partial\omega = \begin{cases} \int_{\Omega} L^{direct}\partial\omega & + \\ \int_{\Omega} L^{indirect}\partial\omega \end{cases} \tag{3.2}$$

The indirect component can be further defined as the sum of the contributions from

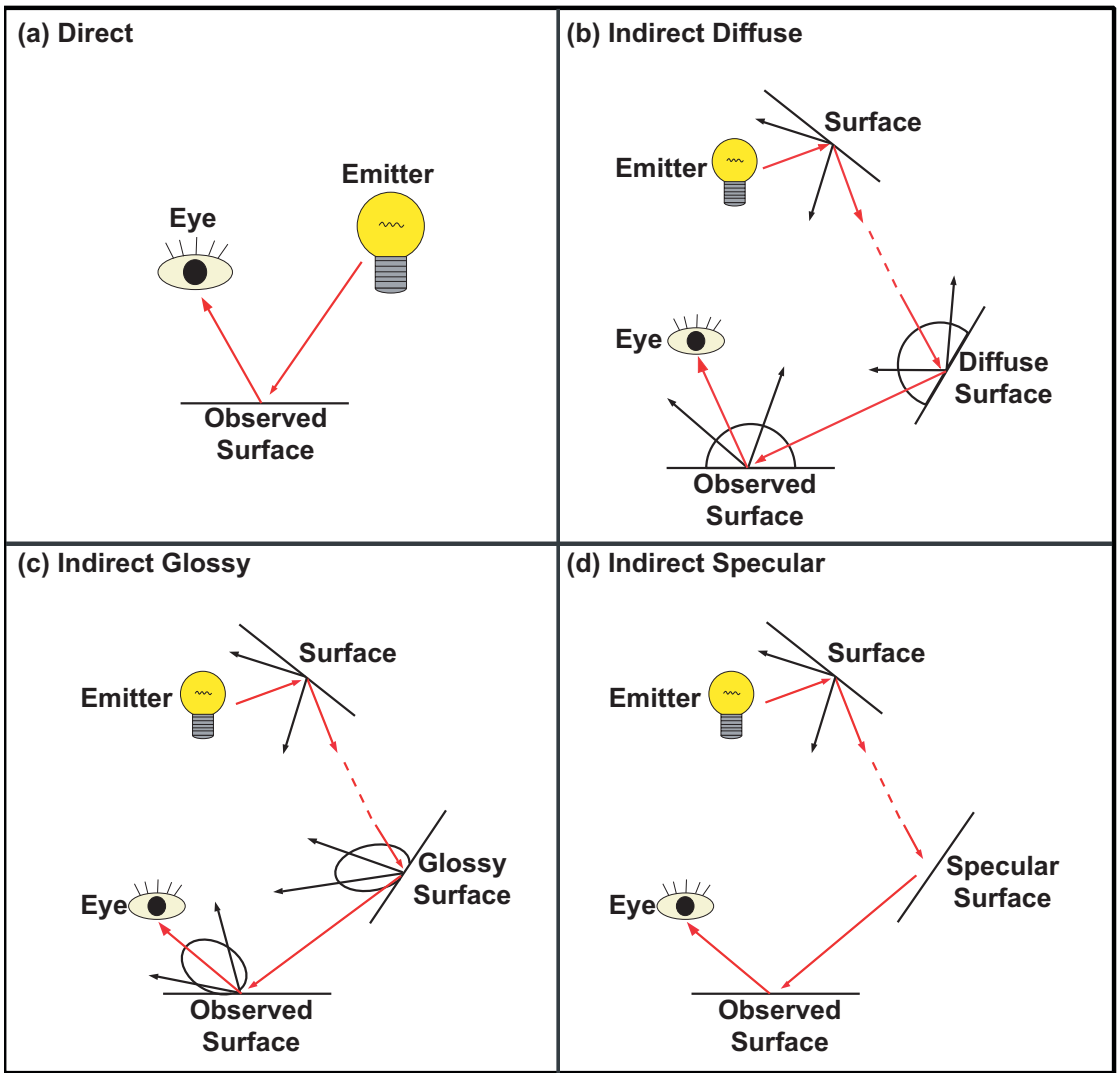


Figure 3.2: (a) Direct and indirect (b) diffuse, (c) glossy, and (d) specular illumination

three pure indirect transport paths (indirect diffuse, LD+E, indirect glossy, LG+E, and indirect specular, LS+E) and a fourth set of hybrid paths that account for the interactions between the pure component paths. If the hybrid path interactions are negligible, the overall expression for the indirect component can be approximated as:

$$\int_{\Omega} L^{indirect} \partial\omega \approx \begin{cases} \int_{\Omega}^{indirect} L_{diffuse} \partial\omega & + \\ \int_{\Omega}^{indirect} L_{glossy} \partial\omega & + \\ \int_{\Omega}^{indirect} L_{specular} \partial\omega & \end{cases} \tag{3.3}$$

Figure 3.2 illustrates the (a) direct, (b) indirect diffuse, (c) indirect glossy, and (d) indirect specular illumination components.

## 3.2 Perceptual illumination components

The key idea explored in this thesis is the observation that the different direct and indirect illumination components are probably not of equal importance with respect to their contributions to the visual quality of global illumination renderings. Speculation is often made about the relative importance of the different components, but to our knowledge, there have not been any attempts to formalize these assertions, to quantify the perceptual importances of the direct and indirect components in global illumination rendering, or leverage such variation in importance when rendering.

Rendering by components has another potential benefit: increases in coherency resulting from individual consideration of indirect reflection components may allow more efficient rendering techniques which converge more quickly to be developed. While higher sampling frequencies are necessary for capturing view dependent specular and glossy effects, indirect diffuse illumination can successfully be evaluated at far lower sampling frequencies and combined using interpolation techniques in order to fill in unsampled regions [WH92]. This stems from the observation that spatial changes in indirect diffuse illumination are gradual. By concentrating sampling on a subset of image pixels, the resulting accuracy for such pixels will be higher than that produced when spreading out the same total number of samples evenly across the entire image plane. Furthermore, interpolation techniques which then must be applied to fill in unsampled regions will produce results free of the noise that plagues naive sampling techniques.

Similar techniques can be developed to leverage path coherency when rendering in-

direct glossy and specular components. For example, one approach to approximating indirect glossy illumination might build on widely accepted importance sampling techniques. Since an indirect glossy approximation technique can ignore all indirect diffuse reflections, sampling will concentrate closer to the primary reflection direction, resulting in reduced noise and faster convergence.

Unfortunately, the rendering by components framework we explore has limitations. Unlike traditional global illumination solutions, the framework as described ignores all hybrid path contributions. Caustics (LS+DE) arise when one or more specular surfaces result in the focusing of light onto a diffuse surface before it is reflected toward an observer. Since such paths include both diffuse and specular reflections, they will be excluded by any pure indirect reflection component. Similarly, specular reflections of surfaces receiving only indirect diffuse illumination will not be captured (LD+SE) by any pure indirect reflection component. This situation arises when an area of a scene which receives only indirect diffuse illumination is reflected by a mirror or glass surface a night. If both the area of the room in question and its reflection are located on the image plane simultaneously, a visible discontinuity will arise. Fortunately, the generic rendering by components framework does not limit us from including such exotic paths during image generation. Separate indirect reflection components for each of these sets of paths could be computed and composited with pure indirect components if desired.

The following chapter describes a series of psychophysical experiments we conducted in order to measure and model the effects of the different illumination components on the perceived quality of a set of global illumination images.

# Chapter 4

# Experiments

The goal of this thesis is to develop a new perceptual metric for efficient, high-quality, global illumination rendering that can predict the consequences for visual image quality of approximations to full global illumination simulations. Having adopted the rendering by components framework described in the previous chapter, our task is first to measure, and then to model how the different illumination components affect judgments of image quality. To accomplish this we have conducted a series of psychophysical experiments.

## 4.1 Stimuli

To measure how different illumination components affect perceived image quality, we needed to define and render a set of images to be used in the experiments. Ideally we would like these images to be representative of typical global illumination renderings so our results can be applied beyond the specific conditions of our experiments. For this reason we rejected the idea of rendering abstract environments and instead generated images from detailed models of two real scenes. Ideally a larger set of test scenes would be used during testing, but currently it is difficult to obtain models with realistic material

Figure 4.1: Office and kitchen scenes

properties. Therefore, creation of both models was necessary for this research. Example images are shown in Figure 4.1. Light choice and properties were chosen to best reflect actual lighting in these environments. Two area lights were used in the office environment, while the more than 70 small lights in the kitchen environment were represented with point lights. Ward [War92] and Phong [Bli77] materials were chosen and tuned to best approximate observations of real world surface reflectance. We believe together the office and kitchen scenes span a significant portion of the range of typical interior environments we encounter in the world, with the office scene having relatively uniform illumination and mostly matte surfaces, and the kitchen scene having more dramatic variations in illumination and a larger proportion of glossy and specular surfaces.

To construct test image sets for our experiments, we rendered six views of each scene using a custom component-based Monte Carlo path tracer. Figures 4.2 and 4.3 show these views, which were chosen to be representative of viewpoints that a real observer might occupy, and to show a range of different objects and materials (i.e. avoiding views of blank walls, ceilings, desktops, etc.).

For each view, we generated separate renderings of the direct and indirect diffuse, glossy, and specular illumination components. We created eight composite images of each view by combining all permutations of the three indirect components with the baseline direct component. In addition, for each view we generated a ”gold standard” full global illumination rendering, so in total there were nine images in each of 12 test sets (6 viewpoints per scene x 2 scenes). Figure 4.4 shows the set of composite images for one view of the kitchen scene.

Due to limitations physical display size and any particular software interface would impose on the ranking task, we conducted experiments using printed images. Each image was rendered at $512 \times 512$ pixels and printed at 4”$\times$4” using a Kodak XLS 8600 dye sublimation printer. Each image was then mounted on $1/4$” foamboard to allow easy manipulation.

Since the rendering-by-components framework is based on the idea that good approximations to full global illumination renderings can be achieved by combining separate simulations of the direct and pure indirect illumination components, one issue we were interested in exploring was whether the hybrid indirect paths neglected by the framework, have a significant effect on visual image quality. This was our rationale for including the gold standard in each test set. By comparing the quality ratings given to this full global illumination rendering and the pure component “silver standard” (a composite of direct, pure indirect diffuse, glossy, and specular), we can determine the perceptual importance (or lack thereof) of the hybrid paths.

