Daylight Simulation with Photon Maps

Roland Schregle

Dissertation zur Erlangung des Grades
des Doktors der Ingenieurwissenschaften
der Naturwissenschaftlich-Technischen Fakultät I
der Universität des Saarlandes

Saarbrücken, 2004

Tag des Kolloquiums: 26. Oktober 2004

Dekan der Naturwissenschaftlich-Technischen Fakultät I: Prof. Dr. Jörg Eschmeier

Prüfungskommission

Vorsitzender: Prof. Dr. Philipp Slusallek, Universität des Saarlandes

1. Berichterstatter: Prof Dr. Hans-Peter Seidel, MPI für Informatik, Saarbrücken

2. Berichterstatter: Priv. Doz. Dr. Volker Wittwer, Fraunhofer ISE, Freiburg

Akademischer Beisitzer: Dr. Karol Myszkowski, MPI für Informatik, Saarbrücken

Abstract

Physically based image synthesis remains one of the most demanding tasks in the computer graphics field, whose applications have evolved along with the techniques in recent years, particularly with the decline in cost of powerful computing hardware. Physically based rendering is essentially a niche since it goes beyond the photorealistic look required by mainstream applications with the goal of computing actual lighting levels in physical quantities within a complex 3D scene. Unlike mainstream applications which merely demand visually convincing images and short rendering times, physically based rendering emphasises accuracy at the cost of increased computational overhead. Among the more specialised applications for physically based rendering is lighting simulation, particularly in conjunction with daylight.

The aim of this thesis is to investigate the applicability of a novel image synthesis technique based on Monte Carlo particle transport to daylight simulation. Many materials used in daylight simulation are specifically designed to redirect light, and as such give rise to complex effects such as caustics. The photon map technique was chosen for its efficent handling of these effects. To assess its ability to produce physically correct results which can be applied to lighting simulation, a validation was carried out based on analytical case studies and on simple experimental setups.

As prerequisite to validation, the photon map's inherent bias/noise tradeoff is investigated. This tradeoff depends on the density estimate bandwidth used in the reconstruction of the illumination. The error analysis leads to the development of a *bias compensating operator* which adapts the bandwidth according to the estimated bias in the reconstructed illumination.

The work presented here was developed at the Fraunhofer Institute for Solar Energy Systems (ISE) as part of the FARESYS project sponsored by the German national research foundation (DFG), and embedded into the RADIANCE rendering system.

Zusammenfassung

Die Erzeugung physikalisch basierter Bilder gilt heute noch als eine der rechenintensivsten Aufgaben in der Computergraphik, dessen Anwendungen sowie auch Verfahren in den letzten Jahren kontinuierlich weiterentwickelt wurden, vorangetrieben primär durch den Preisverfall leistungsstarker Hardware. Physikalisch basiertes Rendering hat sich als Nische etabliert, die über die photorealistischen Anforderungen typischer Mainstream-Applikationen hinausgeht, mit dem Ziel, Lichttechnische Größen innerhalb einer komplexen 3D Szene zu berechnen. Im Gegensatz zu Mainstream-Applikationen, die visuell überzeugend wirken sollen und kurze Rechenzeiten erforden, liegt der Schwerpunkt bei physikalisch basiertem Rendering in der Genauigkeit, auf Kosten des Rechenaufwands. Zu den eher spezialisierten Anwendungen im Gebiet des physikalisch basiertem Renderings gehört die Lichtsimulation, besonders in Bezug auf Tageslicht.

Das Ziel dieser Dissertation liegt darin, die Anwendbarkeit eines neuartigen Renderingverfahrens basierend auf Monte Carlo Partikeltransport hinsichtlich Tageslichtsimulation zu untersuchen. Viele Materialien, die in der Tageslichtsimulation verwendet werden, sind speziell darauf konzipiert, Tageslicht umzulenken, und somit komplexe Phänomene wie Kaustiken hervorrufen. Das Photon Map verfahren wurde aufgrund seiner effizienten Simulation solcher Effekte herangezogen. Zur Beurteilung seiner Fähigkeit, physikalisch korrekte Ergebnisse zu liefern, die in der Tageslichtsimulation anwendbar sind, wurde eine Validierung anhand analytischer Studien sowie eines einfachen experimentellen Aufbaus durchgeführt.

Als Voraussetzung zur Validierung wurde der Photon Map bezüglich seiner inhärenten Wechselwirkung zwischen Rauschen und systematischem Fehler (Bias) untersucht. Diese Wechselwirkung hängt von der Bandbreite des Density Estimates ab, mit dem die Beleuchtung aus den Photonen rekonstruiert wird. Die Fehleranalyse führt zur Entwicklung eines *Bias compensating Operators*, der die Bandbreite dynamisch anhand des geschätzten Bias in der rekonstruierten Beleuchtung anpasst.

Die hier vorgestellte Arbeit wurde am Fraunhofer Institut für Solare Energiesysteme (ISE) als teil des FARESYS Projekts entwickelt, daß von der Deutschen Forschungsgemeinschaft (DFG) finanziert wurde. Die Implementierung erfolgte im Rahmen des RADIANCE Renderingsystems.

Detailed Abstract

The lighting industry is increasingly turning to computer simulations to analyse artificial lighting both in visual and numeric terms. While architectural scale models are still used to some degree, they are time consuming and expensive to construct. CAD models coupled with computer shading techniques offer a viable alternative at a fraction of the cost required for traditional methods. The aim of the simulation is to aid the lighting engineer in deciding over the choice of lighting fixtures and their placement during the planning phase. Computer graphics techniques provide a computer generated prediction of the lighting levels expected for a given lighting configuration. Obviously, physical accuracy is imperative for such an application since the prediction is a decisive factor contributing to the comfort (and therefore the productivity) of the inhabitants once the building is completed and the lighting installed.

Daylight simulation follows the same principles as artificial lighting simulation, but under the utilisation of sunlight (possibly in conjunction with artificial light). Techniques have been developed which exploit as well as manipulate natural light in buildings to reduce power consumption, glare, and heat buildup in summer. These techniques include the installation of *daylight systems* designed to redirect or block direct sunlight while transmitting diffuse skylight, i.e. they are *angularly selective*. These systems are constructed from specular materials which are crucial for their selectivity. Consequently, a reliable daylight simulation requires an accurate model of both the system's geometry and its materials.

Most image synthesis tools cannot adequately simulate the light transport arising from the specular properties of angularly selective daylight systems, and therefore fail to predict lighting levels within reasonable accuracy, as well as locating potential sources of glare. The specular reflections from these systems give rise to *caustics*, which cannot be efficiently sampled with traditional *backward raytracing* techniques. A novel *forward raytracing* approach is required to accurately account for these effects, and the *photon map* discussed in this thesis is one such algorithm.

The aim of this thesis is to develop an efficient and accurate image synthesis tool based on forward raytracing specifically for daylight simulation, but which can also be used for more general visualisation. The primary motivation for doing so is the difficulty imposed by specular daylight systems on already existing lighting simulation tools, specifically the RADIANCE system which reveals shortcomings in simulating the redirecting properties of these systems. The photon map is used as basis for the extensions and integrated into RADIANCE. Its applicability to daylight simulation is assessed in the form of a validation by comparing the results with analytical solutions and measurements from an experimental setup.

As a prerequisite to validation, the problem of bias and noise in the illumination reconstructed from the photon map using nearest neighbour techniques is also investigated, leading to the proposal of a novel bias compensating algorithm which

improves the accuracy of caustics in particular. The bias/noise tradeoff is rarely addressed in detail in the literature. Since this thesis aims to endorse the photon map as a lighting analysis tool, it is imperative to analyse its fundamental limitations and develop a means of compensating for them.

Bias and noise are inversely related to each other and subject to the density estimate bandwidth. In situations involving caustics, low bias is preferred in order to preserve detail. On the other hand, in situations involving uniform irradiance, low noise is preferred. This implies that an optimal bandwidth must be dynamically adjustable to the illumination. The proposed bias compensating operator uses a binary search within a specified range for the optimum bandwidth. This search is governed by error estimates extracted from the reconstructed irradiance in order to identify probable bias using the *central limit theorem*. Unlike previous work, the operator is specifically geared toward quantitative analysis such as applications in lighting design. It is conceptually simple and general enough to be used in most density estimation frameworks because it does not rely on additional information, but rather makes use of what can be deduced from the reconstructed irradiance.

Analytical validation is an effective means of ascertaining the accuracy and fundamental soundness of a global illumination algorithm. Though necessarily simple and constrained in scope, this approach defines a controlled environment which is generally more tractable than even the simplest setup used in a physical validation. A spherical “furnace” type scene is used with diffuse reflection and with a special case of the Lafortune BRDF model to confirm that the photon map produces minimal deviations compared to the analytical solution for the constant indirect irradiance on the sphere’s inner surface. The derivation of the solution is based on a series expansion of the rendering equation which is greatly simplified by the symmetry of the setup.

The experimental validation is based on photometric measurements of a simple scale model using an artificial light source. A validation methodology is proposed which emphasises tractability and error minimisation. To this end, simple component case studies are initially carried out to test individual light transport modes (i.e. single diffuse and specular reflection). These serve as foundation for more complex compound case studies which test interreflection (diffuse only and in conjunction with a single specular reflection). Analytical solutions are drawn upon as reference where possible to assess not only the accuracy of the simulation, but also of the measurements themselves.

Physical accuracy is imperative for a validation, thereby necessitating the integration of the light source EDF (*emission distribution function*) and the material BRDFs into the simulation. The EDF is obtained with a novel method based on extraction from HDR (*high dynamic range*) camera images. The BRDFs of the materials are obtained from goniophotometric measurements. Physical validation requires attention to detail and scrutiny in order to minimise errors on the physical side. Since the simulations use measured BRDF data, errors on the physical side will be carried over into the simulation. Consequently, both data sets are subjected to analysis and verification prior to simulation, as well as correction in the case of

the BRDF data due to limitations of the goniophotometer device. The BRDF data also necessitates the development of a resampling technique based on nearest neighbour lookups in a bilevel kd-tree in order to efficiently evaluate the BRDF at arbitrary incident and exitant directions during the simulation.

The uncertainties inherent in a physical validation are discussed, particularly pertaining to measurement inaccuracies. Primary sources of error are identified and accounted for where necessary. These uncertainties are carried over into the results as error bounds. The most problematic factor of the experimental validation are the material BRDFs. An accurate simulation is accomplished by using the measured BRDF data directly, with a fitted analytical BRDF model providing the PDF (*probability density function*) for the sample ray distribution.

The validation results show that both forward and backward raytracers (represented by photon map and RADIANCE, respectively), deliver very similar results. However, while both algorithms perform similarly as far as accuracy is concerned, the computation times differ substantially. This disparity is largely attributed to the fact that the photon map constructs a complete, reusable global illumination solution in the forward pass. RADIANCE, on the other hand, depends on the recursion governed by the number of *ambient bounces*. If this parameter is set too low, underprediction results. This is particularly important with scenes characterised by high reflectance as shown in the analytical validation. Furthermore, backward raytracers are fundamentally inefficient in adequately resolving caustics such as those produced by daylight systems.

Detaillierte Zusammenfassung

Die Beleuchtungsindustrie wendet in zunehmendem Maße Computersimulationen an zur Analyse künstlicher Beleuchtung hinsichtlich qualitativer sowie auch quantitativer Aspekte. Obwohl architektonische Modelle noch teilweise Verwendung finden, ist deren Aufbau vergleichsmäßig teuer und zeitraubend. CAD-Modelle, die per Computer visualisiert werden, sind eine attraktive Alternative die einen Bruchteil der Kosten traditioneller Methoden aufwenden. Ziel der Simulation ist es, den Beleuchtungstechniker bei der Wahl der Leuchten und deren Plazierung während der Planungsphase zu unterstützen. Verfahren aus der Computergraphik liefern eine Voraussage der zu erwartenden Beleuchtungsstärken für eine gegebene Konfiguration. Offensichtlich ist physikalische Genauigkeit unabdingbar für diese Anwendung, denn die darauf beruhende Simulation entscheidet letztendlich auch über das Behagen (und somit die Produktivität) der Bewohner, wenn das Gebäude errichtet und dessen Beleuchtung installiert wurde.

Tageslichtsimulation beruht auf den gleichen Prinzipien wie die Beleuchtungssimulation mit Kunstlicht, jedoch unter verwendung von Sonnenlicht (teilweise auch in Kombination mit Kunstlicht). Verfahren zur Ausnutzung von Tageslicht innerhalb Gebäuden sind entwickelt worden, die darauf abzielen, Blendung und Erwärmung im Sommer, sowie auch den Stromverbrauch durch Kunstlicht zu reduzieren. Zu diesem Zweck werden *Tageslichtsysteme* installiert, die dazu konzipiert sind, direktes Sonnenlicht zu blockieren oder umzulenken, während diffuses Himmelslicht durchgelassen wird, d.h. sie arbeiten *richtungsselektiv*. Diese Systeme sind aus spiegelnden Materialien beschaffen, die entscheidend für dessen Selektivität sind. Folglich setzt eine akkurate Tageslichtsimulation auch ein genaues Modell der Systemgeometrie sowie dessen Materialien voraus.

Die meisten Visualisierungsverfahren sind nicht in der Lage, den durch die spiegelnden Eigenschaften richtungsselektiver Tageslichtsysteme entstehenden Lichttransport adäquat zu simulieren, und somit auch keine verlässliche Voraussage über Beleuchtungsstärken sowie potentielle Blendungen liefern können. Die spiegelnden Reflexionen dieser Systeme erzeugen *Kaustiken*, die mit traditionellem *Backward Raytracing* nur ineffizient erfasst werden können. Hierzu wird ein *Forward Raytracer* benötigt, um diese Effekte akkurat zu simulieren, und dazu wird das *Photon Map* Verfahren als populärster Vertreter dieses Genres im Rahmen dieser Dissertation untersucht.