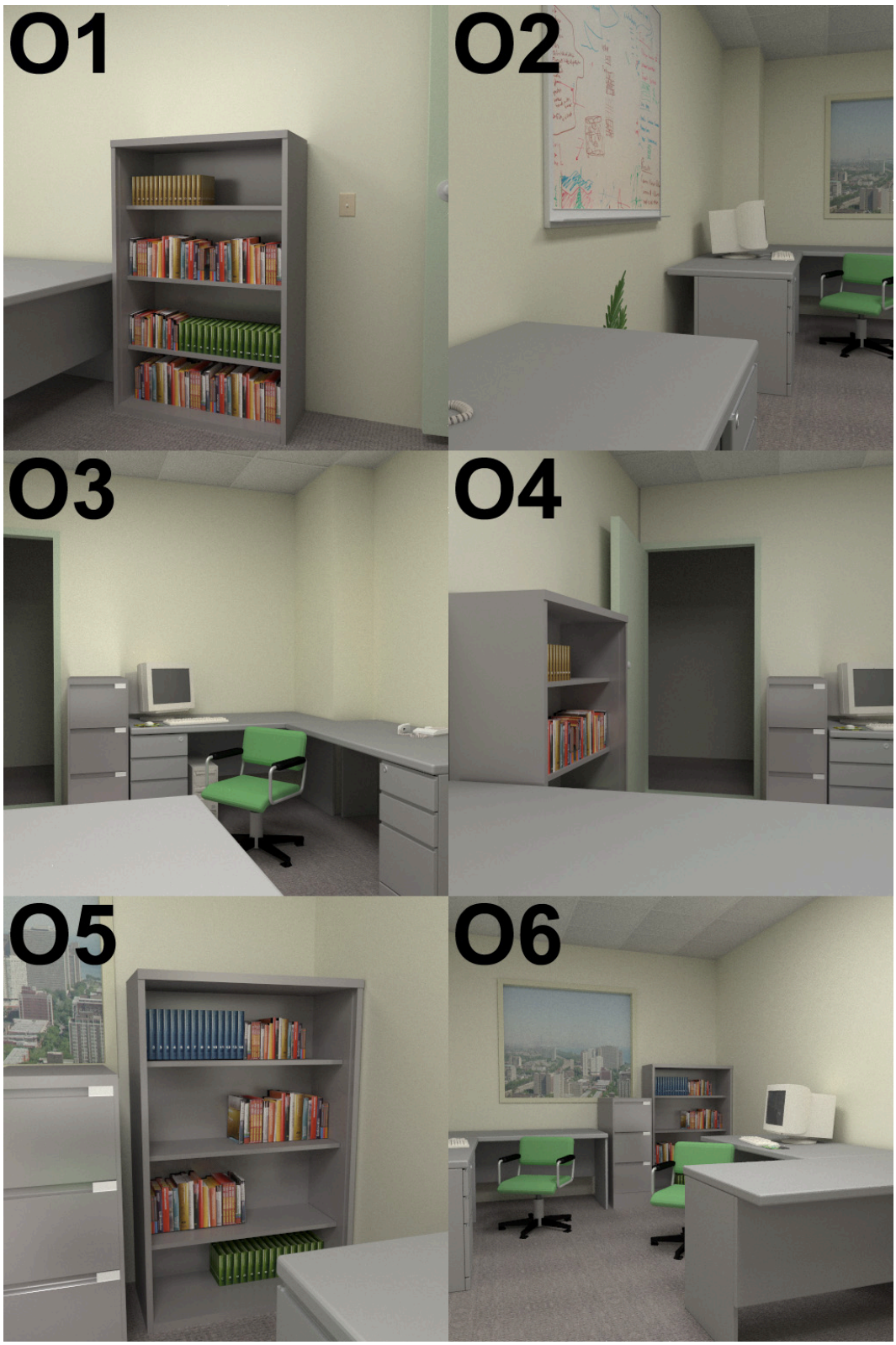


Figure 4.2: Viewpoints tested in the office scene


Figure 4.3: Viewpoints tested in the kitchen scene

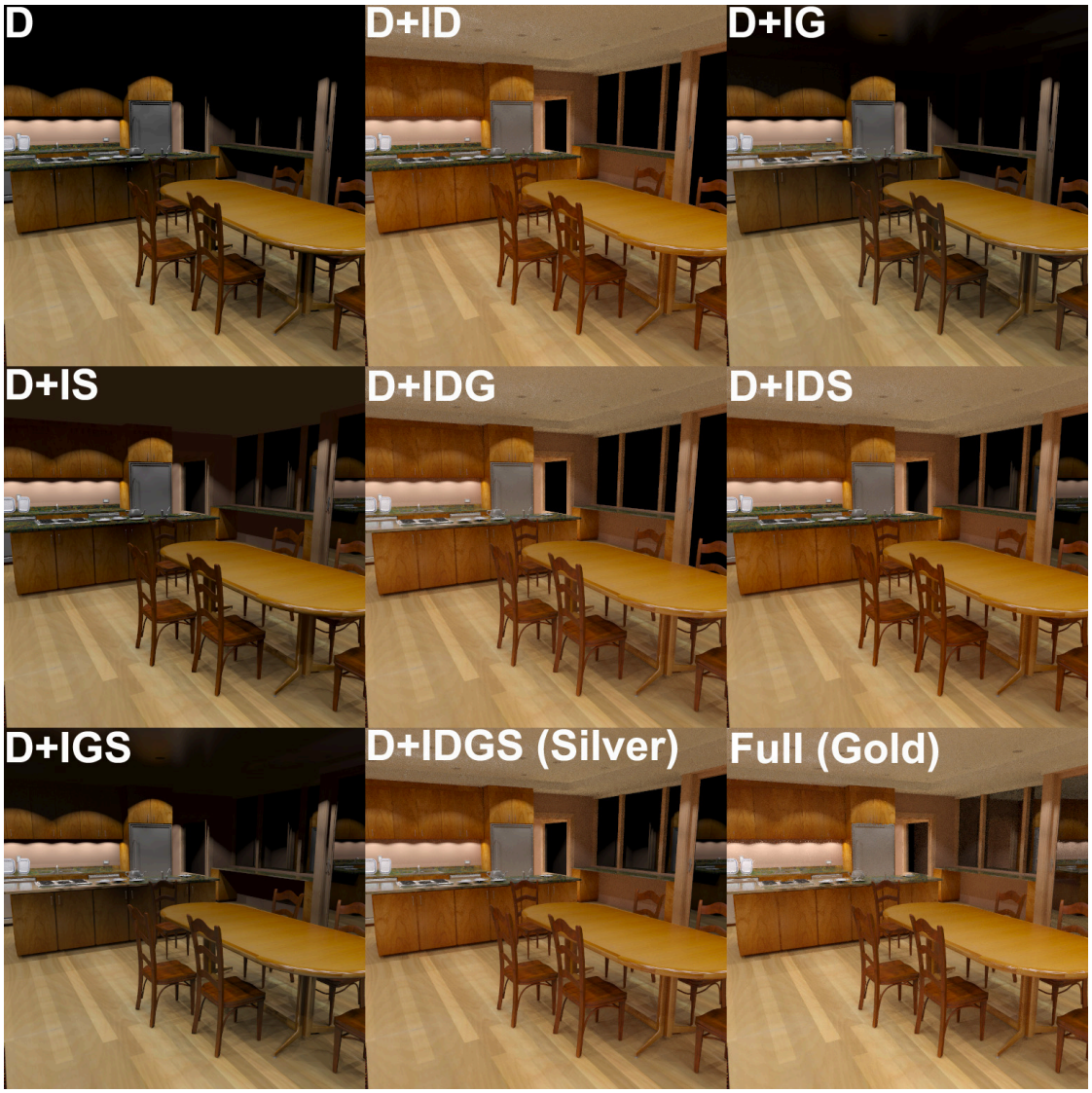


Figure 4.4: A complete set of composites plus the gold standard image for a single viewpoint of the kitchen scene.

From informal observations we expect that the contribution of the hybrid paths to visual image quality will be small in most scenarios because as is shown in Figure 4.5, there is very little visual difference between the gold and silver standard images. For radiometric purposes, by comparing pixel intensities using a difference image, we see that energy provided by hybrid paths is almost nil as is shown in Figure 4.6.

An example of light transport not captured in the silver standard is the reflection of the indirect diffusely illuminated ceiling off the windows in the kitchen scene as shown in viewpoint K2 in Figure 4.3. Another example of hybrid paths are caustics, which occur when specular surfaces focus light onto diffuse surfaces to producing visible highlights. While such effects are not captured by pure indirect components, the rendering-by-components framework we describe would allow for the inclusion of additional indirect components. From our experience caustics are rarely significant contributors in natural environments, and do not make any significant contribution in the test scenes we chose. For these reasons, we leave more accurate modeling of caustic path importance to future work.

In computing the indirect components, we needed to set a cutoff for path depth. Too severe a cutoff could result in large errors in radiance estimates, while too lax a cutoff would be inefficient. We chose to set the cutoff at four bounces. We found that images with greater path depths were visually indistinguishable from the four-bounce images. Figure 4.7 illustrates this point. Additional illumination resulting from inclusion of light transport involving four bounces is barely visible in the rightmost image when compared to the previous image.

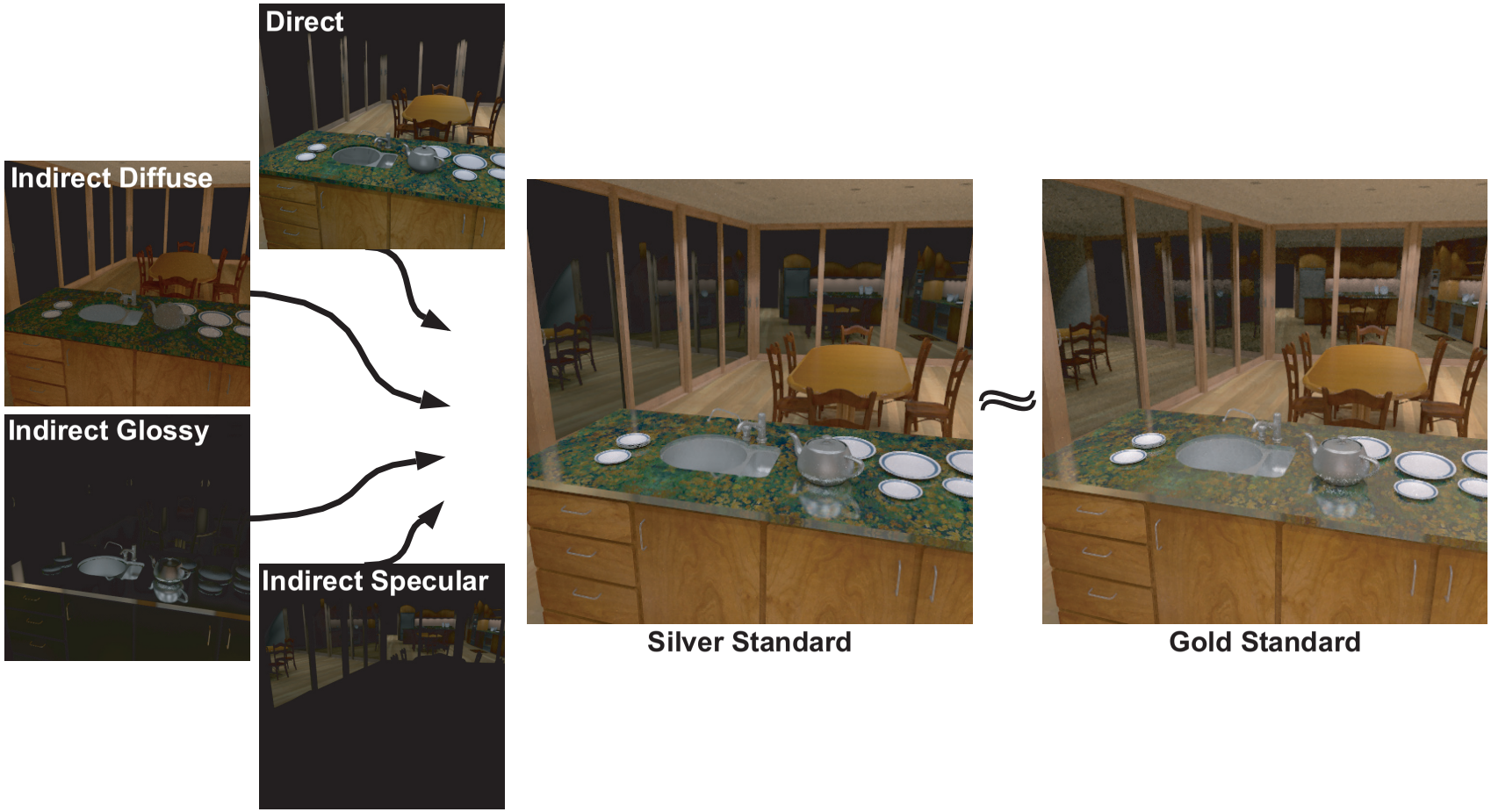


Figure 4.5: Direct, indirect diffuse, indirect glossy, and indirect specular illumination combine to form a silver standard, roughly equivalent to a full global illumination solution or gold standard.


Figure 4.6: Comparing the silver and gold standards using a difference image.

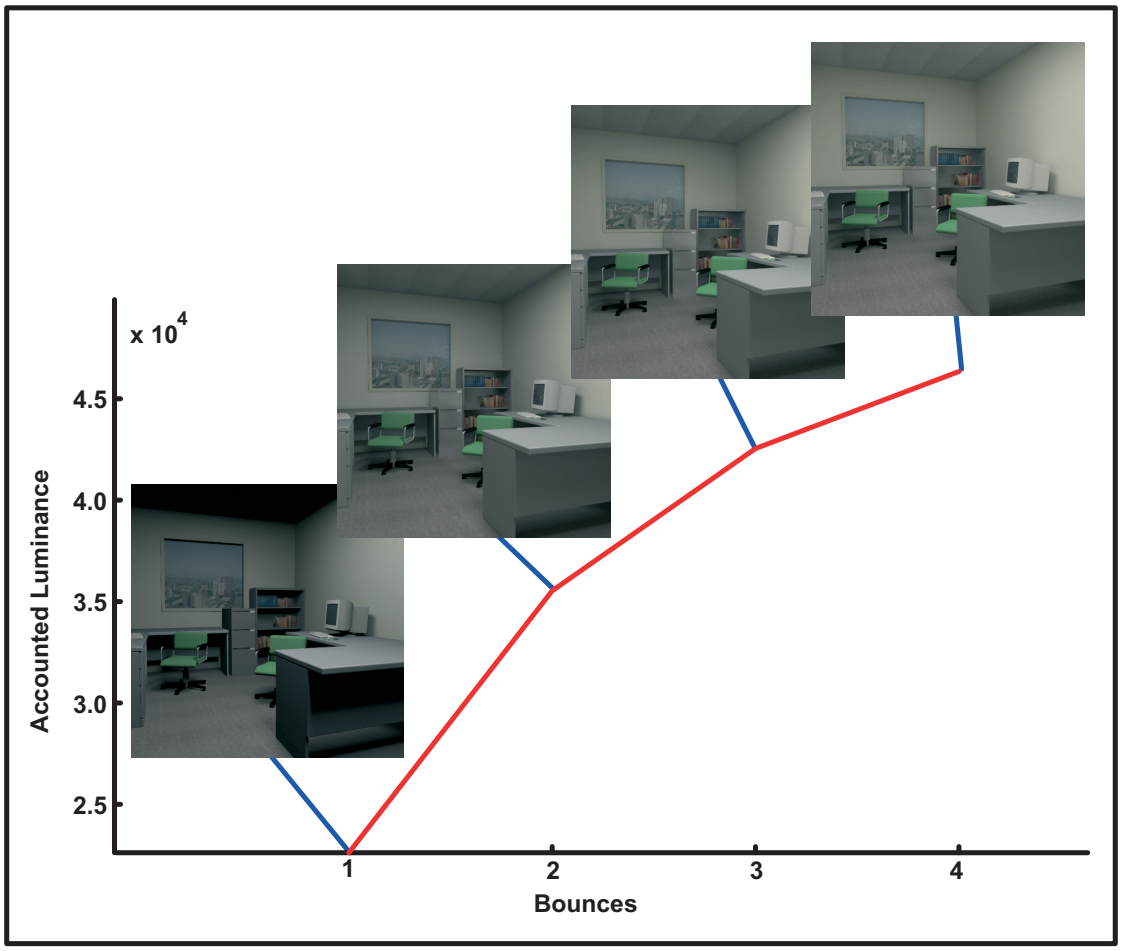


Figure 4.7: Accounted luminance contributions by successive bounces exhibit diminishing returns. Additional luminance resulting from including a fourth bounce is barely detectable.