Ziel dieser Dissertation ist die Entwicklung eines effizienten und akkuraten Visualisierungstools basierend auf Forward Raytracing, daß speziell zur Tageslichtsimulation, aber auch zur allgemeineren Visualisierung einsetzbar ist. Die Motivation liegt hauptsächlich in den Schwierigkeiten, die spiegelnde Tageslichtsysteme bei einer Vielzahl der vorhandenen Lichtplanungstools bereiten, insbesondere das RADIANCE System, das erhebliche Einschränkungen bei der Simulation der lichtlenkenden Funktion dieser Systeme aufweist. Der Photon Map ist die Ausgangsbasis

für die Erweiterungen, die in das RADIANCE System integriert werden. Dessen Anwendbarkeit im Rahmen der Tageslichtsimulation wird in Form einer Validierung beurteilt, in der die Ergebnisse mit analytischen Lösungen sowie auch Messungen an einem experimentellen Aufbau verglichen werden.

Als Voraussetzung zur Validierung wird das Problem des Rauschens und Bias (zufälliger resp. systematischer Fehler) in der vom Photon Map mittels *Nearest Neighbour* Suche rekonstruierten Beleuchtungsstärke untersucht. Dies führt zur Entwicklung eines neuartigen Algorithmus zur Biaskompensierung, der inbesondere die Rekonstruktion von Kaustiken verbessert. Die Wechselwirkung zwischen Rauschen und Bias wird in der Fachliteratur selten ausführlich erläutert. Da diese Arbeit darauf abzielt, den Photon Map als Werkzeug zur Beleuchtungssimulation zu untersuchen, ist eine Analyse dessen Grundsätzlichen Einschränkungen, sowie die Entwicklung eines Verfahrens zu dessen Kompensierung, notwendig.

Bias und Rauschen verhalten sich invers zueinander und sind abhängig von der Bandbreite des *Density Estimates*, mit dem die Beleuchtungsstärke aus dem Photon Map rekonstruiert wird. Im Falle von Kaustiken wird niedriges Bias zugunsten der Auflösung von Details bevorzugt. In Bereichen gleichmäßiger Beleuchtung wird dagegen niedriges Rauschen bevorzugt. Dies deutet darauf hin, daß eine optimale Bandbreite durch dynamische Anpassung an die Beleuchtung erzielbar ist. Der vorgestellte Algorithmus zur Biaskompensierung basiert auf einer Binärsuche nach einer optimalen Bandbreite innerhalb eines vorgegebenen Intervalls. Diese Suche wird durch Fehlerabschätzungen gesteuert, die aus der rekonstruierten Beleuchtungsstärke extrahiert werden, um wahrscheinlichen Bias mittels des *Zentralen Grenzwertsatzes* zu erkennen. Im Gegensatz zu vorangegangenen Arbeiten ist der Algorithmus speziell für quantitative Analysen gedacht, wie z.B. in der Lichtplanung. Das Konzept ist einfach und allgemein, und kann somit im Rahmen der meisten Density Estimation Verfahren angewandt werden, da es nicht auf zusätzliche Informationen beruht, sondern sich allein auf das bezieht, was aus der rekonstruierten Beleuchtung abgeleitet werden kann.

Analytische Validierung ist eine effektive Methode um die Genauigkeit und grundsätzliche Gültigkeit eines *Global Illumination* Algorithmus zu ermitteln. Obwohl notwendigerweise einfach und eingeschränkt im Umfang, definiert dieser Ansatz eine kontrollierte Umgebung die generell nachzuvollziehbarer ist, als den einfachsten experimentellen Aufbau im Rahmen einer physikalischen Validierung. Als Testszene dient der Innenraum einer Einheitskugel (sogenannter “Ofen”), dessen Innenfläche durch diffuse, sowie auch spiegelnde Reflexion (definiert durch ein Sonderfall des Lafortune BRDF-Modells) charakterisiert ist. Hiermit wird vergewissert, daß der Photon Map tatsächlich minimale Abweichungen gegenüber der analytischen Lösung für die konstante Beleuchtungsstärke auf der Innenfläche der Kugel aufweist. Die Herleitung der Lösung basiert auf einer Reihenentwicklung der *Rendering-Gleichung*, die durch die Symmetrie der Szene stark vereinfacht wird.

Die experimentelle Validierung basiert auf photometrischen Messungen an einem einfachen physikalischen Modell unter Verwendung einer Kunstlichtquelle. Der Schwerpunkt der vorgestellten Validierungsmethodologie liegt in der Nachvollzieh-

barkeit und Fehlerminimierung. Zu diesem Zweck werden anfangs einfache Komponentenstudien durchgeführt, um einzelne Lichttransportpfade zu testen (d.h. einfache Diffus- bzw. Spiegelreflexion). Diese Studien dienen als Voraussetzung für darauf folgende, komplexere Kombinationsstudien, die Interreflexion testen (ausschließlich diffus sowie auch kombiniert mit einfacher Spiegelreflexion). Analytische Lösungen werden, soweit möglich, herangezogen, um nicht nur die Genauigkeit der Simulation, sondern auch der Messungen zu überprüfen.

Physikalische Korrektheit ist zwingend für eine Validierung; dies erfordert die Integration der LVK (*Lichtverteilungskurve*) der Lichtquelle sowie der BRDFs der Materialien in die Simulation. Die LVK wird durch ein neuartiges Verfahren aus Aufnahmen einer Leuchtdichtekamera extrahiert. Die BRDFs der Materialien werden dagegen mit Messungen eines Goniophotometers erhalten. Eine physikalische Validierung erfordert sorgfältiges Vorgehen, um Meßfehler zu minimieren. Da die gemessenen BRDF-Daten in die Simulation eingehen, werden diese Meßfehler auch übertragen. Folglich müssen LVK und BRDFs im Vorfeld analysiert, verifiziert, und, im Falle der BRDF aufgrund Einschränkungen des Goniophotometers, auch korrigiert werden. Die Verwendung der BRDF-Daten in der Simulation erfordert auch die Entwicklung eines Resamplingverfahrens basierend auf Nearest Neighbour Suche in einem zweistufigen *kd-Baum*, um die gemessene BRDF effizient für beliebige Ein- und Ausfallswinkel auszuwerten.

Ferner werden die inhärenten Unsicherheiten einer physikalischen Validierung erläutert, besonders in Bezug auf Meßfehler. Die hauptsächlichen Fehlerquellen werden identifiziert und, wenn notwendig, auch berücksichtigt. Die Unsicherheiten gehen in die Ergebnisse ein in Form von Fehlerbalken. Die größten Schwierigkeiten in der experimentellen Validierung bereiten die BRDFs der verwendeten Materialien. Eine genaue Simulation wird erzielt, indem die gemessenen BRDF-Daten direkt verwendet werden, wobei ein darauf optimiertes analytisches BRDF-Modell die Wahrscheinlichkeitsdichte für die Strahlenverteilung liefert.

Die Validierung zeigt, daß Forward sowie Backward Raytracer (repräsentiert durch Photon Map resp. RADIANCE) sehr ähnliche Ergebnisse liefern. Obwohl beide Algorithmen vergleichbar sind bezüglich deren Genauigkeit, unterscheiden sich die Rechenzeiten jedoch erheblich. Dieses Mißverhältnis ist auf die Tatsache zurückzuführen, daß der Photon Map eine komplete, wiederverwendbare Lösung zur globalen Beleuchtung konstruiert. RADIANCE, dagegen, beruht auf Rekursion, die durch die Anzahl der *Ambient Bounces* begrenzt wird. Wird dieser Parameter zu niedrig gesetzt, fällt die errechnete Beleuchtungsstärke zu niedrig aus. Dies spielt besonders eine Rolle in Szenen mit hoher Reflektivität, wie in der analytischen Validierung gezeigt wird. Ferner haben Backward Raytracer grundsätzliche Effizienzprobleme beim adäquaten Auflösen von Kaustiken, wie sie z.B. von Tageslichtsystemen produziert werden.

“Research is what I’m doing when I don’t know what I’m doing.”
– Wernher von Braun

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

Introduction

This chapter introduces the focus and aims of this thesis and how it contributes to computer graphics research. The problem at hand is explained, and an overview of the thesis is presented.

1.1 Daylight Simulation

The lighting industry is increasingly turning to computer simulations to analyse artificial lighting both in visual and numeric terms. While architectural scale models are still used to some degree, they are time consuming and expensive to construct. CAD models coupled with computer shading techniques offer a viable alternative at a fraction of the cost required for traditional methods.

Lighting simulations are generally applied to projects in such diverse fields as:

architecture: offices and public buildings, both interiors and exteriors

industry: factory lighting

transport: lighting for ships, train stations, airports, and roads

emergency: exit signs and emergency markings

The aim of the simulation is to aid the lighting engineer in deciding over the choice of lighting fixtures and their placement during the planning phase. The factors that govern these choices include the characteristics of the fixtures themselves, such as the directional intensity distribution (also known as a *goniometric diagram*) and spectral distribution. Often this planning is carried out by contractors for a client. Computer graphics techniques provide a computer generated prediction of the lighting levels expected for a given lighting configuration. Obviously, physical accuracy is imperative for such an application, and can mean the difference between a satisfied client and an embarrassment once the building is completed and the lighting installed.

Daylight simulation follows the same principles as artificial lighting simulation, but under the utilisation of sunlight (possibly in conjunction with artificial light). Techniques have been developed which exploit as well as manipulate natural light in buildings to reduce power consumption, glare, and heat buildup in summer. These techniques include the installation of *daylight systems* designed to redirect sunlight. To evaluate a daylighting proposal, computer analyses of natural light levels using daylight systems are carried out, often encompassing seasonal changes over an entire year (figure 1.1). Such a simulation requires the exact position (latitude and longitude) of the building as well as the time of day and year to determine the sun's position.

Daylight systems for glare reduction are often integrated or mounted outside windows in the form of façades (figure 1.2). These systems are generally designed to redirect or block direct sunlight while transmitting diffuse skylight. Not only does this increase visual comfort for the inhabitants in the vicinity of the window, but it also reduces the sun's thermal influence on the interior. These daylight systems are constructed from specular materials which are crucial for their function. Consequently, a reliable daylight simulation requires an accurate model of both the system's geometry and its materials. Integrating the material properties into the simulation, particularly in the form of measured reflection / transmission data, is a challenge, since the appropriate parameters must be determined for analytical models which reproduce this behaviour, and only the more complex models account for the majority of material properties one is likely to encounter.

Most image synthesis tools cannot adequately simulate the effects arising from the specular properties of sophisticated daylight systems, and therefore fail to predict lighting levels within reasonable accuracy, as well as locating potential sources of glare. The specular reflections from these systems give rise to *caustics*, which cannot be efficiently sampled with traditional ray tracing techniques. A novel approach with a modified sampling scheme is required to accurately account for these effects, and the *photon map* discussed in this thesis is one such algorithm.

1.1.1 Evaluation Methods

The analysis of a lighting concept can be carried out in *qualitative* or *quantitative* terms. Qualitative analysis concerns the visual impression of the lighting and is primarily geared towards optimising aesthetics. Consequently, the analysis entails generating images of the illuminated scene, typically from different perspectives and with several candidate lighting designs. On the other hand, quantitative analysis considers the actual lighting levels in the design and may not even require rendering images, but rather generating illuminance values at a set of discrete points in the scene corresponding to actual measurements with a luxmeter[1]. In some cases a combined qualitative and quantitative analysis can be presented in the form of contour maps, or falsecolour images such as figure 1.2.

[1] Technically, these measurement points are essentially treated as viewpoints for computer generated images

Summer

Spring/autumn

Winter

Figure 1.1: Seasonal daylight availability study in office space for june, september, and december 20th at 3:00 pm. The sky is based on the CIE clear sky model (described in appendix C.2). Glare at workplane height can severely impair legibility on desks and computer monitors, as is the case here for spring/autumn and winter.

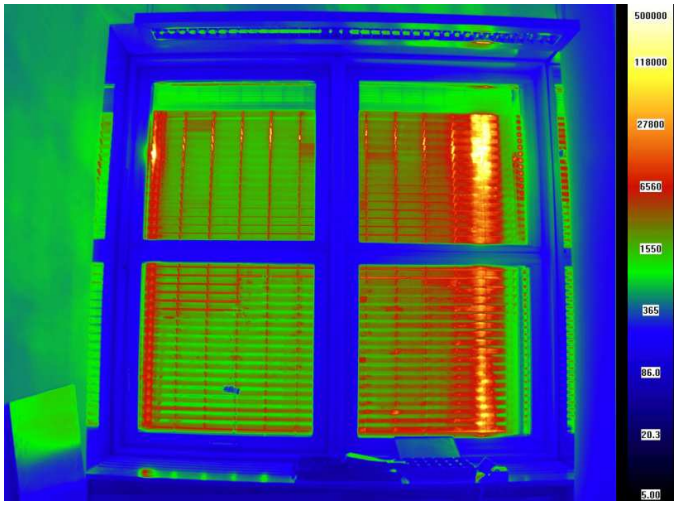


Figure 1.2: Falsecolour image of luminance (cd/m^2) on daylight system. The system is designed to reduce glare and heat buildup from direct sunlight (incident in the upper right), while still admitting diffuse skylight. In this particular system this is achieved through retroreflection from prisms integrated into the lamella. Without such a system, glare would penetrate into the room and dramatically impair the occupants' visual comfort, particularly on desks and monitors in office environments.