## 4.2 Equalizing contrast and hue

One consequence of the rendering-by-components and compositing approach, is that including or excluding components from a composite can introduce contrast and hue differences with respect to the other composites in a set. Composites which include fewer indirect components have darker darks or higher contrast, while composites which

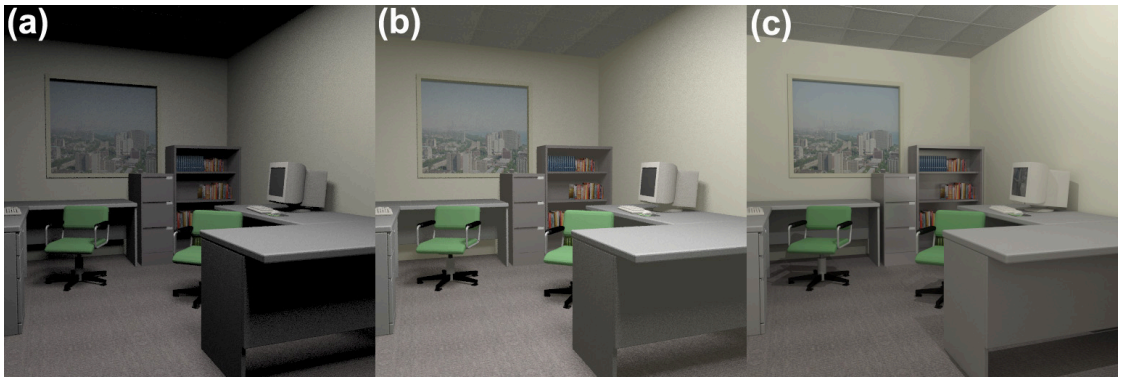


Figure 4.8: (a) Direct illumination image. (b) Contrast equalized version of (a). (c) Full global illumination solution. Notice the contrast differences between (a) and (c) and the similarity of (b) and (c).

include more indirect components often appear warmer. Figure 4.8 illustrates this problem. Image (a) only contains direct illumination. Notice that image contrast is greater than in (c), the full global illumination solution. Previous research [Fai98] has shown contrast discrepancies can bias participant responses. In order to prevent such bias, we developed a contrast equalization technique which minimized this artifact.

In any set of composites, the full global illumination solution (gold standard) has the lowest contrast. This results from dark shadows which are filled with low indirect illumination. In composites which contain subsets of indirect illumination, these regions are darker or completely black in the case of composites only containing direct illumination, and when viewed appear to have higher image contrast. By adding an "ambient" image to the other composites in the set, it is possible to minimize contrast differences. To accomplish this we first calculated an albedo image that represented the diffuse colors of the surfaces visible from that viewpoint. For each composite image + albedo image combination, we used a binary search method to find a scaling factor for the albedo image that minimized contrast differences with respect to the full global

illumination solution.

Image contrasts were calculated using a scale-dependent local contrast measure that [Pel91] has shown is well correlated with perceived contrast. First, we derived a luminance image from RGB values by the technique described in [Poy96], where luminace (L) at a pixel is defined as:

$$L = 0.2125 \cdot R + 0.7154 \cdot G + 0.0721 \cdot B \tag{4.1}$$

where R, G, and B are the intensity of the red, green, and blue channels. Next we constructed an image pyramid and calculated pixel contrast $C_{i,j}$ at different spatial scales using the equation:

$$C_{i,j} = \frac{L'_{i,j} - L_{i,j}}{L'_{i,j}} \tag{4.2}$$

where $L_{i,j}$ is the luminance of pixel $i$ at pyramid level $j$ and $L'_{i,j}$ is the luminance of the pixel in a bi-linearly upsampled version of the next highest pyramid level. To combine the contrast measures at different spatial scales, it is necessary to weight the contrasts at each level by the factor $W_j = \frac{B_w \cdot B_h}{C_w \cdot C_h}$ where $B_w$ and $B_h$ are the image dimensions at the base level and $C_w$ and $C_h$ are the dimensions at the current mip-map level $j$. Thus a summary contrast measure of the image can be defined as:

$$C = \sum_{j=0}^{n} \sum_{i=0}^{m} W_j \cdot C_{i,j} \tag{4.3}$$

where $m$ is the number of samples in level j and $n$ is the number of pyramid levels. The contrast equalization process works as follows. First, the gold standard image contrast is computed. Next, each composite is contrast equalized against the gold standard by adding a scaled version of an ambient image such that both images have equal image contrast. The albedo scale factor is determined using a standard binary search technique.

The effectiveness of this contrast equalization technique can be seen in Figure 4.8 by comparing image (b), the contrast equalized version of (a) with image (c), the full

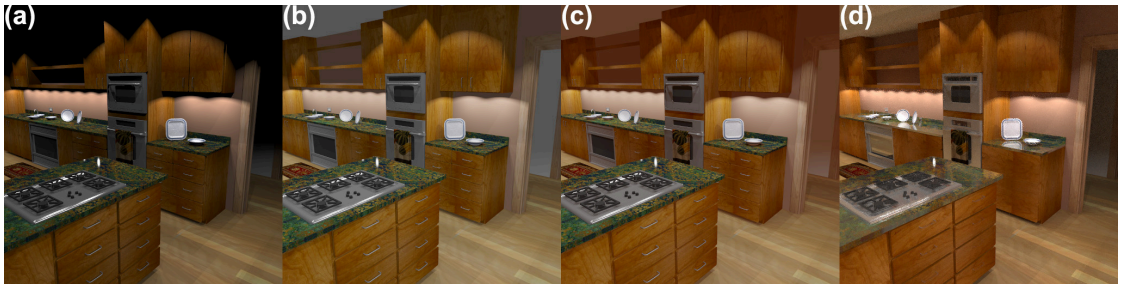


Figure 4.9: (a) Direct illumination, (b) contrast equalized version, (c) contrast equalized using hue shifted albedo image, and (d), full global illumination solution. Notice the hue differences between (a,b) and (d), and the similarity of (c) and (d).

global illumination solution. One area of future exploration would be to weight the contrast at each mip-map level using the contrast sensitivity function. This would likely enable more accurate contrast equalization resulting from more accurate simulation of the human visual systems frequency-dependent contrast sensitivity. The current contrast equalization technique is highly sensitive to contrast at the lowest mip-map level, and as a result has trouble equalizing contrast in some circumstances. Knowledge of the viewer distance to and the print size of the resulting experimental images is required before applying such frequency-dependent weighting techniques.

Another side effect of the compositing process was a hue shift between images containing different combinations of indirect illumination components. To minimize this artifact we hue shifted the albedo image by scaling it by the average hue of a composite of the three pure indirect components. This was done before contrast equalization so any luminance changes in the albedo image due to hue shifting would be compensated for during calculation of the albedo scale factors. Figure 4.9 shows an example of this hue shift artifact and how our procedure minimizes this artifact while equalizing contrast with respect to the full global illumination solution.

## 4.3 Procedure

To measure the relationships between the presence of different illumination components and visual image quality we ran a series of ranking experiments [Gui54]. The procedure is illustrated in Figure 4.10. During each trial, a subject was given the set of composite images for one viewpoint. The subject was then asked to place the images in order from lowest to highest by perceived quality. Subjects were informed that image differences might be subtle and to relax and take their time. While no strict sorting method was required, a simple divide and conquer method was suggested. Subjects were encouraged to pickup and move around the images as much as they liked. Before running through actual trials, subjects were run through three practice trials consisting of randomly selected smaller sets of images, allowing them to discover the differences across images and decide what their effect on image quality meant to them. Throughout each experiment, the subjects were asked to elaborate on what visual discrepancies they noticed and the effects each had on the ranking task. We randomized the order in which each subject received the sets, and the initial ordering of images within each set ranked. The "office" and "kitchen" ranking experiments were conducted in two sessions on different

Figure 4.10: Ranking a set of images

days under equal lighting conditions. Ten subjects participated in the experiments. Both expert (computer graphics graduate students) and non-expert (university graduates and undergraduates) participated. All were naive to the design and goals of the experiments, and all had normal or corrected to normal vision.

### 4.3.1 Results and Preliminary Analysis

The data generated by the experiments consisted of the rank orderings produced by each subject for each image set. Rank data is strictly ordinal by nature which presents significant limitations in terms of quantitative analysis and modeling [Gui54]. For example, while average rankings may indicate that one image is consistently perceived to be of higher quality than another, rankings alone cannot specify *how much* higher in quality one image is than another. To place the images on an interval scale that allows us to quantify the contributions to quality of the different illumination components, we applied Thurstonian scaling techniques from visual psychophysics [Tor58].

In Thurstonian scaling, the variance in the rank position given to an image by different subjects is used to derive rank distributions for each image. The overlap in these distributions is taken as a measure of the psychological distance between the images with respect to visual quality. By determining these parameters for all the images in a set, a perceptually-linear interval scale that indicates the relative differences in quality between the images can be derived.

Using the technique described above, we derived perceived quality scales for each of the 12 image sets (6 viewpoints per scene $\times$ 2 scenes). The scales are summarized in Figure 4.11. There are several preliminary observations that we can make about this data.

- First, across all the viewpoints there is a fairly consistent ordering of the different

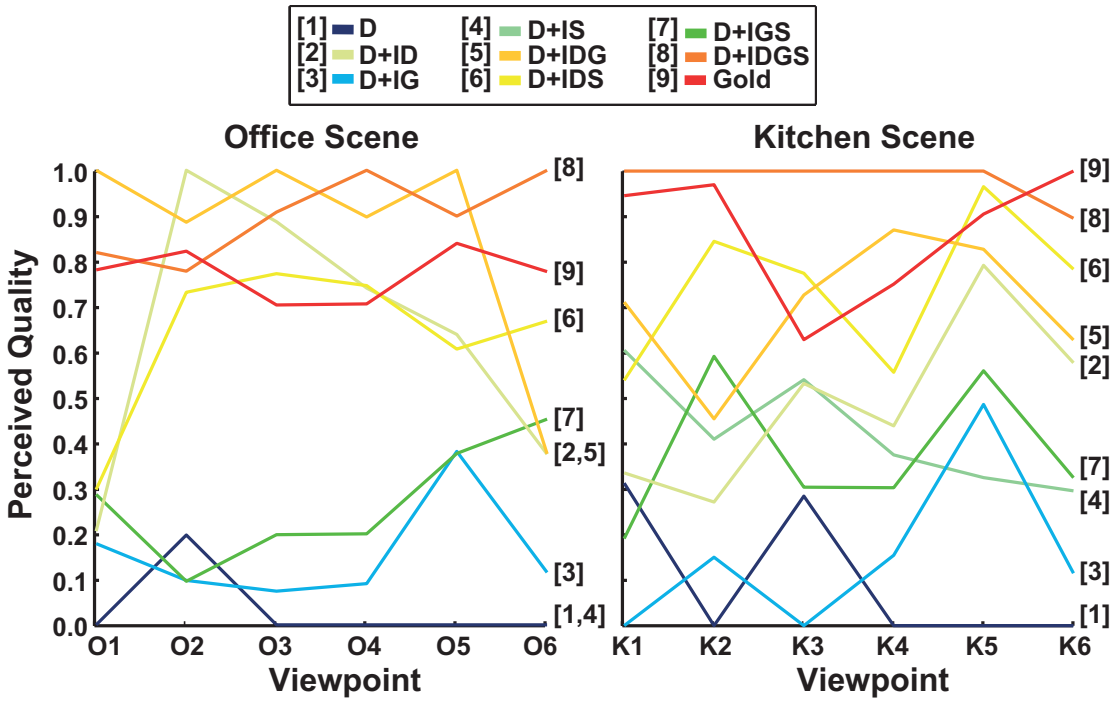


Figure 4.11: Experimental results: quality scores for the different composites. While significant variation in quality across viewpoint is evident, general trends in the ordering of the composites can be seen.

composite images, with direct-only images judged to be lowest in quality, and the silver and gold standard images judged to be highest. Gold standard images are often assigned slightly lower quality scores than the highest ranking composites. We conjecture this is because indirect hybrid lighting masks image detail provided by other lighting components.