Techniques for the generation of computer images, or *rendering*, are usually subject to a tradeoff between accuracy and performance, which affects the predicted lighting levels. The acceptable relative error tolerance for lighting applications is generally considered to be $\pm$10%. However, depending on the particular global illumination algorithm used and its parametrisation, this error margin is not always guaranteed in all conceivable situations.

A popular and intuitive quantitative evaluation method is the *daylight factor*, which characterises the daylight quantity in a building's interior [Lit90]. It is defined as the percentage of the indoor illuminance $E_i(\vec{x})$ at a point $\vec{x}$, usually at workplane height (ca. 0.8 m from the floor) to the outdoor horizontal illuminance E_o of an overcast sky:

$$DF\,(\vec{x}) = 100\frac{E_i\,(\vec{x})}{E_o}. \tag{1.1}$$

Because it is defined for uniform overcast skies, the daylight factor is independent of the building's orientation. Furthermore, since daylight is minimal during overcast conditions, the factor is a measure of interior daylight levels in a worst case scenario.

1.1.2 Simulation Components

A scene description for daylight simulation consists of the following components (see figure 1.3):

- the scene geometry (e.g. room with daylight systems), often obtained from a CAD model
- the materials used by the geometry, which may be obtained from measurements as directional reflection distributions
- optional goniometric data for artificial lights (if any)
- a sky model comprising a localised solar source and a diffuse sky source.

The scene description is passed into the renderer along with either a camera model or a set of measuring points for illuminance data, depending on whether a visual or numeric analysis is performed.

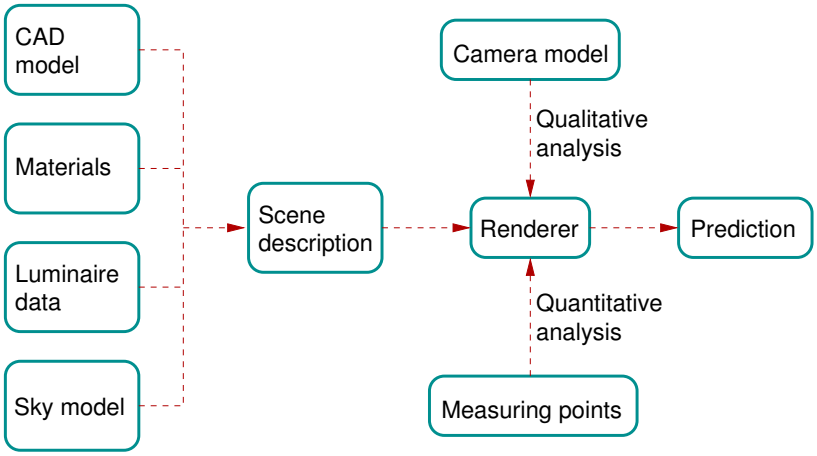


Figure 1.3: Components of daylight simulation

1.1.3 Daylighting Geometry

Daylight simulation differs from artificial lighting simulation mainly in the light sources it uses. While artificial lighting simulation uses local light sources (which are contained within the scene geometry), daylight simulation requires distant light sources in order to model the sun and sky. These light sources are external to the scene geometry and conceptually infinitely distant, such that their illuminance is translationally invariant throughout the scene. They are defined by an incident direction and a solid angle around this direction, which are both independent of the position in the scene. RADIANCE [War94] is one of the few renderers which offer these primitives specifically for daylight simulation.

A typical daylight simulation geometry is depicted in figure 1.4. It consists of the following components:

- an interior space (often an office environment) containing the viewpoint or measuring points

- a daylight system under consideration for glare reduction
- a hemispherical source to account for diffuse skylight, usually subject to a directional luminance distribution obtained from an analytical sky model (see appendix C for some commonly used models)
- a small (ca. 1°) solar source for direct sunlight with very high emission (typically on the order of $10^6\ W/m^2sr$)
- a second hemispherical source to account for sunlight reflected diffusely from the ground (with a typical reflectance of ca. 20%).

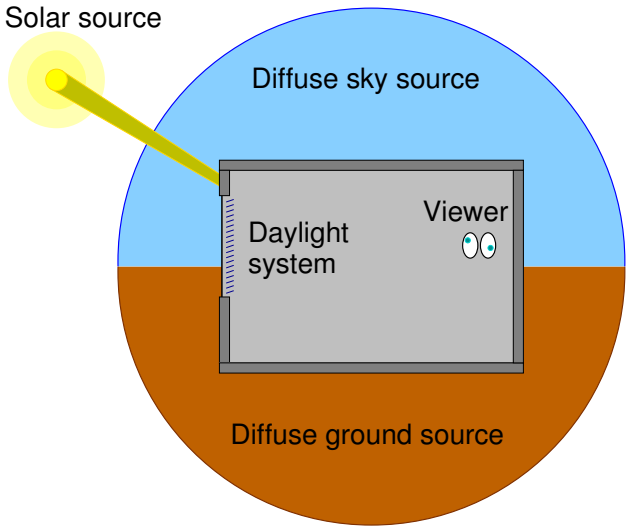


Figure 1.4: Typical daylight simulation geometry

1.2 Thesis Aim and Contribution

The aim of this thesis is to develop an efficient and accurate image synthesis tool for daylight simulation, but also for more general visualisation. The primary motivation for doing so is the difficulty imposed by specular daylight systems on already existing lighting simulation tools, specifically the RADIANCE system which reveals shortcomings in simulating the redirecting properties of these systems. The photon map is used as basis for the extensions and integrated into RADIANCE. Its applicability to daylight simulation is assessed in the form of a validation by comparing the results with analytical solutions and measurements from an experimental setup.

The experimental validation also encompasses the acquisition of material BRDFs and light source emission distributions. This includes the verification, correction (if necessary), and integration of the measured data into the simulation. This nontrivial task requires the development of novel approaches in the context of the

validation. Given the complexity of the task, the validation also aims at developing a sound methodology emphasising tractability and error minimisation. Such methodologies are rarely described in detail in the literature available to date.

Furthermore, the problem of bias and noise in the illumination reconstructed from the photon map using nearest neighbour techniques is also addressed, leading to the proposal of a novel bias compensating algorithm which improves the accuracy of caustics in particular. The bias/noise tradeoff is rarely addressed in detail in the literature. Since this thesis aims to endorse the photon map as a lighting analysis tool, it is imperative to analyse its fundamental limitations and develop a means of compensating them. While some of the other proposals brought forward to address these problems rely on perceptive error metrics (and are therefore solely geared towards qualitative analysis), the method developed in this thesis is specifically designed for quantitative analysis.

While image synthesis tools and techniques abound, few satisfy the demands of daylight simulation, which is still very much a niche in the rendering field. Such an application for the photon map has not yet been explored. Because of its popularity, validating the photon map for demanding applications such as lighting analysis is an undertaking that contributes to computer graphics research. Furthermore, physical validations of global illumination algorithms are very rare, and therefore significant in advancing the state of the art.

1.3 Thesis Outline

The outline of the thesis is as follows:

- Chapter 2 introduces the basics of physically based image synthesis, including radiometric terminology, BRDFs, and Monte Carlo ray tracing, thereby laying the theoretical foundation for the work presented here.
- Chapter 3 introduces the concept of forward ray tracing techniques, and the photon map in particular.
- Chapter 4 describes features and implementation details of the RADIANCE photon map and presents daylight simulation examples generated with the module.
- Chapter 5 presents an analysis of the bias/noise tradeoff inherent in the photon map and proposes a *bias compensating* operator which attempts to minimise this error.
- Chapter 6 describes the analytical validation of the RADIANCE photon map.
- Chapter 7 describes the experimental validation.

- The appendices provide supplementary material and validation data. The supplements include an overview of the RADIANCE rendering system, the BRDF models used in the validation, and sky models.

Chapter 2

Physically Based Image Synthesis

Computer image synthesis, or *rendering*, is the process of generating images of a 3D scene via an algorithm designed for a specific appearance. Many such algorithms exist, having different properties. They fall into a variety of categories to suit different requirements, based on how they model light interaction with the environment. Most algorithms can be classified as being *local* or *global* with respect to the simulated illumination. Local illumination algorithms are based on the rather naïve assumption that the illumination of a surface depends solely on the local material properties and the incident light sources. Interaction with other surfaces in the scene is disregarded. By contrast, global illumination algorithms model this interaction and have become the staple of research in the physically based rendering field, and have been refined over the last two decades.

The ultimate goal of image synthesis is *photorealism*: producing renderings indistinguishable from a photograph of the real object being modeled. However, the human eye is extremely adept at discerning fakes from reality, much to the vexation of computer graphics researchers. To them, photorealism is all about fooling the eye, and through their efforts they've come pretty close. Sophisticated rendering algorithms deliver stunning results, and the motion picture industry employs computer generated imagery extensively nowadays to save costs on props or for special effects which would be impossible to accomplish by traditional means. The fact that the audience is for the most part ignorant of the bogus imagery on the screen which they perceive as reality bears tribute to the success of photorealistic rendering techniques.

Photorealism is the mainstream application of image synthesis. Compromises must usually be made to obtain results in a timely fashion, as long as the imagery is still visually convincing. By contrast, physically based rendering is a niche outside this mainstream, and goes well beyond the goal of visual realism: the images are not only photorealistic, but also physically plausible in terms of the radiometric units quantifying the illumination on the surfaces. To this end, physically based algorithms

operate in radiometric units and reflection functions subject to physical constraints (energy conservation) which can be measured and verified. Consequently, these algorithms are more complex and slower than those tailoured for photorealism alone. Physical basis is the prerequisite for application of rendering techniques to applications such as lighting simulation. Only by applying physical units can a rendering have physical meaning and thus be considered a simulation. But before a physically based renderer can be used with confidence for applications which require accuracy it should be *validated*, i.e. checked for physical correctness by comparing its results with analytically derived as well as measured quantities using simple test scenes.

This chapter introduces physically based rendering concepts, and gives an overview of the major classes of global illumination algorithms. A comprehensive treatment of the techniques presented here can be found in the standard reference by Glassner [Gla95], while Philip Dutré's compendium [Dut03] is an excellent collection of relevant formulae and definitions.

2.1 Radiometry

Radiometric units [IES86, NRH+77] are the conventions by which physical quantities pertaining to light transport are measured. These are actually functions of wavelength, position, direction, time, and polarisation. When applied to computer graphics, these terms are usually simplified by ignoring the last two parameters. Rather than considering all possible wavelengths in the visible spectrum, these are typically narrowed down to the red, green, and blue primaries of the additive colour system.

2.1.1 Solid Angle

The *solid angle* ω is the 3D extension of the familiar 2D concept of angle. The 2D angle subtended by an object is the ratio of the arc length of its projection to the arc's radius. By the same token, the 3D solid angle subtended by an object is the ratio of the area of its hemispherical projection to the hemisphere's squared radius (figure 2.1), and is measured in *steradians* (sr). Note that both definitions are independent of the radius. Since the surface area of a sphere is $4\pi r^2$, the maximum solid angle is 4π sr.

$$\omega = \frac{A}{r^2} \qquad \left[sr = \frac{m^2}{m^2} \right] \tag{2.1}$$

2.1.2 Flux

The basic unit quantifying energy is the *joule* (J), denoted Q. Moving energy such as electromagnetic radiation (including light) is quantified by J/s, or *watts* (W). This

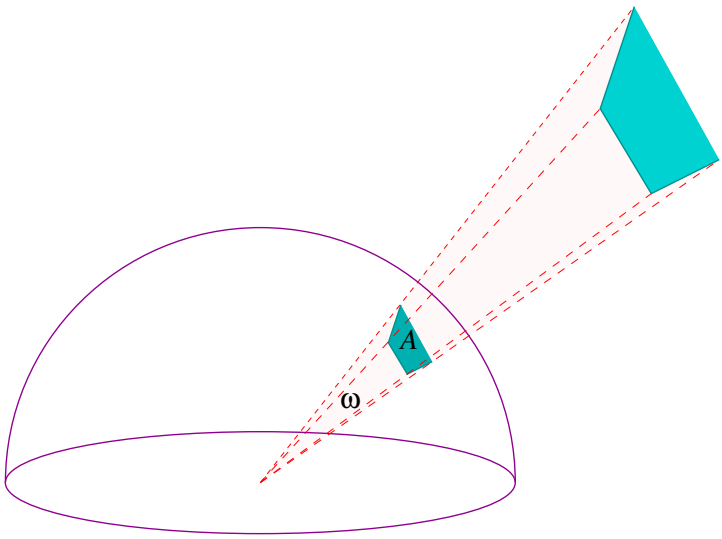


Figure 2.1: Solid angle

is termed the *radiant flux* Φ. It describes the energy passing through an arbitrary surface per unit time.

$$\Phi = \frac{dQ}{dt} \qquad \left[W = \frac{J}{s}\right] \tag{2.2}$$

2.1.3 Irradiance

Irradiance E quantifies the flux entering an area, measured in watts per square meter.

$$E = \frac{d\Phi_i}{dA} \qquad \left[\frac{W}{m^2}\right] \tag{2.3}$$

2.1.4 Radiosity

The counterpart to irradiance is the Radiosity B, which is defined similarly but specifies the flux leaving an area.

$$B = \frac{d\Phi_r}{dA} \qquad \left[\frac{W}{m^2}\right] \tag{2.4}$$

2.1.5 Radiant Intensity

This unit is often encountered when dealing with light source emission distributions (often termed *goniometric diagrams*), which specify the flux emitted per solid angle.

$$I = \frac{d\Phi}{d\omega} \qquad \left[\frac{W}{sr}\right] \tag{2.5}$$

2.1.6 Radiance

The most important radiometric term is the *radiance* L. This expresses the flux arriving at or leaving from a surface per unit solid angle and per unit projected (or foreshortened) area. The latter refers to the projection of the surface onto the plane perpendicular to the direction of flux propagation. It is measured in watts per steradian per square meter.

$$L = \frac{d^2\Phi}{d\omega dA_\perp} = \frac{dI}{dA_\perp} = \frac{dE}{d\omega} \qquad \left[\frac{W}{sr \cdot m^2}\right] \tag{2.6}$$

2.2 Photometry

Radiometry is the objective description of the physics behind radiative transport over the entire electromagnetic spectrum. Conversely, photometry restricts itself to wavelengths visible to humans and deals with the subjective issues involved. Photometry is particularly important in judging the perception of physically based renderings, based on knowledge of the human visual system (HVS). The HVS has a non-uniform response to a very small range in the electromagnetic spectrum, known as the *visual band* from 380 to 780 nm. This is characterised by the *luminous efficiency function* $V(\lambda)$ (figure 2.2).