- Within the broad middle range, the presence of the indirect diffuse component appears to be an important factor, with images that include indirect diffuse generally being ranked higher than those that exclude it. Additionally, the indirect glossy and specular components appear to have smaller modulating effects within these larger trends. It is only in views where specular surfaces such as large windows

take up a large portion of the image plane (e.g. viewpoint K2 in the kitchen as shown in Figure 4.3) that indirect specular importance exceeds that of indirect diffuse.

- Finally, there are also clearly significant variations in the perceptual importances of the different components across viewpoints and scenes that will need to be accounted for by our metric.

## 4.4 Determining the perceptual importances of the illumination components

The scaling procedure we applied in the previous section allowed us to calculate perceived quality scores for the composite images in each set. The next step toward our goal of developing a metric that can predict these scores, is to relate them to the presence or absence of the different indirect illumination components, and derive measures of the perceptual importance of each component with respect to perceived quality.

We do this by performing a linear regression on the quality scores in which the score (Q) is modeled as the linear combination of the components. This model takes the form:

$$Q_i = a_i + a_{i,d} \cdot ID + a_{i,g} \cdot IG + a_{i,s} \cdot IS \tag{4.4}$$

where *ID*, *IG*, and *IS* are binary variables that indicate the inclusion or exclusion of the indirect components, and $a_{i,d}$, $a_{i,g}$, and $a_{i,s}$ are weights returned by the regression that quantify the perceptual importance of each component in viewpoint *i*.

We performed separate regressions for each viewpoint in each scene. Figure 4.12 a and b summarize the perceptual importance factors returned by the regressions for the office and kitchen scenes. We can observe several trends in this data.

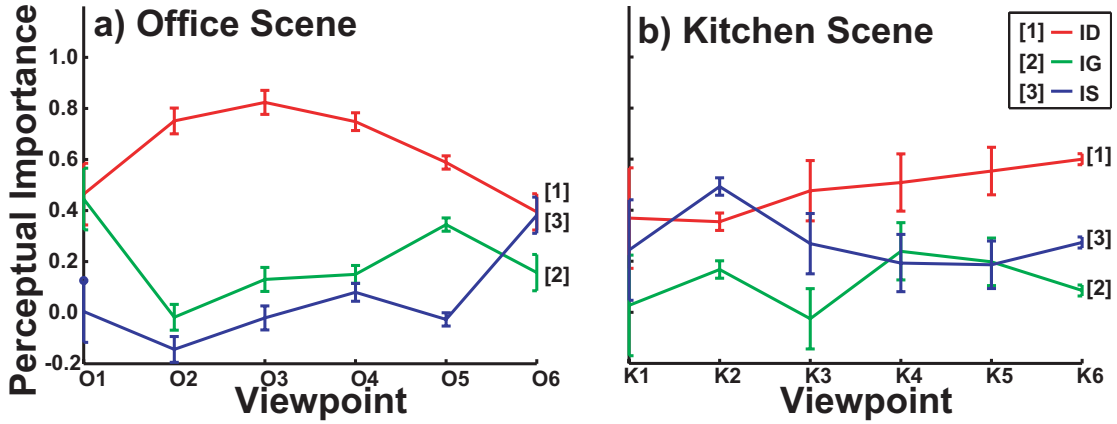


Figure 4.12: Computed perceptual importances for the office and kitchen scenes. Note that indirect diffuse generally has the largest effect on image quality. It is clear that the considerable variation across viewpoints must be taken into account by any perceptual metric.

- First, in all but one viewpoint, the indirect diffuse component (*ID*) has the highest perceptual importance. This is consistent with informal observations and comments made by the subjects that they did not like images containing large black regions (which would be the case for shadowed areas).

- Next, the regression showed that the indirect glossy (*IG*) and indirect specular (*IS*) components were of lower and roughly equal perceptual importance, with some variation across viewpoints. This result is also consistent with observations and subject reports that the absence of reflections on surfaces that should have them (e.g. windows, monitors, etc.), is disturbing and reduces the visual quality of the image, but at the same time, surface reflections that mask underlying detail (such as the reflections on the marble countertop in the kitchen scene), are also sometimes judged to reduce image quality.

- Finally, a t-test was conducted which compared perceptual quality scores of silver

and gold composites. A t-test determines whether the mean of of two groups are statistically different from each other by judging the difference between their means relative to the spread of variability of their scores. The results from this test indicated the addition of the hybrid paths did not have a significant effect on perceived image quality. Thus we can conclude that under the conditions studied their perceptual importance is low. This result provides further support for the rendering-by-components framework.

At this point we have taken several important steps toward our goal of developing a perceptual metric for global illumination rendering-by-components. Through the experiments we have measured the relationships between perceived image quality and the presence or absence of the different illumination components. From our subsequent analysis we have derived values for the perceptual importance of each of the components with respect to the measured quality scores. What remains to be done is: 1) to find a method by which we can predict the perceptual importances of the components from some indicator of the physical properties of a scene, and 2) to formulate the metric so we can predict and/or specify the visual quality of a rendering by knowing the scene characteristics and the illumination components used. This work will be described in the following chapter.

# Chapter 5

# Formulating the metric

From our experiments, we have derived measures of the perceptual importances of the three pure indirect illumination components. To proceed with formulating a perceptual metric that can be used to guide the rendering process, we need indicators that specify how the physical characteristics of the scene are related to the perceptual importance of the different illumination components. While the perceptual importance of these components are ultimately a complex function of scene geometry, lighting, and material properties, we believe that it is possible to construct useful indicators of component perceptual importance based on material properties alone. Our reasoning is as follows. In a scene without specularly reflecting materials, the illumination contribution of the indirect specular path would be null, and therefore the perceptual importance of that component should be zero. Similarly, in a scene with more matte than glossy materials, the indirect diffuse component will on average have greater influence on the scene radiances than the glossy component, and therefore its influence on the final appearance of the image, and its perceptual importance, are likely to be greater. Similar arguments in support of material-based indicator variables can be constructed for the other illumination components. Undoubtedly one can create scenes where these principles will break

down, but we believe that this approach will be useful for a wide range of scenes.

One distinct advantage of employing material-based indicators, is that indicator values can be calculated online during computation of the direct illumination component. Since any global illumination rendering system will have to compute the direct component anyway, this approach adds negligible overhead to the rendering process and should be of great advantage for implementing interactive systems.

## 5.1 Defining the indicators

As inputs to the perceptual metric we intend to develop, several material statistics-based indicators must be explored and chosen. Our goal is to accurately predict the perceptual importances of the indirect illumination components, while adding as little overhead as possible in order to be incorporated into an interactive system. Considering these requirements, we explored several indicator variables.

The first indicator variable, lobe counts, reports the percentage of pixels at which the intersected object's material includes a diffuse, glossy, or specular reflecting lobe. [1] A greater presence of a particular lobe may be enough to accurately predict the increased importance of a particular indirect component which include reflections off such surfaces.

Unfortunately, presence alone may not be enough to accurately predict the perceptual importance for all indirect components, especially those which include surfaces which vary greatly in absolute reflectance. In a second indicator, image reflectivities, we extended the previous indicator to take into account lobe reflectance by measuring the percentage of total image reflectance contributed by each indirect component. This

[1] See Figure 3.2 and Section 3.1 for examples of material lobes and an explanation of their relation to surface reflection.

calculation can be summarized as:

$$X_{\{d,g,s\}} = \frac{\Sigma_i^{\#pixels} \rho_{i,\{d,g,s\}}}{\Sigma_i^{\#pixels} \rho_{i,d} + \rho_{i,g} + \rho_{i,s}} \tag{5.1}$$

where $\rho_{i,j}$ is the reflectance at pixel $i$ resulting from including component $j$.

Finally, in an attempt to introduce a notion of locality, we explored a third indicator variable, pixel reflectivities, where percentage reflectance is considered on a pixel basis, and averaged across an entire image. This calculation can be summarized as:

$$X_{\{d,g,s\}} = \frac{1}{\#pixels} \cdot \Sigma_i^{\#pixels} \frac{\rho_{i,\{d,g,s\}}}{\rho_{i,d} + \rho_{i,g} + \rho_{i,s}} \tag{5.2}$$

Figure 5.1 shows the three indicator variables across different views in the office and kitchen scenes.

## 5.2 Indicator response functions

Before we can construct and test different metrics, we must define response functions for the indicators that specify how sensitive the metric is to any particular indicator. We experimented with three types of response functions. The simplest was a linear function of the form,

$$f_1(x) = b_1 + b_2 \cdot x \tag{5.3}$$

Here sensitivity to the indicator is controlled by the constant $b_2$ which is fixed for all values of an indicator.

The second response function we tested,

$$f_2(x) = b_1 - b_2^{\frac{1}{x+1}} \tag{5.4}$$

is also linear for low and moderate indicator values, but asymptotes to a ceiling to limit the influence of large values.

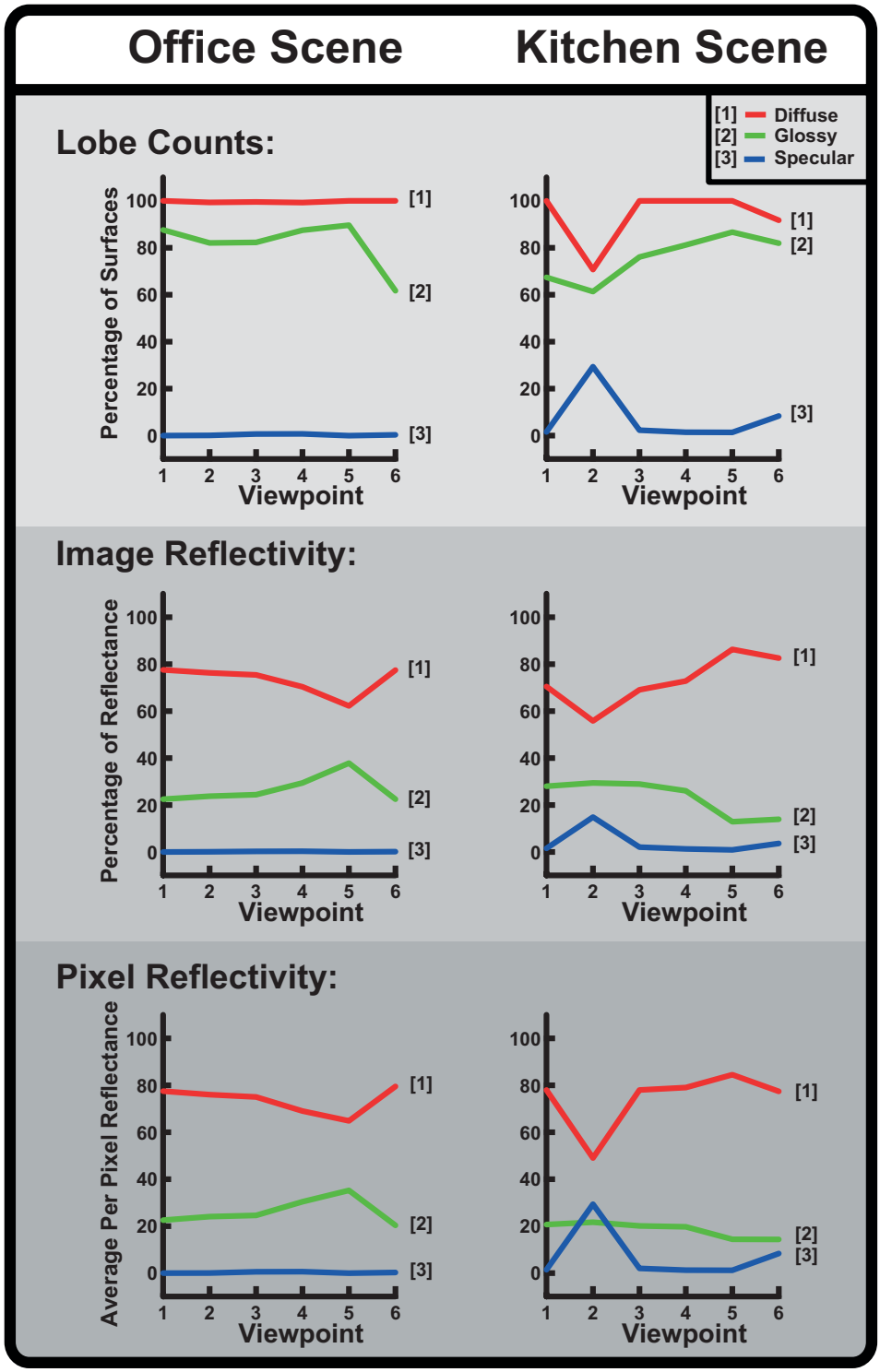


Figure 5.1: Indirect diffuse (red), glossy (green), and specular (blue) indicator response for office and kitchen scenes.

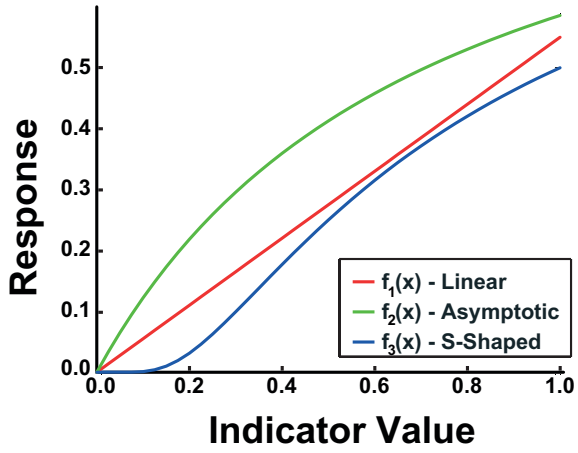


Figure 5.2: Linear (red), asymptotic (green), and s-shaped (blue) response functions.

Finally we also experimented with an s-shaped response function of the form,

$$f_3(x) = b^{\frac{-1}{x}} \text{ where we define } f_3(0) = 0 \tag{5.5}$$

In addition to limiting the response to large values, this function also limits the influence of small values that might be contaminated by sampling noise. Examples of these three response functions are illustrated in Figure 5.2.

## 5.3 Modeling perceptual importance

We are now ready to construct and test different metrics for predicting the visual quality of component-rendered global illumination images. Since our experiments have shown that the perceptual importance of the components varies across viewpoint and scene, we need to include these factors in our metrics.

Using the indicator variables and response functions defined above, we first attempt to model the perceptual importance $a_{i,j}$ of the $j$th indirect illumination component for viewpoint $i$ as:

$$\hat{a}_{i,j} = w_{1,j} + w_{d,j} \cdot f_i(x_{i,d}) + w_{g,j} \cdot f_i(x_{i,g}) + w_{s,j} \cdot f_i(x_{i,s}) \tag{5.6}$$

In this model $x_{i,d}$, $x_{i,g}$, and $x_{i,s}$ are diffuse, glossy, and specular indicator values for viewpoint $i$, $w_{1,j}$, $w_{d,j}$, $w_{g,j}$, and $w_{s,j}$ are weights determined by regressing on the perceptual importances, and the $\hat{}$ symbol indicates we are modeling an experimentally obtained value.

## 5.4 Selecting the appropriate model

Using this general model, we ran a series of regressions to find the combination of indicators and response functions that optimally predicted the perceptual importance factors we measured in the experiments. Of the three indicators we tested, the regressions showed that the second indicator, image reflectivities, provided the best prediction of the experimental results.

The lobe counts indicator worked reasonably well for the specular component, but failed to predict the importances of the diffuse and glossy components. This result is understandable. A lobe counts indicator does not take into consideration variation in surface reflectance across samples. Since there is little variation in reflectivity across specular surfaces throughout both test scenes this indicator captured enough viewpoint information necessary to accurately predict indirect specular importance. In contrast, sampled diffuse and glossy surfaces had considerably more reflectance variation, and as a result all importance models based on lobe counts indicators performed poorly. The performance of the pixel reflectivity indicator was similar to, but slightly worse than the image reflectivity indicator.

Of the three response functions we tested, the regressions also indicated that the simple linear response function yielded the best-fitting models. While overall, this model was reasonably good at predicting the perceptual importances of the different compo-

nents, its ability to predict the importance of the indirect glossy component was poor relative to the others. We believe this is because glossy materials have a wide range of appearances, from almost matte to almost mirror-like. We found that by defining separate high gloss and low gloss indicators, we were able to produce a model that was better at predicting the perceptual importance of the indirect glossy component. In order to divide glossy materials in two halves, we first needed to match Phong and Ward glossy reflections. We constructed a simple scene consisting of a glossy sphere in a checkered environment. The checkered background was chosen to produce sharp reflections with high gloss materials and blurry reflections with low gloss materials. First, a series of images were rendered by varying the Phong exponent for the sphere from 1 to 163840. Next, a similar set of images were rendered by using a similar Ward material where lobe width varied from 0.01 to 0.30. Rendered images were placed in two rows from high gloss (sharp reflections) on the left to low gloss (blurry reflections) on the right. By sliding the row of images corresponding to Phong materials across the row of images corresponding to Ward materials, we were able to align the images such that reflection contrast was most nearly matched in every column. We experimented with various divisions for high and low gloss materials and found that the best fit occurred when Ward materials with lobe widths $> 0.05$, and Phong materials with exponents $< 320$ were used to calculate the low gloss indicator ($x_{lg}$). Materials with lobe widths $\leq 0.05$ and exponents $\geq 320$ were used to calculate the high gloss indicator ($x_{hg}$) [War92, Bli77]. Figure 5.3 illustrates this calibration and division process. High contrast reflections produced by high gloss materials are shown on the left, while blurry reflections produced by low gloss materials can be found on the right. It should be noted that Phong exponent and Ward lobe width are inversely related.

An expanded form of the model which subdivides low and high gloss indicators is

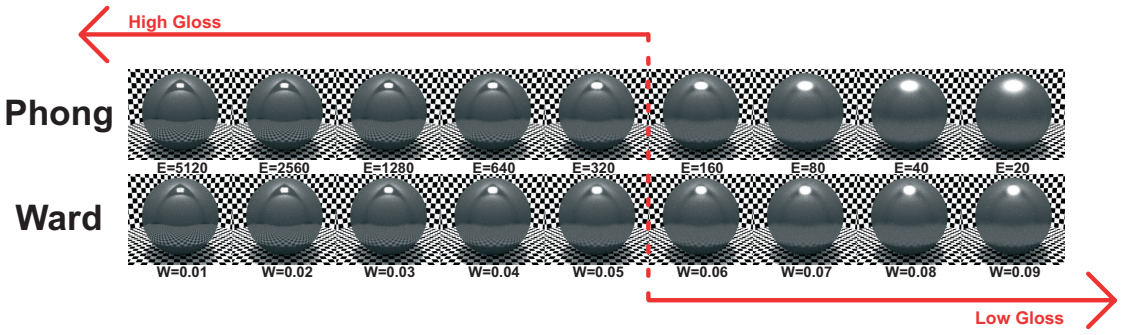


Figure 5.3: Breaking up low and high gloss

shown below. Since we have decided to use linear response functions for the indicators, the $f_i()$ terms in Equation 5.6 can be dropped, because the constants contained in these terms can be folded into the indicator weights $w_{i,j}$. Also it is not necessary to explicitly include the specular indicator variable in this formulation, because when considering image reflectivity $x_{i,d} + x_{i,g} + x_{i,s} = 1$, and therefore $x_{i,s}$ can be derived from the other two variables. This produces the simplified expression:

$$\hat{a}_{i,j} = w_{1,j} + w_{d,j} \cdot x_{i,d} + w_{hg,j} \cdot x_{i,hg} + w_{lg,j} \cdot x_{i,lg} \tag{5.7}$$

Finally, we found that we could further improve the model's ability to predict the perceptual importances of the different components by adding a factor $r^2$ that is a statistical measure of the reliability of the data we are attempting to model. Since we are modeling perceptual importances that are themselves derived from regressions on the quality scores measured in the experiments, there will be higher $r^2$ values associated with viewpoints where these regressions provided a better fit to the quality scores. The higher $r^2$ values indicate that the reliability of the perceptual importance estimates are better for those views. Adding $r^2$ allows us to desensitize the regression process to viewpoints for which importance data is less accurate. Incorporating this factor into the model produces

the expression:

$$\hat{a}_{i,j} = w_{1,j} + (1 - r^2) \cdot w_{2,j} + r^2 \cdot \left[ w_{d,j} \cdot x_{i,d} + w_{hg,j} \cdot x_{i,hg} + w_{lg,j} \cdot x_{i,lg} \right] \tag{5.8}$$

In viewpoints where initial regressions provided poor fits, the $1 - r^2$ term will approach 1 and as a result an average term $w_{1,j}$ will dominate. For viewpoints where initial regressions provided the best fits, the $r^2$ term will dominate and as a result the metric will be more sensitive to indicator values. With this model we were able to predict more than 70% of the variance in the perceptual importances of the components we measured in the experiments.

Unfortunately, these $r^2$ values would not normally be available to a rendering algorithm, since they are the product of the analysis of the data from the ranking experiments. If we want to use this model in our final metric we need to be able to estimate $r^2$ from other, more accessible data. We found that the following expression:

$$\hat{r}_i^2 = c_1 + c_d \cdot x_{i,d} + c_{hg} \cdot x_{i,hg} + c_{lg} \cdot x_{i,lg} \tag{5.9}$$

is a good estimator of the $r^2$ values we obtained from our analysis of the experimental data. In this model, $x_{i,j}$ are the same indicators used by the overall perceptual metric, while $c_1$, $c_d$, $c_{hg}$, and $c_{lg}$ are weights fit using original $r^2$ values obtained while modeling image quality.

By estimating $r^2$ values and calculating the indicator values for each viewpoint in each scene, we were able to run regressions to determine the weights ($w_{i,j}$) in Equation 5.8 that best model the perceptual importance of the indirect illumination components for each scene. These parameters are tabulated in Appendix A.

We found that the model's predictive power improved slightly when we excluded viewpoint six of the office scene. In this particular viewpoint, the contrast adjustment procedure (described in Section 4.1) added a significantly larger ambient term to the

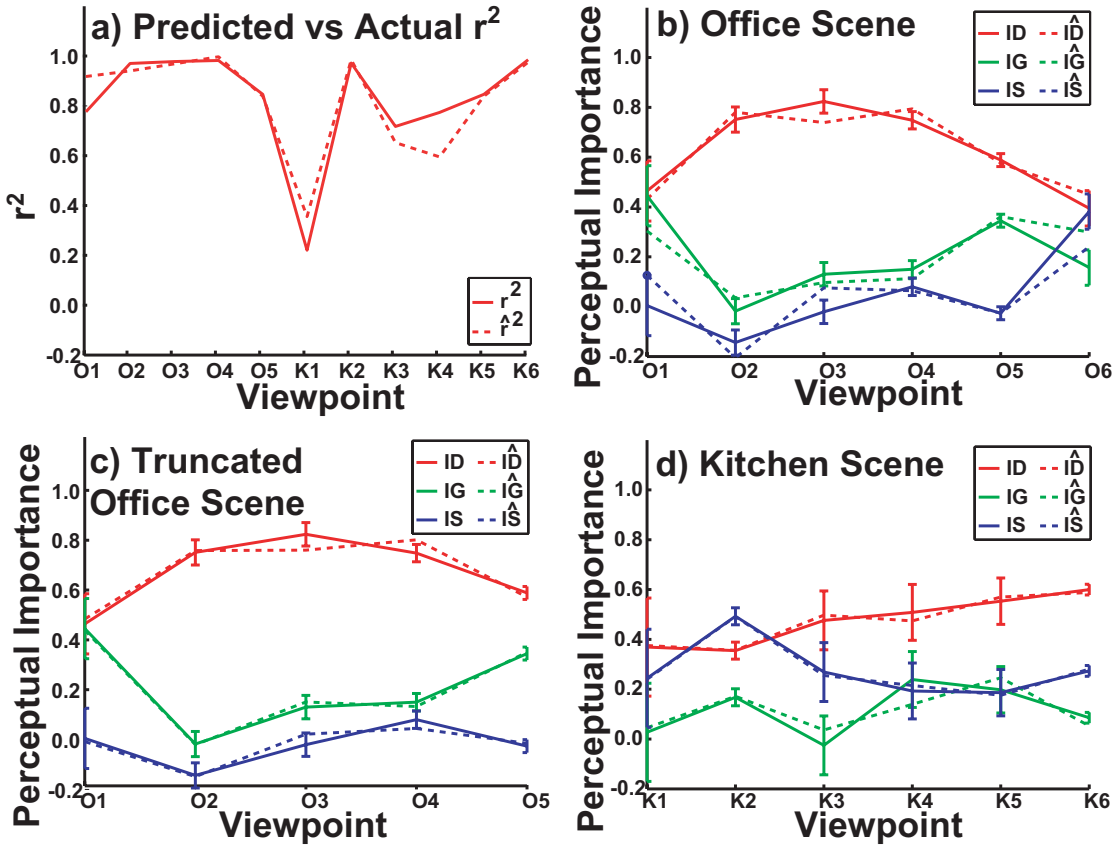


Figure 5.4: Computed and predicted $r^2$ values and perceptual importances for the office and kitchen scenes.

images than in the other cases, as a result of initial high image contrast measures. We suspect this large addition of ambient illumination distorted experimental measures of perceptual importances of the illumination components for this view, and made it an outlier in subsequent analysis and models. We believe future work which improves upon our contrast measuring technique by applying the contrast sensitivity function when weighting contrast across spatial frequencies will avoid this distortion resulting most likely from oversensitivity to contrast at higher spatial frequencies.

Figure 5.4 illustrates the predictive abilities of the models we have developed. Figure 5.4a compares observed $r^2$ values(solid line) and those predicted (dashed line) by

Equation 5.9. Figures 5.4b and 5.4c show the correspondence between the experimentally measured importances (solid lines) and the importances predicted by Equation 5.8 (dashed lines) for the office scene. In Figure 5.4c viewpoint six has been removed. Finally, Figure 5.4d shows the correspondence between the measured and predicted importances for the kitchen scene. It is clear from the close correspondence between the solid and dashed lines in these graphs, that the models are very good at predicting the perceptual importances we measured in our experiments, and therefore should perform well as the foundation of our perceptual quality metric.