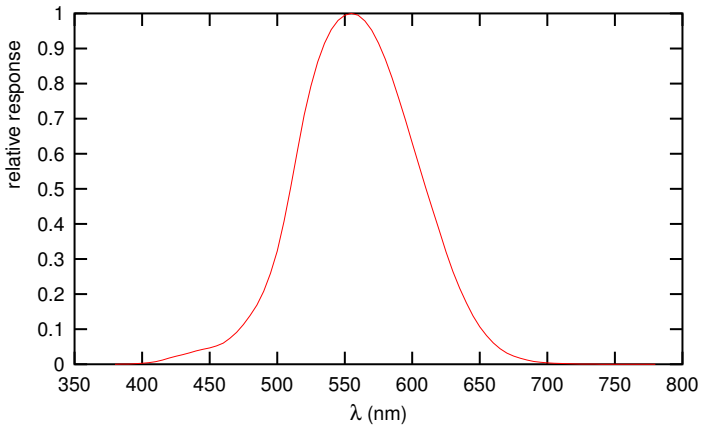


Figure 2.2: Luminous efficiency function of the HVS

Since physically based renderings are computed in radiometric units they are independent of the HVS and require adaptation by weighting them with the response curve. Photometry has developed its own set of units which are derived from their radiometric counterparts by weighting them according to $V(\lambda)$. These units are

listed in table 2.1. Photometric quantities are often encountered in physical light measurements and will be used extensively in this document.

Radiometric		Photometric	
Radiant flux	$[W]$	Luminant flux	$[lumen\ (lm)]$
Irradiance	$\left[\frac{W}{m^2}\right]$	Illuminance	$\left[\frac{lm}{m^2} = lux\ (lx)\right]$
Radiosity	$\left[\frac{W}{m^2}\right]$	Luminosity	$\left[\frac{lm}{m^2} = lux\ (lx)\right]$
Radiant intensity	$\left[\frac{W}{sr}\right]$	Luminant intensity	$\left[\frac{lm}{sr} = candela\ (cd)\right]$
Radiance	$\left[\frac{W}{sr \cdot m^2}\right]$	Luminance	$\left[\frac{cd}{m^2} = nit\right]$

Table 2.1: Radiometric and photometric units

Adapting computer generated images to subjective viewing is a complex task which requires a mapping from radiometric to photometric quantities capable of modelling the reponse of the HVS to the rendered scene. This *tone mapping* process is still under research and some very sophisticated methods have been developed.

The most important aspect of a tone mapping operator is the ability to compress the dynamic range of the rendering into a range suitable for display on an output device. Real word scenes have very high dynamic ranges; for example, a typical sunlit scene spans six orders of magnitude. A physically based renderer will reproduce this range. The HVS is capable of adapting to this dynamic range, hence a tone mapping operator is indispensable in simulating this adaptation process, particularly for applications such as lighting. Additionally, a tone mapping operator may simulate deficiencies of the HVS such as scattering within the lens and reduced acuity in low lighting. These limitations are also exploited by perceptionally based rendering techniques by optimizing computations to omit details which would not be perceivable by a human observer.

2.3 The BRDF

The *bidirectional reflectance distribution function* f_r, or BRDF, provides a general mechanism for describing arbitrary surface properties. Unlike empirical reflectance models, the BRDF has physical basis and can be measured. It is the most flexible representation of material properties, ranging from simple (e.g. Lambertian reflection) to complex (e.g. anisotropic reflection from brushed metal or varnished wood).

Given a point $\vec{x}$, an incident direction $\vec{\omega}_i$, and direction of reflection $\vec{\omega}_r$, the BRDF $f_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)$ specifies the amount of radiance incident on $\vec{x}$ along $\vec{\omega}_i$ that is reflected along $\vec{\omega}_r$ (figure 2.3). It is formally defined as the ratio of the reflected radiance $L_r(\vec{x}, \vec{\omega}_r)$ to the irradiance $E(\vec{x}, \vec{\omega}_i)$.

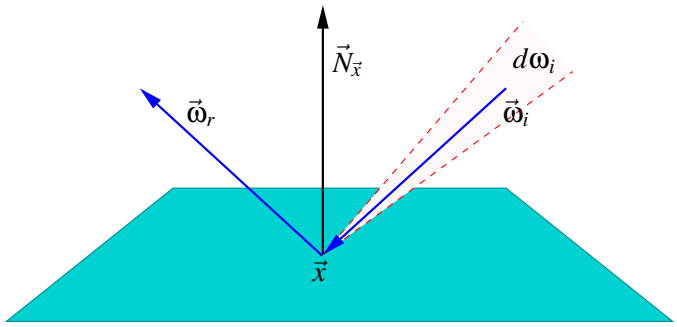


Figure 2.3: BRDF geometry

$$f_r(\vec{x},\vec{\omega}_i,\vec{\omega}_r) = \frac{L_r(\vec{x},\vec{\omega}_r)}{E(\vec{x},\vec{\omega}_i)} = \frac{dL_r(\vec{x},\vec{\omega}_r)}{L_i(\vec{x},\vec{\omega}_i)\,|\vec{\omega}_i\cdot\vec{N}_{\vec{x}}|d\omega_i}, \tag{2.7}$$

where $\vec{N}_{\vec{x}}$ is the surface normal at point $\vec{x}$ and $d\omega_i$ is a differential solid angle around $\vec{\omega}_i$. The counterpart for transmission is the *bidirectional transmission distribution function* (BTDF) f_t.

The *directional-hemispherical reflectance* ρ_{dh} is defined as the integral of the BRDF over all reflected directions for a given incident direction:

$$\rho_{dh}(\vec{x},\vec{\omega}_i) = \int_{\vec{\omega}_r\in\Omega_r} f_r(\vec{x},\vec{\omega}_i,\vec{\omega}_r)\,|\vec{\omega}_r\cdot\vec{N}_{\vec{x}}|d\omega_r, \tag{2.8}$$

where Ω_r is the set of all reflected directions in the hemisphere centered at $\vec{x}$. The *directional-hemispherical transmittance* τ_{dh} is defined similarly:

$$\tau_{dh}(\vec{x},\vec{\omega}_i) = \int_{\vec{\omega}_t\in\Omega_t} f_t(\vec{x},\vec{\omega}_i,\vec{\omega}_t)\,|\vec{\omega}_t\cdot\vec{N}_{\vec{x}}|d\omega_t. \tag{2.9}$$

In this document, these quantities will simply be referred to as the reflectance ρ and transmittance τ. Furthermore, unless noted otherwise, reflection and transmission are analogous and interchangeable for the rest of this discussion.

In order to be physically plausible a BRDF must meet the following requirements:

- Helmholtz reciprocity: the BRDF is invariant with respect to inversion of the direction of light transport, i.e.

$$f_r(\vec{x},\vec{\omega}_i,\vec{\omega}_r) = f_r(\vec{x},-\vec{\omega}_r,-\vec{\omega}_i) \tag{2.10}$$

- Conservation of energy: the radiance reflected in all directions can never exceed the irradiance, i.e.

$$\rho(\vec{x},\vec{\omega}_i) \leq 1, \qquad \forall\vec{\omega}_i\in\Omega_i \tag{2.11}$$

BRDFs can be represented in a number of ways:

- As analytical models [LFTG97, War92]
- As decompositions into basis functions, such as spherical harmonics [WAT92] or spherical wavelets [SS95].

While some analytical models are expensive to evaluate they generally allow a closed form representation to generate reflected directions for Monte Carlo sampling purposes. Decompositions into basis functions usually do not offer this luxury, but they are more general and can represent complex BRDFs which would be difficult to model by analytical means. Decomposition methods incur a tradeoff between accuracy and memory requirements, as they are capable of lossy compression.

BRDFs can be acquired from several sources, including:

- Theoretical models [HTSG91]
- Simulated microfacet geometry [WAT92]
- Measurements from actual material samples with a goniophotometer [AB95] or an imaging reflectometer [War92].

Measurents are the most useful BRDF source, however these are prone to teething problems such as noise and inherent limited precision on the part of the measuring device and the light sources used. Currently BRDF measurement is a laborious process. An overview of BRDF representation and acquisition techniques and their respective merits can be found in a survey by Rusinkiewicz [Rus97].

2.4 The EDF

The *emission distribution function* $L_e(\vec{x}, \vec{\omega}_e)$, or EDF, describes the distribution of radiance emitted from a point $\vec{x}$ on a light source in a direction $\vec{\omega}_e$. It can be considered as an analogon to the BRDF for light sources. It is related to the light source's radiant intensity I (whose distribution is usually available as goniometric data) as in equation 2.6.

Integrating the EDF over all points on the light source surface A and all emission directions Ω_e gives the total emitted flux Φ_e:

$$\Phi_e = \int_{\vec{x} \in A} \int_{\vec{\omega}_e \in \Omega_e} L_e(\vec{x}, \vec{\omega}_e) \cos\theta_e d\vec{\omega}_e dA. \tag{2.12}$$

2.5 The Rendering Equation

A formalised representation of the global illumination problem was proposed by Kajiya, known as the *rendering equation* [Kaj86]. It expresses the radiance $L_r(\vec{x}, \vec{\omega}_r)$ reflected by a point $\vec{x}$ on a surface along $\vec{\omega}_r$ in terms of the incident radiance from all surfaces in the scene. The radiance reflected from these latter surfaces is, in

turn, defined by the rendering equation, hence the equation provides a recursive definition of light transport. The radiance emitted by a surface is also included, thus catering for light sources.

$$L_r(\vec{x},\vec{\omega}_r) = L_e(\vec{x},\vec{\omega}_r) + \int_{\vec{\omega}_i\in\Omega_i} L_r(\vec{x},\vec{\omega}_i,\vec{\omega}_r)\,d\omega_i \tag{2.13}$$

$$= L_e(\vec{x},\vec{\omega}_r) + \int_{\vec{\omega}_i\in\Omega_i} f_r(\vec{x},\vec{\omega}_i,\vec{\omega}_r)\,L_i(\vec{x},\vec{\omega}_i)\,|\vec{\omega}_i\cdot\vec{N}_{\vec{x}}|d\omega_i \tag{2.14}$$

where:

- $L_e(\vec{x},\vec{\omega}_r)$ is the radiance emitted by the surface at point $\vec{x}$ along $\vec{\omega}_r$.
- $L_i(\vec{x},\vec{\omega}_i)$ is the radiance incident at $\vec{x}$ along direction $\vec{\omega}_i$. This incoming radiance is itself defined by recursively applying the equation.

The rendering equation forms the theoretical basis for global illumination methods. The goal of a global illumination algorithm is to solve this equation.

2.6 Raytracing

The bulk of global illumination algorithms is based on the *raytracing* paradigm originally proposed by Whitted in 1980 [Whi80], which is inspired by geometric optics. It is based on the assumption of point light transport in the form of infinitessimal rays along which the radiance L is invariant. Raytracing follows a recursive scheme in which *primary* rays are traced starting from the viewer followed by a series of *secondary* reflected or refracted rays from intersections with objects in the scene. For each intersected object, the irradiance due to light sources is evaluated and the radiance reflected according to the BRDF is accumulated and propagated along the ray back to the viewer. This is essentially the reversed photographic process; since rays are traced from the viewer back into the scene, the traditional raytracing algorithm is also termed *backward* raytracing to distinguish it from the *forward* raytracing algorithms introduced in the next chapter.

Raytracing's popularity stems from its ability to efficiently handle specular reflections due to its point light transport paradigm. The rendering equation translates literally into the raytracing algorithm. Raytracing solves the rendering equation numerically by distributing rays based on *Monte Carlo* methods.

2.6.1 Monte Carlo Methods

The rendering equation is an integral over all incident directions $\vec{\omega}_i$ in the hemisphere Ω_i centered at point $\vec{x}$. Each such direction gives rise to recursion, making an evaluation by analytical methods difficult.

Numerical methods such as Monte Carlo integration [HH64, Rub81] provide an alternative. Disregarding the emission term, the application of Monte Carlo methods to equation 2.14 consists of estimating the integral in the right hand term:

$$I = \int_{\vec{\omega}_i \in \Omega_i} L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)\, d\omega_i. \tag{2.15}$$

The essence of Monte Carlo lies in taking random samples $\vec{\omega}_i$ from the domain of incident directions Ω_i and recursively evaluating $L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)$ by tracing rays along $-\vec{\omega}_i$ to sample the environment. Formally, this is represented by a *primary estimator* $\hat{I}_1$ for the integral I:

$$\hat{I}_1 = L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r). \tag{2.16}$$

The distribution of the sample ray directions $\vec{\omega}_i$ is characterised by a *probability density function* (PDF) $p(\vec{\omega}_i)$ defined over Ω_i which specifies the likelihood for the choice of a particular direction to sample. Given this sample distribution, we can determine the *expected value* E of the estimator:

$$E\left[\hat{I}_1\right] = \int_{\vec{\omega}_i \in \Omega_i} L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)\, p(\vec{\omega}_i) d\omega_i. \tag{2.17}$$

The basic primary estimator uses uniformly distributed samples, i.e. $p(\vec{\omega}_i) = 1$, which implies that its expected value is equal to the integral I.