### 5.4.1 Formulating the perceptual metric

Given the model for the perceptual importance of the different illumination components defined in Equation 5.8, we can now finally formulate our full perceptual quality metric as:

$$\hat{Q}_i = a_i + \hat{a}_{i,d} \cdot ID + \hat{a}_{i,g} \cdot IG + \hat{a}_{i,s} \cdot IS \tag{5.10}$$

where $\hat{Q}_i$ is the quality score predicted by the metric for an image of viewpoint $i$, (defined by some composite of illumination components), $\hat{a}_{i,j}$'s are the perceptual importances of the different indirect illumination components, and $a_i$ is a general offset approximately equal to the quality of a composite consisting only of direct illumination.

## 5.5 Using the metric

In the previous section we formulated a new perceptual metric for predicting the visual quality of component-rendered global illumination images. In this section we will demonstrate the metric's predictive abilities and illustrate how the metric could be used in an interactive rendering system.

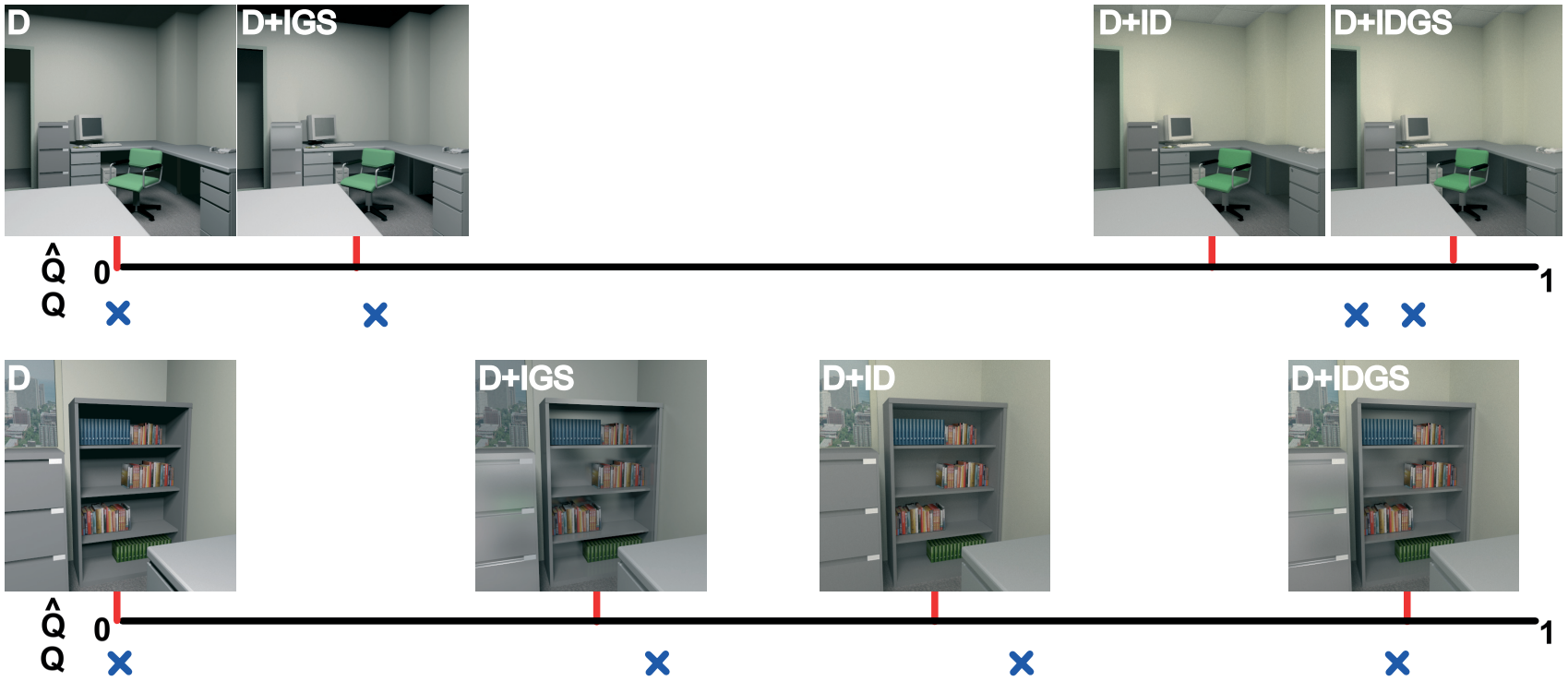


Figure 5.5: Actual versus predicted image quality for various global illumination composites of viewpoints three and five from the office scene. Not only do predicted quality values (indicated with red lines) closely match experimental results (represented by blue X's), our perceptual metric accurately models variation in perceptual importance across viewpoints.

Figure 5.5 shows two sets of component-rendered images of the office scene. Each of the images is positioned with respect to the visual quality scale $\hat{Q}$ predicted by the new perceptual metric. There are several factors to observe in this figure.

- First, is the correspondence between the quality values ($\hat{Q}$) predicted by the metric for these images, and the values measured in our experiment ($Q$), indicated by the X's below each scale. The similarity of the measured and predicted values is a confirmation that the metric is indeed capable of modeling the perceptual importances of the different illumination components and their contributions to perceived quality.

- Next, the metric's ability to quantify the quality of these images can also be confirmed by noticing the correspondence between the locations of the images along the quality scale and the similarities and differences in their visual appearances. Notice that on the upper scale the composites $D+ID$ and $D+IDGS$ placed near the high end of the scale are both similar to each other in visual quality, and are of distinctly higher quality than the composites $D$ and $D+IGS$ that the metric placed near the lower end of the scale. Similarly, the relatively equidistant positions of the images on the lower scale, accurately reflect the moderate increments in visual quality that can be observed by comparing adjacent images on the scale.

- Finally, the power of the metric to predict view-specific differences in the perceptual importance of the illumination components can be seen by comparing the central pair of images on each scale. On the upper scale, because of the material characteristics of the objects in this view, the metric correctly predicts that the indirect diffuse component makes a much greater contribution to quality than the glossy or specular components; thus composite $D+ID$ is of substantially higher

quality than composite $D+IGS$. However, for the view shown on the lower scale, where the material characteristics are substantially different, adding the the indirect diffuse component $D+ID$ does not produce as great an improvement in image quality, and conversely, adding the indirect glossy and specular components $D+IGS$ produces a relatively greater improvement than it did for the other view.

The three capabilities described above confirm that our new metric can accurately predict both the perceptual importances of different illumination components, and the resulting perceived quality values of component-rendered images.

## 5.6 A perceptually-based component renderer.

Figure 5.6 illustrates how our new metric could be used in an interactive global illumination rendering system. Assuming there are not sufficient resources to complete full global illumination rendering in real time, a user could specify whether to optimize system performance to a constant quality level, or a constant frame rate. Given these user preferences, novel viewpoints could be rendered by first gathering information about the materials visible in the current view as part of an initial direct illumination pass. Then calculating indicator values and the perceptual importance of the indirect illumination components. Lastly, dynamically allocating system resources to the computation of the different components (as shown by the pie chart on the lower right of Figure 5.6); and compositing the components for display. Taking advantage of frame-to-frame coherence in estimating the importances and allocating resources would be likely to lead to even further improvements in performance. Unfortunately dynamic allocation of compu-


Figure 5.6: A hypothetical perceptual component rendering system.

tational resources using perceptual importances is a nontrivial problem. Possible solutions will be explored in Chapter 6.

# Chapter 6

# Conclusions/Future Work

Based on our experiments we have developed a new perceptual metric for efficient, high quality, global illumination rendering. Using the rendering-by-components framework, the metric can predict the perceptual importances of the indirect illumination components, and their contributions to the visual quality of the resulting image. We have demonstrated the predictive capabilities of the metric and have shown how it could be used in a global illumination rendering system.

An important aspect of this perceptual metric compared to others that have previously been developed, is that because it is based on simple measures of scene reflectance values that can be gathered during calculation of the direct illumination component, it adds negligible overhead to the global illumination rendering process. This should make it attractive for use in interactive rendering systems.

A major goal of our research is to use such a perceptual metric to adaptively allocate resources for rendering the indirect components in an interactive setting. In order to build a system that achieves this goal it will be necessary to develop scalable approximation techniques for computing indirect illumination components and apply triage to determining the optimal resource allocation for rendering indirect illumination compo-

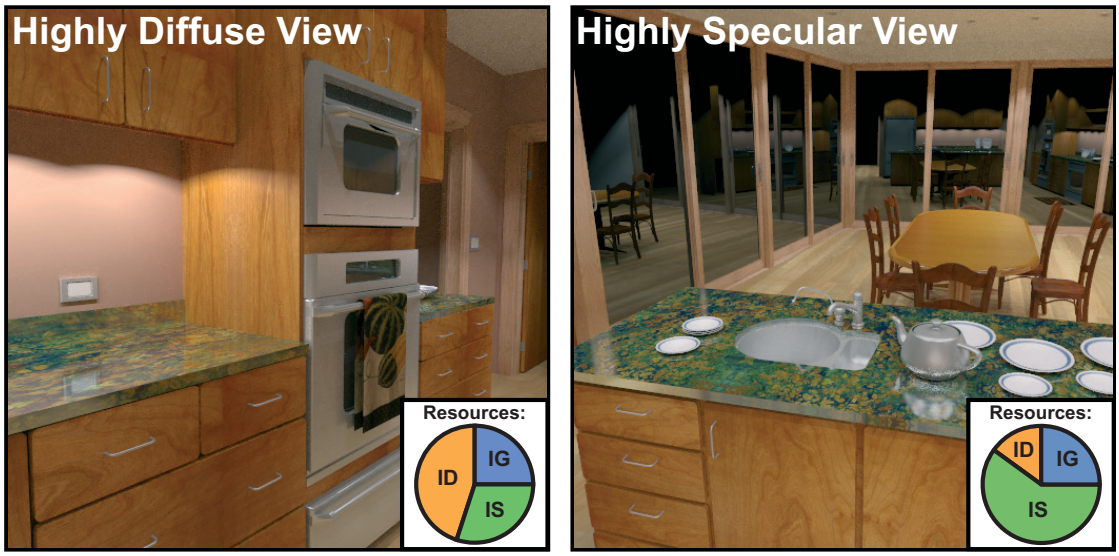


Figure 6.1: Possible optimal resource allocations for predominantly diffuse (left) and predominantly specular (right) viewpoints within the same scene. For image-based solutions, as visible material statistics change, it is likely optimal resource allocations will exhibit significant changes as well.

nents. It should be noted that resource allocations have been shown to be viewpoint and scene dependent, and will need to be computed on the fly for walkthrough and interactive systems, as illustrated in Figure 6.1.

## 6.1 Applying a Decision Theoretic

The use of decision theoretics in rendering systems are not new. Dumont et. al. [DPF03] used a decision theoretic for their texture cache management scheme in which they adaptively degraded mip-map quality in order to maintain interactive system performance. Our work differs from their approach in that we propose subdividing the rendering of indirect illumination into components and adaptively allocating resources to the rendering of each.

Our intention is to formulate a rendering-by-components framework such that it can be implemented using any number of orthogonal[1] indirect approximation techniques. Unfortunately, this requirement disallows applying the same triage approach used by Dumont et. al. in which a maximal quality solution is degraded until system constraints are met. While sampling-based indirect approximation techniques ideally converge to correct solutions over time, it is impossible to specify when solution quality is maximized due to their asymptotic convergence nature. Fortunately, three possibly valid alternative approaches to triage exist.

The simplest approach, hill climbing, is an iterative technique that can be applied by beginning without allocating any computational cycles to any of the indirect components. Additional computational cycles are iteratively allocated to the approximation technique that will most benefit from increased computational resources. This process continues until all computational cycles have been allocated. This iterative approach is illustrated in Figure 6.2.

A second approach to triage requires searching for the optimal resource allocation by transitioning between complete resource allocations using more intelligent search methods such as simulated annealing. Such techniques are more likely (or in the case of simulated annealing, guaranteed) to find the optimal resource allocation instead of getting stuck in a local maximum, as illustrated in Figure 6.3. In addition, if one begins the search process at a state close to the optimal solution, faster convergence will take place. Assuming viewing positions change slowly from frame to frame as a user explores an environment, and visible material statistics are a function of viewpoint, changes in such data can safely be assumed to be continuous. Since our perceptual metric is based on visible material statistics, it follows that deviation of optimal resource allocation from

[1]When compositing indirect components to produce global illumination images, no two indirect reflection components are allowed to capture the same energy transfer.

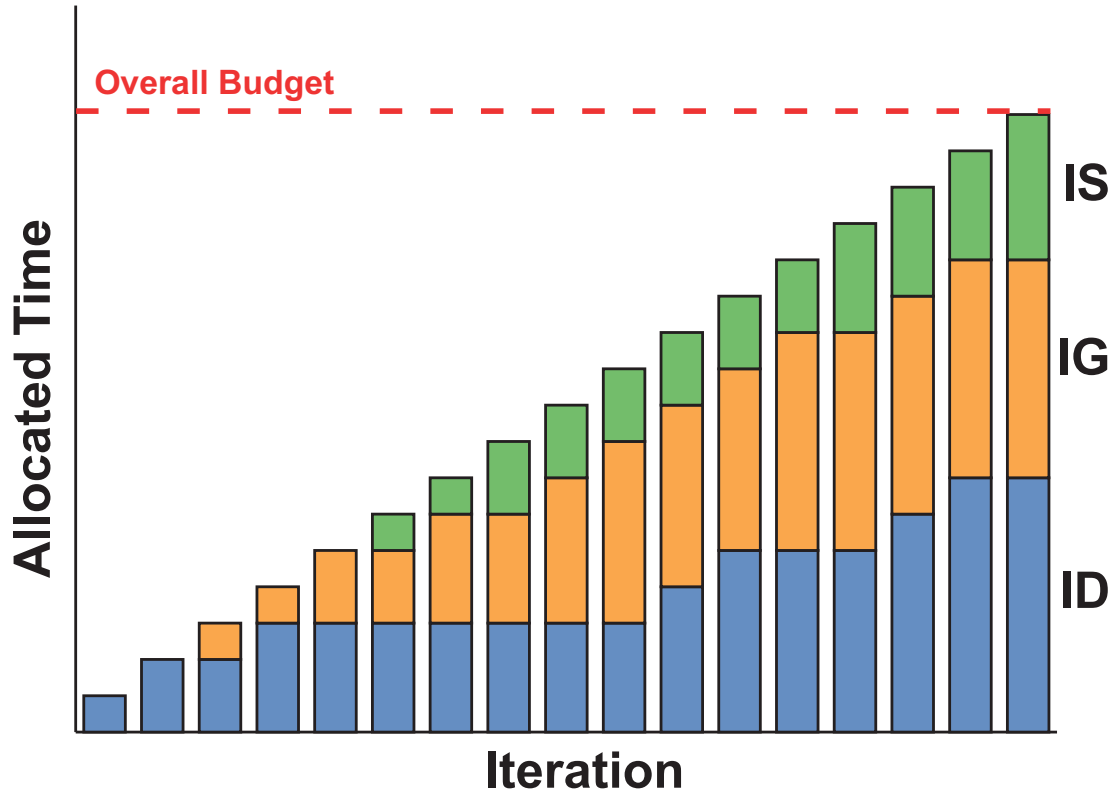


Figure 6.2: An example of an iterative hill climbing approach to resource allocation. Additional resources are allocated until an overall budget has been exhausted.

frame to frame will be small, or more simply put, the optimal resource allocation will most likely be very similar to that from the previous frame. By starting with the last known optimal resource allocation, it is likely that significantly faster convergence can be achieved.

Finally, it may be possible to solve analytically for the optimal resource allocation. It should be noted that none of these hypotheses have been tested within the scope of this thesis.

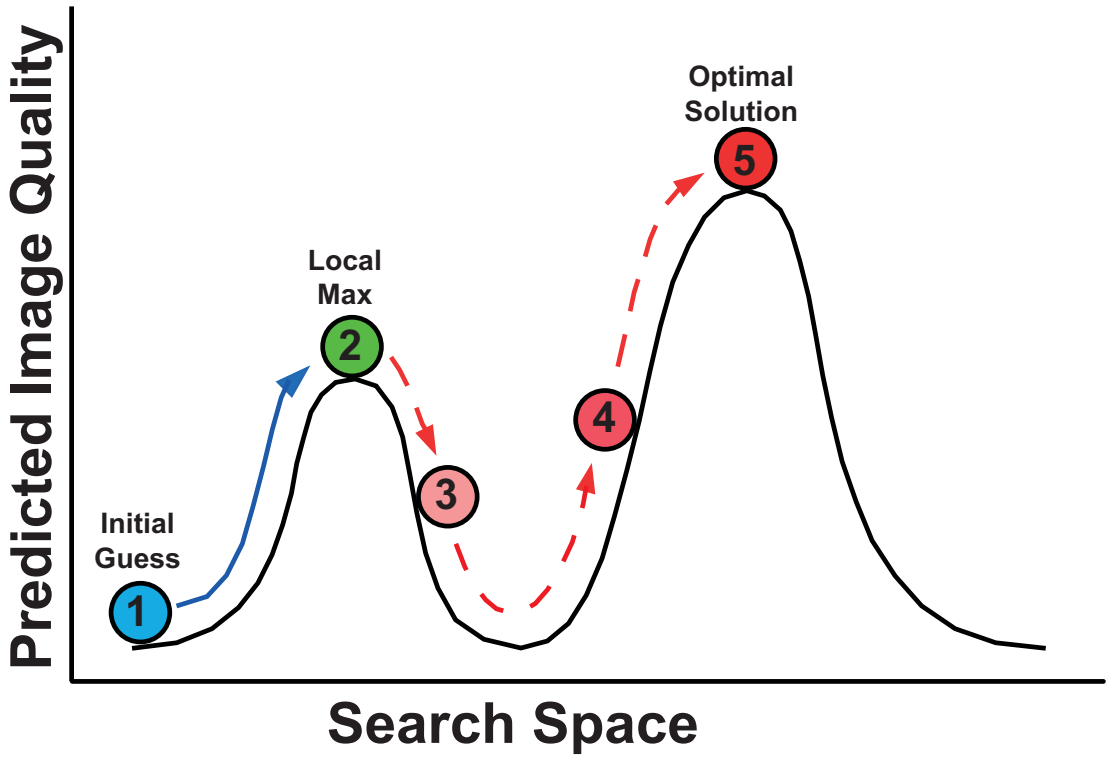


Figure 6.3: When searching for the optimal distribution of resources, hill climbing approaches can get stuck in local maximum. When applied properly, simulated annealing techniques are guaranteed to find the optimal solution.

## 6.2 Cost Quality Functions

In order to conduct triage, cost-quality functions for each approximation technique will be required. Cost-quality functions measure perceived quality of an approximated indirect component after a specified number of computational cycles have been utilized. While cost-quality curves for a particular indirect approximation technique may vary slightly across viewpoints, they are likely to retain similar shape and size. A series of psychophysical rating experiments could be run in which subjects rate the perceived quality of a series of approximations of an indirect illumination component which are computed using varying budgets. Linear interpolation, or more advanced fitting meth-

ods, can then predict the perceived quality of an approximated indirect component for a given allocation of computational resources. Figure 6.4 provides an example of a cost quality function, and how one might interpolate sampled quality data in order to predict approximate solution quality in between sampled budget allocations. We suggest future work verify the effects on image quality that result from using intermediate indirect component approximations.

In order to take into consideration approximation convergence rates in addition to predicted perceptual importance values, we suggest future research combine these measures by scaling each cost-quality curve by the perceptual importance of the particular indirect illumination component using the perceptual metric. For example, if individuals are likely to find indirect diffuse illumination significantly more important than the other indirect components, it follows that a decrease in quality of an indirect diffuse approximation will result in a larger decreases in overall image quality than those which result from equal decreases in quality of the other indirect component approximations. Our research has shown that perceptual importance is a function of viewpoint and scene, thus we suggest performing this scaling operation every frame following evaluation of the perceptual metric, before the optimal resource allocation is found. Figure 6.5 illustrates this scaling process.

## 6.3 Additional Future Work

While our research has concentrated on modeling the perceptual importances of the pure indirect diffuse, glossy, and specular illumination components, it is important to note that our rendering-by-components framework allows for future systems to approximate global illumination solutions using larger sets of indirect approximation techniques. Ad-

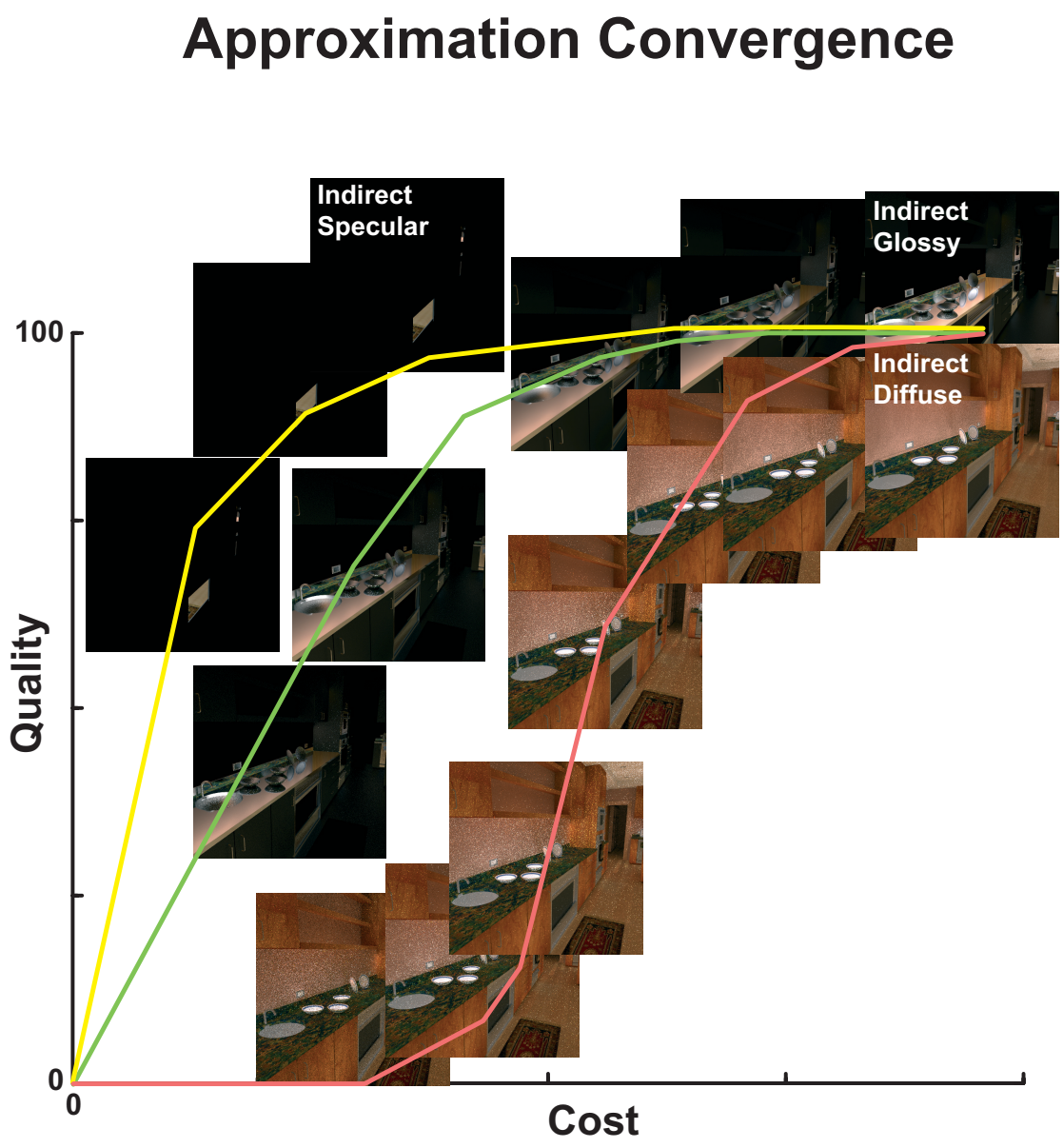


Figure 6.4: In addition to modeling perceptual importance values, it is necessary to take into consideration the quality of approximated components as they improve over time.