The accuracy of a Monte Carlo estimator can be expressed by the two quantities *bias* and *variance*. The bias β indicates the deviation of the estimator $\hat{I}_1$ from the actual integral I:

$$\beta\left[\hat{I}_1\right] = E\left[\hat{I}_1\right] - I. \tag{2.18}$$

Since $E[\hat{I}_1] = I$, the primary estimator is *unbiased*, which is of course optimal. If an estimator does introduce bias, it must be compensated for.

The variance V characterises the random noise in the estimator:

$$V\left[\hat{I}_1\right] = \sigma^2\left[\hat{I}_1\right] = E\left[\hat{I}_1^2\right] - E^2\left[\hat{I}_1\right] \tag{2.19}$$

$$= \int_{\vec{\omega}_i \in \Omega_i} L_r^2(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)\, d\omega_i - I^2, \tag{2.20}$$

where σ is the standard deviation. Obviously, the variance of the estimator is related to the variance of the illumination function L_r. For this reason, a uniform distribution of ω_i cannot handle specular components efficiently.

The primary estimator $\hat{I}_1$ based on a single sample will inevitably have a high variance, which can be reduced by averaging the primary estimators for N samples. Equivalently, we can break up the integral I into a sum of integrals I_j:

$$I = \int_{\vec{\omega}_i \in \Omega_i} L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)\, d\omega_i \tag{2.21}$$

$$= \sum_{j=1}^{N} \int_{\vec{\omega}_{ij} \in \Omega_i} \frac{L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r)}{N} d\omega_{ij} \tag{2.22}$$

$$= \sum_{j=1}^{N} I_j. \tag{2.23}$$

This leads to a *secondary estimator* $\hat{I}_2$ which is based on an average of primary estimators:

$$\hat{I}_2 = \sum_{j=1}^{N} \hat{I}_{1j} = \frac{1}{N} \sum_{j=1}^{N} L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r). \tag{2.24}$$

The expected value of this estimator is:

$$\begin{aligned} E\left[\hat{I}_2\right] &= E\left[\frac{1}{N} \sum_{j=1}^{N} L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r)\right] && (2.25)\\ &= \frac{1}{N} \sum_{j=1}^{N} E\left[L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r)\right] && (2.26)\\ &= \frac{1}{N} \sum_{j=1}^{N} E\left[\hat{I}_{1j}\right] && (2.27)\\ &= I, && (2.28) \end{aligned}$$

since $E[\hat{I}_{1j}] = I$, and hence this secondary estimator is also unbiased. The variance now becomes:

$$\begin{aligned} \sigma^2\left[\hat{I}_2\right] &= \sigma^2\left[\frac{1}{N} \sum_{j=1}^{N} L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r)\right] && (2.29)\\ &= \frac{1}{N^2} \sum_{j=1}^{N} \sigma^2\left[L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r)\right] && (2.30)\\ &= \frac{1}{N} \sigma^2\left[L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)\right] && (2.31)\\ &= \frac{1}{N} \sigma^2\left[\hat{I}_1\right]. && (2.32) \end{aligned}$$

Equation 2.32 is essential to classical Monte Carlo as it expresses the rate of convergence using N samples based on the primary estimator $\hat{I}_1$. From this we can see that the error is reduced by a factor of $1/\sqrt{N}$. This is a slow convergence rate, for which Monte Carlo methods in general are notorious. However, a notable feature of Monte Carlo integration is the fact that the rate of convergence is independent of the dimension of the integral. For this reason, Monte Carlo is also viable for high-dimensional integrals.

More sophisticated Monte Carlo methods exist which strive to reduce the variance of the primary estimator $\hat{I}_1$ by distributing the samples $\vec{\omega}_i$ more intelligently. Equation 2.32 serves as a basis to compare these methods. The major representatives found in practice are *stratified* and *importance sampling*.

2.6.1.1 Stratified Sampling

The Monte Carlo estimator $\hat{I}_2$ introduced in the last section has an implicit drawback: the samples may cluster in one part of the domain, while other (possibly

more significant) parts may not be sampled at all. To avoid this, we can divide the domain Ω_i of incident directions into smaller subdomains, or *strata*, and sample each in turn (figure 2.4). The underlying assumption here is that the illumination is approximately constant within each stratum, which is why stratified sampling is often employed to sample diffuse illumination.

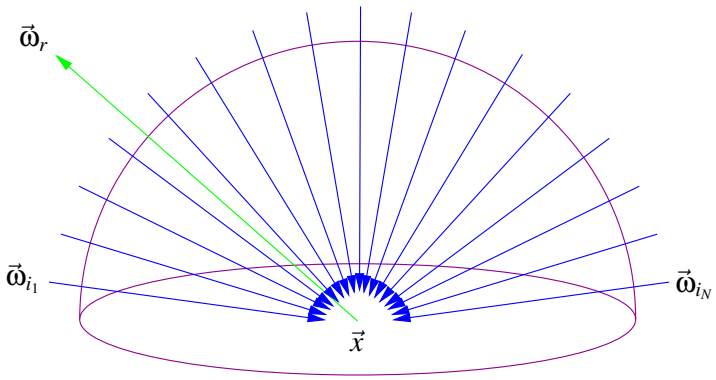


Figure 2.4: Stratified sampling

In basic stratified sampling, the strata have uniform size and are each represented by a single sample. We can break the integral I under consideration up into a sum of N subintegrals for each stratum Ω_{ij}:

$$I = \int_{\vec{\omega}_i \in \Omega_i} L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)\, d\omega_i \tag{2.33}$$

$$= \sum_{j=1}^{N} \int_{\vec{\omega}_{ij} \in \Omega_{ij}} L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r)\, d\omega_{ij} \tag{2.34}$$

$$= \sum_{j=1}^{N} I_j. \tag{2.35}$$

Assuming the samples are distributed according to the PDF $p(\vec{\omega}_{ij}) = N$, the unbiased primary estimator is:

$$\hat{I}_{1j} = \frac{L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r)}{N}, \qquad \vec{\omega}_{ij} \in \Omega_{ij}. \tag{2.36}$$

The secondary estimator is then simply the sum of the primary estimators:

$$\hat{I}_2 = \sum_{j=1}^{N} \hat{I}_{1j} = \frac{1}{N} \sum_{j=1}^{N} L_r(\vec{x}, \vec{\omega}_{ij}, \vec{\omega}_r). \tag{2.37}$$

While this is identical to equation 2.24, the sample distribution is different, since $\vec{\omega}_{ij} \in \Omega_{ij}$. The expected value is also identical to equation 2.28, and hence this estimator is also unbiased. However, the variance is now the sum of the primary

estimator variances:

$$\sigma^2\left[\hat{I}_2\right] = \sum_{j=1}^{N} \sigma^2\left[I_{1j}\right] \tag{2.38}$$

$$= \frac{1}{N^2}\sum_{j=1}^{N}\left[E\left[L_r^2\left(\vec{x},\vec{\omega}_{ij},\vec{\omega}_r\right)\right] - E^2\left[L_r\left(\vec{x},\vec{\omega}_{ij},\vec{\omega}_r\right)\right]\right] \tag{2.39}$$

$$= \frac{1}{N^2}\sum_{j=1}^{N}\left[\int_{\vec{\omega}_{ij}\in\Omega_{ij}} L_r^2\left(\vec{x},\vec{\omega}_{ij},\vec{\omega}_r\right) N d\omega_{ij} - \left[\int_{\vec{\omega}_{ij}\in\Omega_{ij}} L_r\left(\vec{x},\vec{\omega}_{ij},\vec{\omega}_r\right) N d\omega_{ij}\right]^2\right] \tag{2.40}$$

$$= \frac{1}{N}\int_{\vec{\omega}_i\in\Omega_i} L_r^2\left(\vec{x},\vec{\omega}_i,\vec{\omega}_r\right) d\omega_i - \sum_{j=1}^{N} I_j^2. \tag{2.41}$$

Since

$$\sum_{j=1}^{N} I_j^2 \geq \frac{1}{N} I^2, \tag{2.42}$$

this variance is lower than that of the standard secondary estimator (equation 2.32). This can be improved further by making the strata nonuniform, such that smaller strata (denser samples) are used where the illumination exhibits an increase in variance. This of course requires some knowledge of the illumination. If this knowledge is available, or the illumination is at least sufficiently predictable, we can apply importance sampling instead of stratified sampling.

2.6.1.2 Importance Sampling

A more effective approach to Monte Carlo sampling is to base the sample ray distribution on some a priori knowledge of the indirect illumination. Some regions of the incident hemisphere may be more important than others, particularly those which have high irradiance or high gradients compared to their neighbours. Concentrating samples in these regions can reduce the variance and lead to faster convergence, depending on the accuracy of the information available about the integral. This is the underlying concept of *importance sampling* [KW86] (figure 2.5).

The PDF $p(\vec{\omega}_i)$ provides the information required for the sampling density and should thus correlate with the incident indirect illumination. Since this knowledge is generally not available a priori, the surface's reflection characteristics are often used instead to approximate this information [BSS94, Lan91]. Specifically, the BRDF may serve as a PDF and guide the raytracing algorithm by associating high probabilities with those incident directions $\vec{\omega}_i$ which yield high reflectivity for the given reflecting direction $\vec{\omega}_r$. Importance sampling is therefore often applied to specularly reflected rays. However, this does not rule out situations in which the directional distributions of the BRDF and the unknown irradiance are vastly disparate, resulting in increased variance.

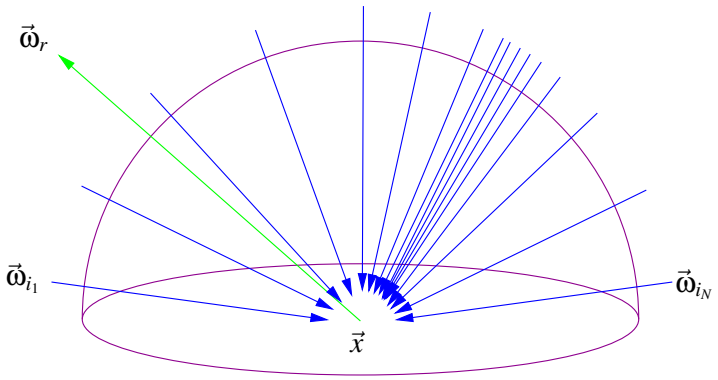


Figure 2.5: Importance sampling

Since the sample ray distribution is nonuniform, the primary estimator requires adaptation to avoid introducing bias. This is implied when rearranging the integral I under consideration:

$$I = \int_{\vec{\omega}_i \in \Omega_i} L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)\, d\omega_i = \int_{\vec{\omega}_i \in \Omega_i} \frac{L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)}{p(\vec{\omega}_i)} p(\vec{\omega}_i)\, d\omega_i. \tag{2.43}$$

The primary estimator $\hat{I}_1$ is then

$$\hat{I}_1 = \frac{L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)}{p(\vec{\omega}_i)}. \tag{2.44}$$

This estimator is unbiased, as its expected value is still equal to the integral I. The variance becomes

$$\sigma^2\left[\hat{I}_1\right] = \int_{\vec{\omega}_i \in \Omega_i} \left[\frac{L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)}{p(\vec{\omega}_i)}\right]^2 p(\vec{\omega}_i)\, d\omega_i - I^2 \tag{2.45}$$

$$= \int_{\vec{\omega}_i \in \Omega_i} \frac{L_r^2(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)}{p(\vec{\omega}_i)} d\omega_i - I^2. \tag{2.46}$$

Comparing this to equation 2.20 implies that the variance of importance sampling depends on the PDF; a good choice of PDF can reduce the variance compared to that of the standard primary estimator, while a poor choice can actually increase it. An optimal PDF would be $p(\vec{\omega}_i) = L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r)/I$, in which case the variance would be zero. However, since this requires the integral I this is utopian. Hence in practice the PDF must approximate I as closely as possible.

In order to apply importance sampling, the PDF must satisfy the following requirements:

- $p(\vec{\omega}_i) > 0$ when $L_r(\vec{x}, \vec{\omega}_i, \vec{\omega}_r) \neq 0$
- the PDF is normalised: $\int_{\vec{\omega}_i \in \Omega_i} p(\vec{\omega}_i) d\omega_i = 1$

- the PDF is invertible with respect to its *cumulative density* (or *cumulative distribution*) *function* (CDF) $P(\vec{\omega}_i)$, which is defined as:

$$P(\vec{\omega}_i) = \int_{\vec{\omega} \leq \vec{\omega}_i} p(\vec{\omega})\, d\omega. \tag{2.47}$$

  This assumes some ordering of the directions $\vec{\omega}$, e.g. by mapping their corresponding polar coordinates (θ, ϕ) to a one-dimensional index. The CDF effectively expresses the likelihood that samples below $\vec{\omega}_i$ will be chosen from the domain. Samples can be generated according to the PDF by evaluating $P^{-1}(\xi)$ for a uniform random number $\xi \in [0,1]$. Unfortunately, not all CDFs are invertible and thus not amenable to importance sampling.

Monte Carlo techniques have become the standard for physically based raytracers, but their slow convergence also led to the evolution of an alternative branch of physically based rendering called *Radiosity*.