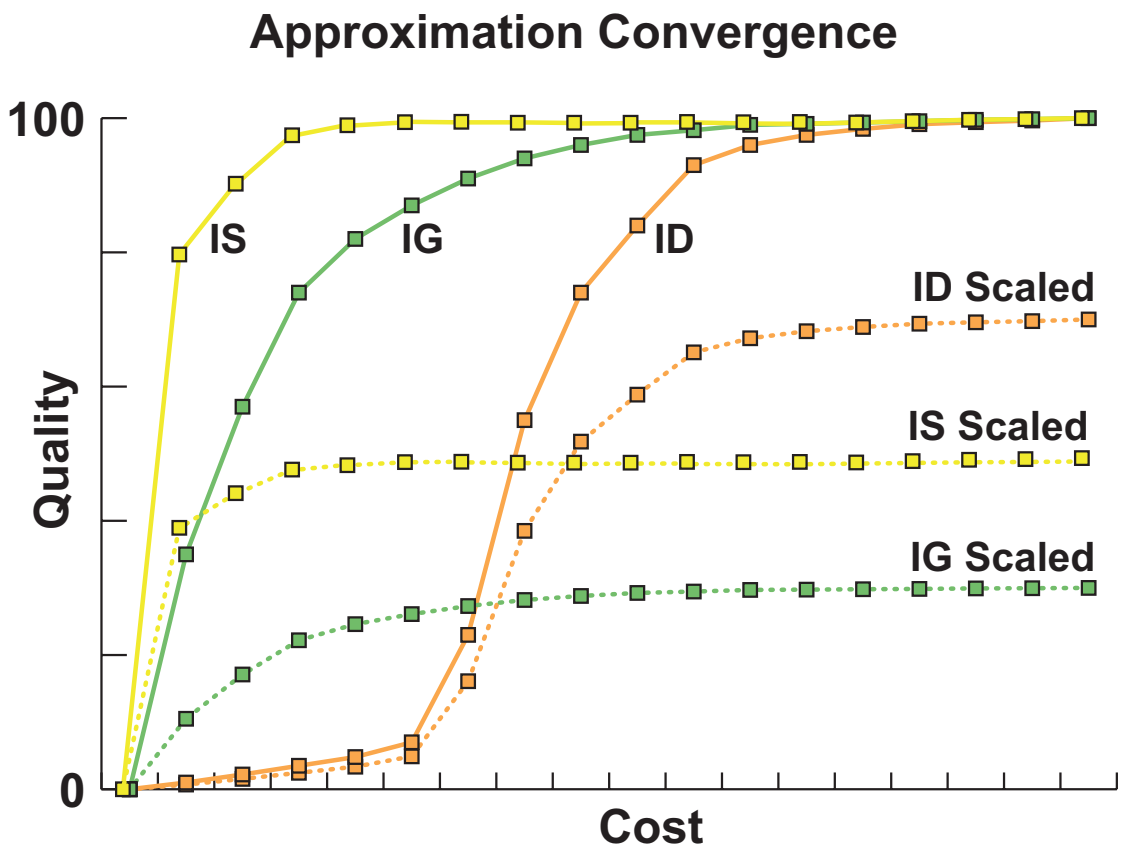


Figure 6.5: In order to incorporate perceptual importance data, cost-quality curves might be scaled before solving for the optimal resource allocation.

ditionally, while we outlined a system which must perform triage every frame, running triage asynchronously at a reduced frame rate and interpolating resource allocations for in between frames should be explored. Not only would this remove the requirement that triage be completed fast enough to facilitate interactive rates, an asynchronous approach would force changes in resource allocations to be gradual, possibly reducing any temporal artifacts that may result from changing computational resource allocations over time. Computational budgets provided by either approach must then guide approximation of indirect illumination components in the proposed rendering by components framework. The results of these approximation techniques combine with a direct illumination solu-

tion to form an optimal global illumination solution given system constraints.

In addition to developing triage and approximation techniques, a number of tasks remain before the rendering-by-components framework we envision becomes a reality. Improvements to the material-based indicators used to estimate the importance of the indirect glossy component should be explored, in addition to finding more meaningful ways to divide the wide range of surface reflectance properties that are currently defined as glossy. Previous work at Cornell's Program for Computer Graphics [PFG00] might provide a useful starting point. Unfortunately, parameters used in the current metric are viewpoint independent, but must be tuned for a particular scene. As a next step, it would be interesting to generalize the metric so it can automatically adapt to different scene characteristics. Finally, there is always much more work to be done to increase our understanding of human perception to develop more sophisticated and effective perceptual metrics for global illumination rendering and other aspects of computer graphics.

While it is clear there is much work to be done, we believe the perceptual metric and associated rendering by components framework discussed in this thesis will enable the development of an important new class of efficient global-illumination rendering systems that can intelligently allocate limited computational resources between components of the global illumination simulation to provide high quality renderings at interactive rates.

# Appendix A

# Parameters used in the metric

## Predicted $r^2$ values:

| $c_1$ | $c_d$ | $c_{hg}$ | $c_{lg}$ |
|---|---|---|---|
| 6.07 | -5.26 | -8.52 | -4.74 |

## Predicted importance factors: office scene

| | $w_1$ | $w_2$ | $w_d$ | $w_{hg}$ | $w_{lg}$ |
|---|---|---|---|---|---|
| Indirect Diffuse: | -3097.21 | 2566.22 | 3206.90 | 0 | 2911.57 |
| Indirect Glossy: | 5344.02 | -4439.80 | -5529.39 | 0 | -5026.76 |
| Indirect Specular: | 1730.64 | -1446.58 | -1789.08 | 0 | -1630.97 |

## Predicted importance factors: kitchen scene

| | $w_1$ | $w_2$ | $w_d$ | $w_{hg}$ | $w_{lg}$ |
|---|---|---|---|---|---|
| Indirect Diffuse: | -0.509 | 0.80 | 1.08 | 0.49 | 1.70 |
| Indirect Glossy: | -12.12 | 11.20 | 13.28 | 21.64 | 6.37 |
| Indirect Specular: | 5.45 | -4.97 | -5.54 | -7.66 | -3.65 |

# Bibliography


[Ahu93] Jr. Albert J. Ahumada. Computational image quality metrics: A review. *Society for Information Display International Symposium Digest of Technical Papers*, 24:305–308, 1993.

[Bli77] James F. Blinn. Models of light reflection for computer synthesized pictures. In *Proceedings of SIGGRAPH 77*, Computer Graphics Proceedings, Annual Conference Series, pages 192–198, July 1977.

[BM95] Mark R. Bolin and Gary W. Meyer. A frequency based ray tracer. In *Proceedings of SIGGRAPH 95*, Computer Graphics Proceedings, Annual Conference Series, pages 409–418, August 1995.

[BM98] Mark R. Bolin and Gary W. Meyer. A perceptually based adaptive sampling algorithm. In *Proceedings of SIGGRAPH 98*, Computer Graphics Proceedings, Annual Conference Series, pages 299–310, July 1998.

[BM99a] Mark R. Bolin and Gary W. Meyer. A perceptually based adaptive sampling algorithm. In *Proceedings of SIGGRAPH 99*, Computer Graphics Proceedings, Annual Conference Series, pages 299–310, July 1999.

[BM99b] Mark R. Bolin and Gary W. Meyer. A visual difference metric for realistic image synthesis. *SPIE Human Vision and Electronic Imaging IV*, 4(3644):106–120, July 1999.

[Dal93] Scott Daly. *The Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity*. Digital Images and Human Vision. MIT Press, 1993.

[DPF03] Reynald Dumont, Fabio Pellacini, and James A. Ferwerda. Perceptually-driven decision theory for interactive realistic rendering. *ACM Transactions on Graphics*, 22(2):152–181, April 2003.

[Fai98] Mark D. Fairchild. *Color Appearance Models*. Addison-Wesley, Reading, MA, 1998.

[FPSG97] James A. Ferwerda, Sumanta N. Pattanaik, Peter S. Shirley, and Donald P. Greenberg. A model of visual masking for computer graphics. In *Proceedings of SIGGRAPH 97*, Computer Graphics Proceedings, Annual Conference Series, pages 143–152, August 1997.

[GH97] S. Gibson and R. J. Hubbold. Perceptually-driven radiosity. *Computer Graphics Forum*, 16(2):129–141, 1997.

[GH00] S. Gibson and R. J. Hubbold. A perceptually-driven parallel algorithm for efficient radiosity simulation. *IEEE Transactions on Visualization and Computer Graphics*, 6(3):220–235, July 2000.

[GMY97] A. Gaddipatti, R. Machiraju, and R. Yagel. Steering image generation with wavelet based perceptual metric. *Computer Graphics Forum*, 16(3):241–252, August 1997.

[Gui54] J.P. Guilford. *Psychometric methods*. McGraw-Hill, New York, 1954.

[Hec90] Paul S. Heckbert. Adaptive radiosity textures for bidirectional ray tracing. In *Proceedings of SIGGRAPH 90*, Computer Graphics Proceedings, Annual Conference Series, pages 145–154, August 1990.

[HMYS01] Jörg Haber, Karol Myszkowski, Hitoshi Yamauchi, and Hans-Peter Seidel. Perceptually guided corrective splatting. *Computer Graphics Forum*, 20(3):142–152, 2001.

[HWP97] David Hedley, Adam Worrall, and Derek Paddon. Selective culling of discontinuity lines. In *Eurographics Rendering Workshop 1997*, pages 69–80, June 1997.

[IKN98] Laurent Itti, Christof Koch, and Ernst Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11):1254–1259, 1998.

[Lub95] Jeffery Lubin. A visual discrimination model for imaging system design and evaluation. In *Vision Models for Target Detection and Recognition*, World Scientific, pages 245–283, August 1995.

[Mit87] Don P. Mitchell. Generating antialiased images at low sampling densities. In *Proceedings of SIGGRAPH 87*, Computer Graphics Proceedings, Annual Conference Series, pages 65–72, July 1987.

[MK00] Karol Myszkowski and Tosiyasu L. Kunii. A case study towards validation of global illumination algorithms:. *The Visual Computer*, 16(5):271–288, 2000.

[ML92] Gary W. Meyer and Aihua Liu. Color spatial acuity control of a screen subdivision image synthesis algorithm. In *Proceedings of SPIE 92*, Human Vision, Visual Processing, and Digital Display, pages 387–399, February 1992.

[MRT99] Karol Myszkowski, Przemyslaw Rokita, and Takehiro Tawara. Perceptually-informed accelerated rendering of high quality walk-through sequences. In *Eurographics Rendering Workshop 1999*, pages 5–18, June 1999.

[MTAS01] Karol Myszkowski, Takehiro Tawara, Hiroyuki Akamine, and Hans-Peter Seidel. Perception-guided global illumination solution for animation rendering. In *Proceedings of ACM SIGGRAPH 2001*, Computer Graphics Proceedings, Annual Conference Series, pages 221–230, August 2001.

[Mys98] Karol Myszkowski. The visible differences predictor: Applications to global illumination problems. In *Eurographics Rendering Workshop 1998*, pages 223–236, June 1998.

[NMP98] Laszlo O. Neumann, Kresimir Matkovic, and Werner Purgathofer. Perception based color image difference. *Computer Graphics Forum*, 17(3):233–242, 1998.

[Pel91] E. Peli. Contrast in complex images. *J. Opt. Soc. Am. A*, 7(10):2032–2040, 1991.

[PFG00] Fabio Pellacini, James A. Ferwerda, and Donald P. Greenberg. Toward a psychophysically-based light reflection model for image synthesis. In *Proceedings of ACM SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, pages 55–64, July 2000.

[Poy96] Charles A. Poynton. *A Technical Introduction to Digital Video*. Wiley, New York, 1996.

[PP99] J. Prikryl and Werner Purgathofer. Perceptually-driven termination for stochastic radiosity. In *Seventh International Conference in Central Europe on Computer Graphics and Visualization (Winter School on Computer Graphics)*, February 1999.

[Pur87] Werner Purgathofer. A statistical method for adaptive stochastic sampling. *Computers and Graphics*, 11(2):157–162, 1987.

[Ram00] Mahesh Ramasubramanian. A perceptually based physical error metric for realistic image synthesis. Master's thesis, Cornell University, 2000.

[RE98] K. Brunnstrom R. Eriksson, B. Andren. Modeling of perception of digital images: A performance study. *SPIE Human Vision and Electronic Imaging*, 3:88–97, 1998.

[RPG99] Mahesh Ramasubramanian, Sumanta N. Pattanaik, and Donald P. Greenberg. A perceptually based physical error metric for realistic image synthesis. In *Proceedings of SIGGRAPH 99*, Computer Graphics Proceedings, Annual Conference Series, pages 73–82, August 1999.

[TJ97] Rasmus Tamstorf and Henrik Wann Jensen. Adaptive sampling and bias estimation in path tracing. In *Eurographics Rendering Workshop 1997*, pages 285–296, June 1997.

[Tor58] Warren S. Torgerson. *Theory and methods of scaling*. Wiley, New York, 1958.

[TR93] Jack Tumblin and Holly E. Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics & Applications*, 13(6):42–48, November 1993.

[VMKK00] Vladimir Volevich, Karol Myszkowski, Andrei Khodulev, and Edward A. Kopylov. Using the visual differences predictor to improve performance of progressive global illumination computations. *ACM Transactions on Graphics*, 19(2):122–161, April 2000.

[War92] Gregory J. Ward. Measuring and modeling anisotropic reflection. In *Proceedings of SIGGRAPH 92*, Computer Graphics Proceedings, Annual Conference Series, pages 265–272, July 1992.

[War94] Greg Ward. A contrast-based scalefactor for luminance display. In *Graphics Gems IV*, pages 415–421. Academic Press Professional, Inc., 1994.

[WDP99] Bruce Walter, George Drettakis, and Steven Parker. Interactive rendering using the render cache. In *Eurographics Rendering Workshop 1999*, June 1999.

[WH92] Gregory J. Ward and Paul S. Heckbert. Irradiance gradients. In *Eurographics Rendering Workshop 1992*, pages 85–98, May 1992.

[WHSG97] Bruce Walter, Philip M. Hubbard, Peter Shirley, and Donald P. Greenberg. Global illumination using local linear density estimation. *ACM Transactions on Graphics*, 16(3):217–259, July 1997.

[WPG02] B. Walter, S. N. Pattanaik, and D. P. Greenberg. Using perceptual texture masking for efficient image synthesis. *Computer Graphics Forum*, 21(3):393–399, 2002.

[YPG01] Hector Yee, Sumanta Pattanaik, and Donald P. Greenberg. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics*, 20(1):39–65, January 2001.