2.7 Radiosity

Radiosity [GTGB84] is another popular global illumination method based on radiation transfer in a closed system. Its name is derived from the radiometric unit upon which it operates. It departs from the point transport approach of raytracing and assumes all reflections are diffuse only; specular reflection is entirely disregarded. While conceptually different from raytracing, Radiosity can be expressed in terms of the rendering equation, thus providing the fundamental link between the two approaches [CW93].

The method models diffuse interreflection between surfaces by setting up and solving a system of linear equations

$$B_{r_i} = B_{e_i} + \rho_i \sum_{j=1}^{N} B_{r_j} F_{ij} \qquad i \in [0, N], \tag{2.48}$$

where

- N is the number of surfaces in the scene
- B_{r_i} and B_{r_j} is the radiosity reflected from surface i and j, respectively
- B_{e_i} is the radiosity emitted from surface i
- ρ_i is the reflectance of surface i
- F_{ij} is the *form factor* characterising the exchange between surfaces i and j. This is a geometric term that specifies the fraction of the radiosity reflected from surface j that arrives at surface i. It takes into account the relative orientation, distance, and visibility of the two surfaces.

While standard Gaussian techniques can be applied in solving this system, the performance and storage requirements are substantial. Iterative methods which refine on an initial guess for B_{r_i} and ultimately converge provide a practical alternative [GCS94].

There are some advantages to Radiosity over raytracing: it can be supplemented by rendering hardware, and since the reflection is diffuse and therefore view-independent, the solution of the radiosity equation can be reused to create real-time, interactive walk-throughs. However, radiosity has the disadvantage of necessitating tesselation of curved objects into polygons, since the algorithm relies on planar surfaces which may also be adaptively subdivided into *patches* where high radiosity gradients are detected. This may introduce visible artifacts due to discontinuities at the seams. This tesselation cannot be applied to arbitrary geometries, in particular procedurally defined objects like fractals. Thus radiosity is restricted with respect to the scene geometries it can handle.

Classical radiosity has quadratic memory and time requirements due to the matrix representing the linear equations. The development of *progressive refinement* [CCWG88], which operates on one matrix column at a time and computes form factors on the fly, marked a turning point. More recent developments include hierarchical [HSA91] and clustering [SAG94] techniques, which have substantially improved radiosity's standing. Specular effects have also been supplemented via directional form factors [ICG86], but this once again incurs a substantial increase in storage requirements and is rarely used in practice.

Radiosity has established itself as a mainstream rendering method alongside raytracing and is still actively researched. It is a fast method capable of producing stunning images; the quality of diffuse reflections surpasses that of Monte Carlo raytracing methods. However, the lack of specular effects restricts its applicability. Daylight simulation requires specular components to model glare, precluding radiosity for this specific task.

2.8 Hybrid Renderers

The caveat that raytracing excels at specular and radiosity at diffuse effects leads to the assumption that a combination of these methods yields the best of both worlds. Hybrid methods exist which combine radiosity and raytracing [WCG87, SP89, RT90, SAWG91], however simply serializing the two methods does not account for all light transport paths. The more sophisticated hybrid algorithms do infact model all light paths, but they still require object tesselation which is plagued by the problems discussed above.

Chapter 3

Forward Raytracing

The classical raytracing approach by Whitted and its Monte Carlo derivates including RADIANCE share a common characteristic: they are all backward raytracers, i.e. they trace rays from the viewer to the light sources.

A major drawback of backward raytracing is the inability to adequately handle *caustics*. These are highly directional global illumination components which occur through specular reflections onto diffuse surfaces. They are often noticeable as iridescent highlights on otherwise dull materials from nearby specular objects. A popular example is the cardioid pattern seen in rings, glasses, mugs, etc (figure 3.1).

Figure 3.1: Photograph of metal ring caustic

Since a backward raytracer traces rays from the observer toward the light sources it cannot predict the occurrence of caustics, as this would imply knowledge of which paths contribute to the caustic and ultimately reach the light source (figure 3.2). Without such knowledge, rendering caustics with a backward raytracer

borders on an exercise in futility, as the chances of finding paths which contribute to a caustic are miniscule.

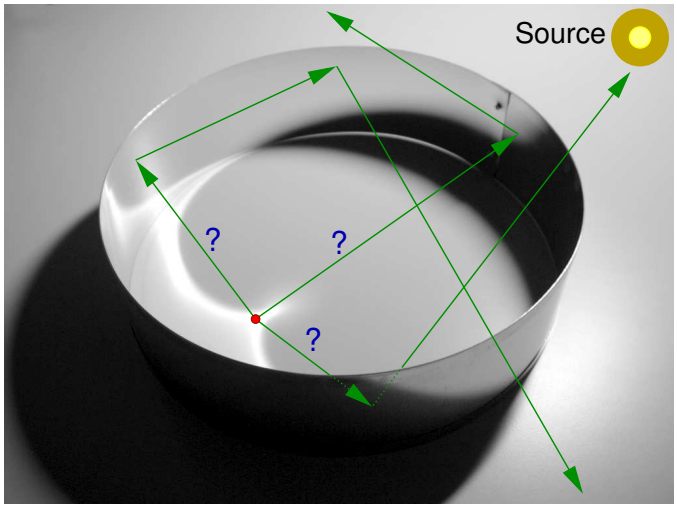


Figure 3.2: Rendering caustics with backward raytracing. A backward raytracer cannot predict which paths contribute to the caustic

In the case of RADIANCE, the caustic is only partially captured by those rays sampling the indirect diffuse component which happen to strike the metal ring. The left rendering in figure 3.3 is an example of the metal ring caustic using RADIANCE. The caustic is sampled by the stratified ambient rays and exhibits severe noise despite high rendering times. By contrast, reversing the backward raytracing process using a *forward raytracer* performs the task of sampling caustics far more efficiently. The right rendering in figure 3.3 is an example of forward raytracing applied to the metal ring, showing a clearly defined caustic. To put the efficiency issue into perspective: this rendering took under 8% of the computation time required for the RADIANCE rendering.

The idea behind forward raytracing is simple: to complement the backward raytracing process starting from the viewer with a forward pass starting from the light sources and coupling them by some means. In essence this connects light transport paths originating at the light sources with those terminating at the viewer, thus accounting for all possible paths (figure 3.4). This is a more natural and intuitive approach since in reality light is propagated from the sources, not the observer. In general, forward raytracers perform a separate forward pass in object space followed by a backward pass in image space, but they differ in the methods used to link the two. Most existing physically based backward raytracers can be supplemented with a forward pass in a fairly straightforward manner. The computation of direct illumination is not affected by this modification.

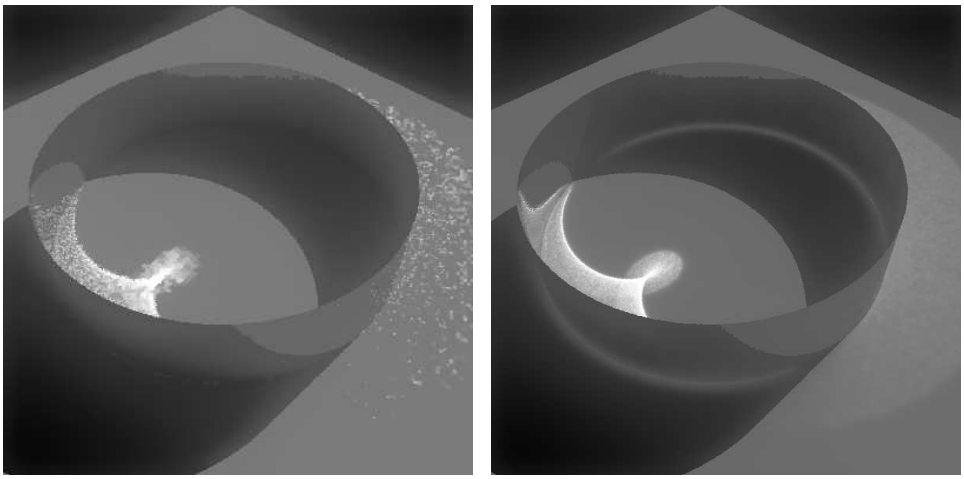


Figure 3.3: Metal ring caustic rendered with RADIANCE's standard backward raytracing algorithm (left) and forward raytracer (right).

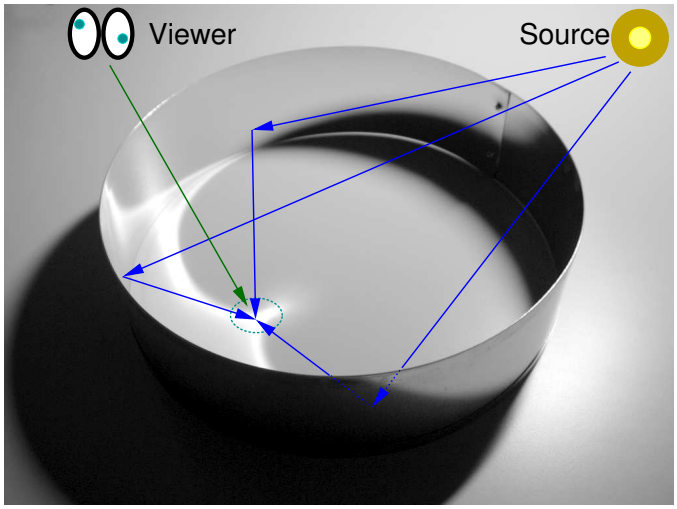


Figure 3.4: Linking forward and backward raytracing

3.1 The Photon Map

The *photon map* developed by Wann Jensen [JC95, Jen96, JC00] is a forward raytracing algorithm that has gained considerable popularity. It differs from other forward raytracing methods (notably the geometry-bound method developed by Shirley and Walter et al [SWH+95, WHSG97]) by decoupling the information generated by the forward raytracer from the geometry. Wann Jensen's primary intent was to isolate the photon map from the objects in the scene so that it may be applied to arbitrary geometries, such as fractal topologies and even volumetric data such as participating media [JC98]. This is a clear advantage over geometrically more con-

strained finite element methods such as Radiosity. A further consequence of this separation is that the photon map scales well to the complexity of the geometry [Suy02, p.106].

The method is based on a Monte Carlo simulation of light particle (photon) transport in the rendered scene, from which it derives its name. The forward pass is a preprocess and consists of distributing photons from the light sources, while the backward (rendering) pass performs lookups in the photon map to evaluate the indirect illumination. This effectively establishes a link between the forward and backward raytracers and takes all light transport paths into account.

3.1.1 Forward Pass

In the forward pass photons are emitted from each light source into the scene and traced as they are scattered at the surfaces. Upon emission, each photon is assigned an initial flux, direction, and origin on the light source. The photon is then traced as it undergoes a series of intersections with objects in the scene. Upon striking a surface, it is stored along with its flux, 3D location (intersection point), and surface normal at that location in a space subdividing data structure for efficient lookup. No reference to the intersected object is made, thus decoupling the photons from the geometry; photons just dangle in space. Once it has been stored, the photon is subjected to a probabilistic scattering event, the outcome of which is either diffuse scattering[1], specular scattering, or absorbtion. If the photon is scattered, it can interact with the scene and still contribute to indirect illumination. If it is absorbed, the path is terminated probabilistically (a statistical process aptly named *russian roulette*) and a new photon is emitted.

The photon map's contribution to indirect illumination is governed by the photon flux and density. Standard particle tracing methods modulate both attributes, but a constrained variant can also be used which maintains a constant photon flux in order to reduce the variance of the reconstructed illumination.

3.1.1.1 Particle tracing

A general procedure for particle tracing entails the following steps:

Emission: photons are emitted according to a PDF p_e based on the position on the light source $\vec{x}$ and the emitted direction $\vec{\omega}_e$. The emitted photon flux Φ_p is then proportional to the EDF L_e [Suy02, p.28]:

$$\Phi_p(\vec{x},\vec{\omega}_e) = \frac{L_e(\vec{x},\vec{\omega}_e)\cos\theta_e}{N_e p_e(\vec{x},\vec{\omega}_e)} dA d\omega_e = \frac{d\Phi_e(\vec{x},\vec{\omega}_e)}{N_e p_e(\vec{x},\vec{\omega}_e)}, \tag{3.1}$$

where N_e is the number of emitted photons.

[1] The term *scattering* is used here to generalise reflection and transmission.

Absorbtion: photons could be absorbed once their flux falls below some threshold. While this seems intuitive, it introduces the risk of bias. A workaround is to apply *russian roulette* and absorb photons with a probability p_a while bloating the flux of those which survive during scattering (see below), thus compensating for the photons which were terminated [AK90]. It can be shown that this does not alter the expected value of the photon flux and that the method is therefore unbiased [Jen01, p.62].

Scattering: photons are scattered according to a PDF p_r. The photon flux Φ_p is attenuated by the surface's BRDF f_r and weighted by a factor $\frac{1}{p_a}$ to compensate for russian roulette [Suy02, p.28]:

$$\Phi_p(\vec{\omega}_r) = \Phi_p(\vec{\omega}_i)\frac{f_r(\vec{\omega}_i,\vec{\omega}_r)\cos\theta_i}{p_a p_r(\vec{\omega}_i,\vec{\omega}_r)}. \tag{3.2}$$

This general approach has the advantage that it employs arbitrary emission and scattering PDFs. This property is particularly useful when exact Monte Carlo inversion of the EDF and BRDF is not possible, thus precluding importance sampling. In the simplest case the PDFs can be uniform, yielding a stratified sampling scheme which is particularly easy to implement.

The drawback with this approach is that the variable flux increases the variance in the reconstructed illumination [Jen01, p.64], since on a receiving surface both flux and density will fluctuate (even with uniform sampling), and the reconstructed illumination inherits this variance as implied by equation 2.20.

3.1.1.2 Constant Photon Flux

Particle tracing with constant flux is a constrained case of the above in which all photons share a uniform flux Φ_p, while the density is variable. This method is also known as *analog simulation* [Suy02, p.28]. It differs from the general approach in the following steps:

Emission: the emission PDF p_e is based on the emitted radiance distribution L_e to concentrate photons in regions of high emission, thereby implementing an importance sampling distribution. The constant photon flux Φ_p is derived from the total flux Φ_e emitted from all light sources [Suy02, p.28]:

$$\Phi_p = \frac{L_e(\vec{x},\vec{\omega}_e)\cos\theta_e}{N_e p_e(\vec{x},\vec{\omega}_e)} dA d\omega_e \tag{3.3}$$

$$= \frac{d\Phi_e(\vec{x},\vec{\omega}_e)}{N_e \frac{d\Phi_e(\vec{x},\vec{\omega}_e)}{\int_{\vec{x}\in A}\int_{\vec{\omega}_e\in\Omega_e} d\Phi_e(\vec{x},\vec{\omega}_e)}} \tag{3.4}$$

$$= \frac{\Phi_e}{N_e}. \tag{3.5}$$

Absorption: the photon density is attenuated according to the surface's reflectance ρ by absorbing photons with a probability $p_a = \rho$.

Scattering: the scattering PDF p_r is based on the surface's BRDF f_r to generate a new photon direction, thereby also implementing an importance sampling distribution as for emission. The photon flux Φ_p is not modified [Suy02, p.29]:

$$\begin{aligned} \Phi_p &= \Phi_p \frac{f_r(\vec{\omega}_i, \vec{\omega}_r)\cos\theta_i}{p_a p_r(\vec{\omega}_i, \vec{\omega}_r)} && (3.6)\\ &= \Phi_p \frac{f_r(\vec{\omega}_i, \vec{\omega}_r)\cos\theta_i}{\rho(\vec{\omega}_i)\frac{f_r(\vec{\omega}_i,\vec{\omega}_r)\cos\theta_i}{\rho(\vec{\omega})}} && (3.7)\\ &= \Phi_p. && (3.8) \end{aligned}$$

This approach is used in most photon map implementations, and is preferable to the unconstrained particle tracing paradigm because the flux is uniform at receiving surfaces and only the density fluctuates. Consequently, the reconstructed illumination will exhibit less variance [Suy02, p.28,112]. The price to pay is a higher implementational and computational overhead due to Monte Carlo inversion of the BRDF and EDF.

While spectral effects are easily accounted for with variable flux, this is not straightforward with a constant flux. To overcome this difficulty, we can extend the concept by splitting the constant photon flux Φ_p into a set of components $\Phi_{p,c}$ for each colour channel c such that their average is Φ_p [Jen01, p.63]. This of course implies that the components themselves are variable and can introduce some spectral variance in the reconstructed illumination. We maintain a constant average flux Φ_p by normalising and scaling the photon flux at every emission and scattering event[2]. The particle tracing steps are adapted to this convention as follows:

Emission: the emitted photon's flux components are weighted by the spectral EDF $L_{e,c}$:

$$\Phi_{p,c}(\vec{x}, \vec{\omega}_e) = \Phi_p \mathrm{norm}(L_{e,c}(\vec{x}, \vec{\omega}_e)), \quad (3.9)$$

with the spectral normalisation factor defined as

$$\mathrm{norm}(g_c) = \frac{g_c}{\overline{g}}, \quad (3.10)$$

where $\overline{g}$ denotes the average over all colour channels.

Scattering: the scattered photon's flux components are weighted by the incident flux modified by the spectral BRDF $f_{r,c}$:

$$\Phi_{p,c}(\vec{\omega}_r) = \Phi_p \mathrm{norm}(\Phi_{p,c}(\vec{\omega}_i) f_{r,c}(\vec{\omega}_i, \vec{\omega}_r)). \quad (3.11)$$

[2] Depending on the implementation, this scaling can only be done subsequent to completion of the photon distribution pass, since the number of emitted photons N_e and therefore the uniform photon flux Φ_p are unknown up to that point.

3.1.1.3 Photon Types

Wann Jensen's original proposal included a separate *global* and *caustic* photon map. The former is dedicated to photons accounting for global diffusely reflected illumination (specularly reflected global illumination is sampled separately using standard Monte Carlo methods). The latter is a subset of the global map accounting only for caustics. Both photon maps differ in the means by which their illumination is evaluated. The caustic photon density is usually much higher than that of global photons to account for detail in the caustics, since these effects tend to have higher gradients than indirect diffuse illumination.

A global photon is stored in the global map on every indirect intersection (i.e. excluding those occuring immediately after emission) with a diffuse[3] surface. Using Heckbert's regular expression notation to classify transport path vertices [Hec90], global photons account for $L(D|S)^+D$ paths [Jen01, p.99] [Suy02, p.37], where L denotes the light source, and D and S denote diffuse resp. specular scattering. Note that these paths start at the light sources and terminate on diffuse surfaces, thus characterising a forward raytracer. It is the backward pass that supplies the viewpoint, or eye vertex (denoted E in Heckbert's notation).

If a diffuse surface is struck immediately following a series of specular scattering events, a caustic photon is stored in addition to the global photon[4]. This translates to paths of the form LS^+D [Jen01, p.97] [Suy02, p.37].

Christensen [Chr00] suggested extending this to account for *secondary* caustics: if reradiated by diffuse scattering, the photon can form additional, albeit less intense, caustics after subsequent specular scattering events. As he claims, these can also lead to significant contributions to the indirect illumination. Suykens [Suy02, p.48] shows an example of this, referring to the phenomenon as *indirect* caustics. These caustics are characterised by paths of the form $L(D|S)^*SD$. The remaining discussion takes secondary caustics into consideration.

Figure 3.5 summarizes the photon distribution process in flowchart form. Figure 3.6 depicts an example of photon distribution in a simple Cornell box with a glass sphere.

3.1.2 Backward Pass

The backward pass traces rays from the viewpoint, or eye vertex E, towards the objects in the scene. It links to the forward pass at the intersected objects, thereby completing the transport paths computed by the latter.

[3] In this document, the terms *diffuse* and *specular* refer to the presence of such components in the BRDF. They do not necessarily imply lambertian resp. mirror reflection. Diffuse or specular scattering is subject to the relative magnitudes of these components and thus stochastic. Terms like *glossy* or *directional diffuse* are not explicitly used here for simplicity, since they are implied by combinations of diffuse and specular.

[4] This 1:1 correspondence is somewhat simplified, but serves to illustrate the principle. In practice, a mechanism is used to control the relative densities of the two photon types, and caustic and global photons are selectively stored based on their density requirements.

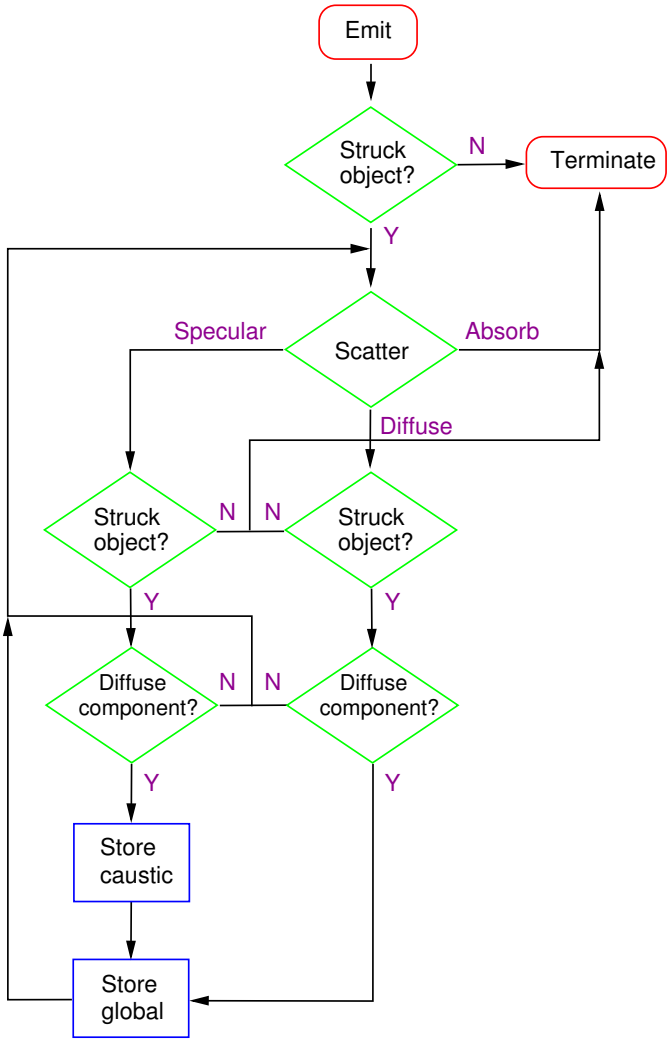


Figure 3.5: Photon distribution flowchart. The distribution step is driven by the outcome of scattering events, which can be either diffuse scattering, specular scattering, or absorbtion.

The backward pass uses standard Monte Carlo backward raytracing for the direct illumination (LE and $L(D|S)E$) and primary specular reflections (LS^+E), but the photon map for all indirect illumination on diffuse surfaces with partial path (DS^*E), i.e. those seen directly by the eyepoint or via a series of primary specular reflections[5] [Jen01, p.85].

Contributions from the photon map are obtained by *nearest neighbour search*, i.e. locating a number of nearby photons in the neighbourhood of the point whose illumination is under consideration (figure 3.7). The underlying data structure for

[5]Note that these paths are evaluated from right to left in the backward pass.

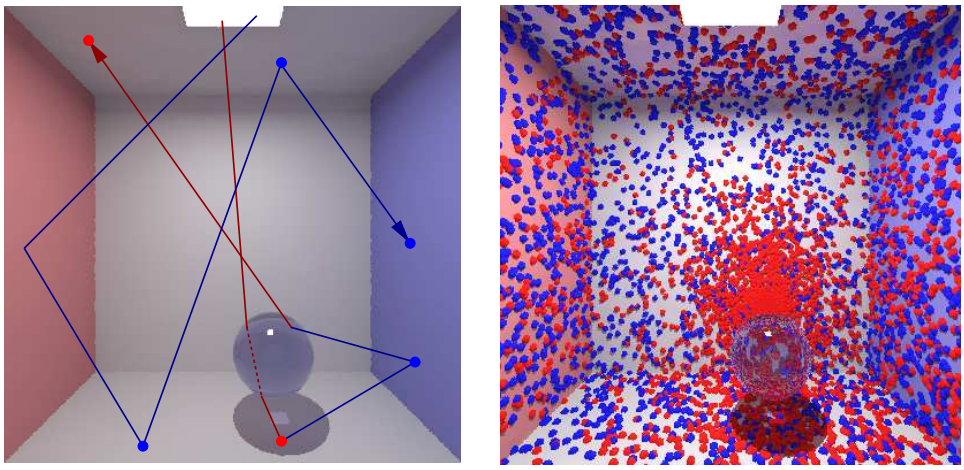


Figure 3.6: Global photon distribution in the Cornell box. Left: global and caustic photon paths during forward pass. Dots indicate stored photons, where blue represents global photons, and red represents caustic photons. Right: photon distribution after completion of forward pass. Note the concentration of caustic photons on the back wall reflected from the glass sphere.

this search is explained in section 3.1.4 below. The resulting irradiance from the photons is evaluated using a technique known as *density estimation* covered in section 3.1.3.

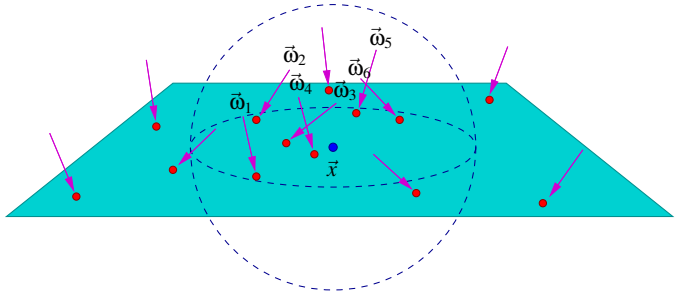


Figure 3.7: Photon gathering via nearest neighbour search

3.1.2.1 Direct Visualisation

The simplest variant of rendering photon contributions is direct visualisation [Suy02, p.39]. The global photon map (which also contains caustics, albeit typically of lower density) is visualised directly on all diffuse surfaces visible at primary rays or via primary specular reflections. Linking the corresponding partial paths from the forward and backward raytracers results in complete paths of type $L(D|S)^{+}DS^{*}E$. A separate caustics photon map is not even needed, since it would introduce redundant light paths. While this visualisation strategy is fast, the inherent noise in global

photon irradiance is evident. Unless a very large number of global photons is used, the quality of the resulting images will be poor.

3.1.2.2 Final Gathering

A more sophisticated approach to photon map visualisation is to mask the noise in the global photon irradiance by relaying its evaluation via an intermediate (usually stratified) diffuse sampling step, adding a secondary diffuse reflection to the partial transport paths provided by the backward pass (i.e. DDS^*E). This approach, known as *final gathering* [Suy02, p.39,115], is inherently slower than direct visualisation, but improves the quality of the indirect illumination on diffuse surfaces substantially. Combining partial forward and backward paths, global photons account for $L(D|S)^+DDS^*E$. While this component is evaluated indirectly, the irradiance from caustic photons is evaluated directly at the primary diffuse reflection, accounting for combined paths of type $L(D|S)^*SDS^*E$. Figure 3.8 illustrates final gathering in schematic form.

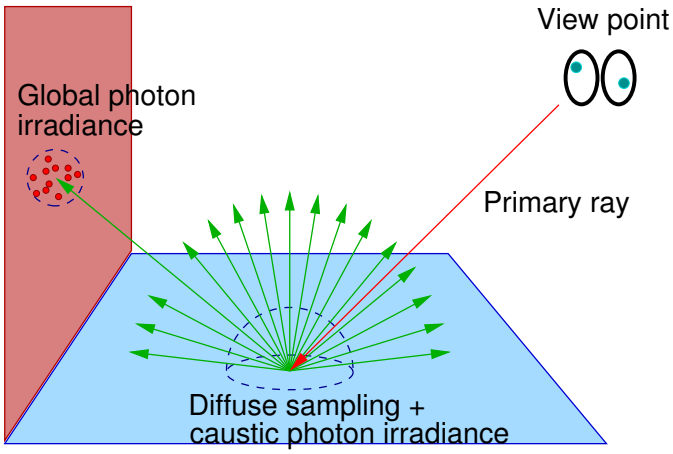


Figure 3.8: Photon visualisation during backward pass using final gathering. Caustic photons are visualised at the primary ray, while global photons are visualised via an intermediate diffuse sampling step.

Having formulated the paths associated with each component, we proceed to verify that the sum of all components actually accounts for all transport paths (see also [Suy02, p.116] for a similar breakdown). Combining the contributions from global and caustic photons, we obtain:

$$(L(D|S)^+DDS^*E) \quad | \quad (L(D|S)^*SDS^*E) \quad = \quad L(D|S)^+DS^*E, \tag{3.12}$$

i.e. all indirect illumination on diffuse surfaces visible directly or via specular reflections is accounted for. We now add the primary specular reflections obtained with standard Monte Carlo backward raytracing:

$$(L(D|S)^+DS^*E) \quad | \quad (LS^+E) \quad = \quad L(D|S)^+E. \tag{3.13}$$

All indirect illumination is therefore accounted for. Finally, we add the direct illumination obtained by explicitly sampling the light sources:

$$(L(D|S)^{+}E) \quad | \quad (L(D|S)E) \quad | \quad (LE) \quad = \quad L(D|S)^{*}E. \tag{3.14}$$

This implies that *all* possible transport paths are accounted for, confirming – in theory – that the photon map coupled with final gathering is a full global illumination solution.

It is important to suppress redundant light paths during the backward pass [Suy02, p.39]. This is accomplished by truncating all backward paths beyond the primary diffuse reflection, i.e. limiting these to $DS^{*}E$. Failure to do so would result in overcounting and introduce bias in the solution. The fact that these paths can be neglected during the backward pass even constitutes a significant bonus: it can result in a dramatic speedup over a standard backward raytracer, since it prunes all branches below the first diffuse reflection in the ray tree.

3.1.2.3 Importance Guided Final Gathering

Instead of using stratified sampling for the intermediate diffuse reflection, Wann Jensen proposed distributing the sample rays based on importance obtained from the global photons [Jen95]. In this variant, the global photon map supplies an estimate of the spatial distribution of the indirect illumination to guide the sampling process by identifying those directions which actually yield high indirect irradiance, effectively amounting to importance sampling. This requires tabulating the accumulated photon flux according to the photons' incident directions. While this approach is more sophisticated than simple stratified sampling, there is a high computational overhead incurred by constructing and inverting the PDF for every primary diffuse intersection. Futhermore, the photon incident directions require additional storage.

3.1.2.4 Precomputed Photon Irradiance

Per Christensen proposed a modification to the photon map using precomputed irradiance from global photons [Chr00, Jen01]. Christensen realised that there is considerable spatial overlap in photon map lookups during the backward pass, implying a considerable degree of redundancy in density estimates. His proposal was therefore to precompute the global photon irradiance at the photon locations and to store it with the photons in place of the photon flux. He suggested precomputing about $\frac{1}{4}$ of the global photons while discarding the rest after precomputation. With this approach, Christensen reports speedups by a factor of 5–8.

During the backward pass global photons are visualised by looking up the single closest global photon and using its irradiance. This effectively forms a Voronoi diagram with each cell shaded with the irradiance from its associated precomputed photon (figure 3.9). This piecewise constant illumination is a valid approximation under the assumption that this irradiance has a low gradient. This optimisation is

therefore not suitable for caustic photons, since their irradiance typically has a high gradient.

Figure 3.9: Precomputed global photon irradiance. Note the caustic from the glass sphere.

3.1.3 Density Estimation

The irradiance contributed by the photons is reconstructed based on their incident flux and density [Jen97] by applying a technique known as *density estimation*. Given the N_p nearest photons found around the point $\vec{x}$ under consideration, the density estimate $\hat{f}(\vec{x}, N_p)$ can be derived from the definition of the irradiance $E(\vec{x})$ under the assumption that each photon contributes flux $d\Phi$ along its incident direction $\vec{\omega}_j$:

$$E(\vec{x}) = \int_{\vec{\omega}_j \in \Omega_i} \frac{d^2\Phi(\vec{\omega}_j)}{dA d\omega} d\omega \tag{3.15}$$

$$\approx \sum_{j=1}^{N_p} K(\|\vec{x} - \vec{x}_j\|)\, d\Phi(\vec{\omega}_j) \tag{3.16}$$

$$= \hat{f}(\vec{x}, N_p), \tag{3.17}$$

where K is a *kernel function* used to weight the photons based on their distance from $\vec{x}$. The number of photons N_p used for the density estimate is often termed the *kernel width* or *bandwidth*, as it defines the domain on which the kernel operates. The kernel function K assigns weights based on the proximity of the photons in order to reduce blurring in the reconstructed irradiance. This function is radially

symmetric and normalised over the domain $[0, r_p]$:

$$\int_0^{r_p} K(s)\, ds = 1, \tag{3.18}$$

where r_p is the maximum distance from $\vec{x}$ to the N_p photons. In 2D density estimates, the normalisation factor accounts for the surface area covered by the kernel's bandwidth, which is assumed to be planar. Asymptotically this will be the case as the photon density approaches infinity.

Typical 2D kernels for density estimation on surfaces include:

- The uniform kernel
$$K_0(s) = \frac{1}{\pi r_p^2}, \tag{3.19}$$
which assigns a constant weight to all photons and consequently does not compensate at all for blurring.

- The cone filter
$$K_c(s) = \frac{1 - \frac{s}{k r_p}}{\pi r_p^2 \left(1 - \frac{2}{3k}\right)} \qquad k \geq 1, \tag{3.20}$$
where k is a constant defining the filter slope. This weighting method was suggested by Wann Jensen [Jen97, Jen01].

- The 2D Epanechnikov kernel
$$K_e(s) = \frac{2}{\pi r_p^2}\left[1 - \left[\frac{s}{r_p}\right]^2\right], \tag{3.21}$$
which Walter et al [WHSG97] employed in their geometry-bound particle tracing method.

- The 2D Silverman kernel
$$K_s(s) = \frac{3}{\pi r_p^2}\left[1 - \left[\frac{s}{r_p}\right]^2\right]^2, \tag{3.22}$$
used by Shirley et al [SWH$^+$95] in their particle tracing implementation. This kernel has higher-order smoothness properties than the Epanechnikov kernel, which is desirable as they are inherited by the density estimate.

These kernels are plotted in figure 3.10. While this document focuses specifically on 2D density estimation, generalised forms of the kernels for higher dimensional cases (e.g. volume photon visualisation) can be found in [Sil86].

Density estimation is a straightforward means of reconstructing photon irradiance, but the blurring effect which the kernels try to suppress is its inherent drawback. It affects the renderings not only visually, but also quantitatively, implying the potential for bias in the reconstructed irradiance. This will be elaborated on in chapter 5.

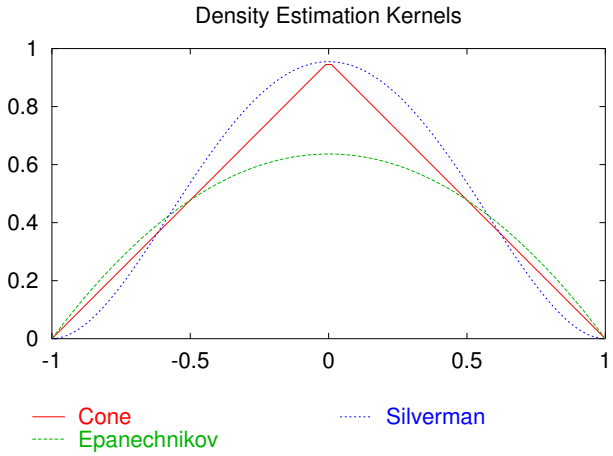


Figure 3.10: Density estimation kernels

3.1.4 Data Structure

The photon map is implemented as a spatial data structure which maintains a set of 3D keys corresponding to the photon positions which are sorted by space subdivision. It must be capable of efficient retrieval, particularly nearest neighbour queries [FBF77]. Several such data structures exist for geometric databases. Wann Jensen applied the *k-d tree* [Ben75, Ben79, BF79] as it is one of the more general spatial data structures, being a logical extension of the conventional binary search tree.

Each node S in the tree has a k-dimensional key (where $k = 3$ in this application) $(s_0, \ldots, s_{k-1})$ and a *discriminator* $d \in [0, k-1]$ such that:

- $l_d < s_d$ for every node L in the left subtree with key $(l_0, \ldots, l_{k-1})$, and
- $r_d > s_d$ for every node R in the right subtree with key $(r_0, \ldots, r_{k-1})$.

Equality between s_d and another key in dimension d is resolved by cyclic comparison in each dimension. This implies that no two nodes in the tree may have identical keys which match in all dimensions.

The discriminator can be graphically interpreted as subdividing the space containing all nodes below S in dimension d.

As with many tree structures, balancing guarantees logarithmic search times and is imperative for optimum performance. This is particularly viable for the photon map since it remains static once the forward pass is complete.

Chapter 4

The RADIANCE Photon Map

This chapter describes implementation details of the forward ray tracing module embedded in the RADIANCE package which are specific to daylight simulation. The photon map module will be referred to as the *RADIANCE photon map*, while RADIANCE's standard backward raytracing method will be referred to as *RADIANCE Classic*. Results obtained with the module are presented in the form of daylight simulation examples at the end of the chapter.

4.1 Implementation

The RADIANCE photon map[1] is implemented as two components; the forward pass is a separate preprocessor (*mkpmap*) which produces a portable photon heap file (essentially a compact balanced k-d tree), while the backward pass is incorporated into the original RADIANCE ray tracer (*rpict, rtrace, rview*) and reads the photon heap. This allows reusing the distribution for a series of renderings in which the geometry remains static, e.g. walkthroughs.

The implementation employs a number of optimisations mentioned in chapter 3. The forward pass uses constant photon flux extended to the red, green, and blue colour channels (cf. section 3.1.1.2) and accounts for secondary caustics (cf. section 3.1.1.3). The backward pass can visualise photons either directly or via final gathering (cf. section 3.1.2.2).

Apart from global and caustic photons, the RADIANCE photon map also supports volume photons as described in [JC98], which can be used with the RADIANCE *mist* primitive.

4.1.1 Photon Distribution

In order to meet the requirements of lighting simulation, the photon distribution step must be capable of accepting arbitrary EDFs for light sources, and emit photons

[1] Available at `http://www.ise.fhg.de/radiance/photon-map`

accordingly. Furthermore, the distribution step must be capable of photon emission from distant light sources used specifically for daylight.

The forward pass emits N_e photons from all light sources in the scene. To accurately simulate the emission characteristics of a source, an explicit integration over finite elements on the light source surfaces is performed to obtain the emitted flux. The reason for this overhead arises from the necessity to account for EDFs which may not only vary in direction, but also in location. Light sources with uniform emission over the surface and emitting angle (which are unrealistic) of course don't require this overhead, and can be optimised with a trivial analytical derivation of the emitted flux.

The total flux Φ_e emitted from all sources is determined by integrating the EDF $L_e(\vec{x},\vec{\omega}_e)$ over the set of exitant directions Ω_e and over the source surfaces A (cf. equation 2.12):

$$\Phi_e = \int_{\vec{x}\in A}\int_{\vec{\omega}_e\in\Omega_e} L_e\left(\vec{x},\vec{\omega}_e\right)\left|\vec{\omega}_e\cdot\vec{N}_{\vec{x}}\right| d\vec{\omega}_e dA. \tag{4.1}$$

We can approximate the intergral by a Monte Carlo estimator $\hat{\Phi}_e$ using partitions ΔA of approximately equal size on the light source surfaces generated with RADIANCE's light source sampling code (normally used to simulate penumbrae). For each partition i, we choose a random point $\vec{x}_i$ within the partition and integrate $L_e(\vec{x}_i,\vec{\omega}_e)$ to yield the flux $\hat{\Phi}_{e,i}$ emitted from the partition:

$$\begin{aligned}
\hat{\Phi}_e &= \sum_{i=1}^{N_{\Delta A}}\sum_{\vec{\omega}_e\in\Omega_e} L_e\left(\vec{x}_i,\vec{\omega}_e\right)\left|\vec{\omega}_e\cdot\vec{N}_{\vec{x}_i}\right|\Delta\vec{\omega}_e\Delta A && (4.2)\\
&= \sum_{i=1}^{N_{\Delta A}}\sum_{\vec{\omega}_e\in\Omega_e} \Delta\hat{\Phi}_e\left(\vec{x}_i,\Delta\vec{\omega}_e\right) && (4.3)\\
&= \sum_{i=1}^{N_{\Delta A}} \Delta\hat{\Phi}_{e,i}, && (4.4)
\end{aligned}$$

where $N_{\Delta A}$ is the number of partitions. Highly peaked distributions may require very small surface partitions ΔA or differential solid angles $\Delta\vec{\omega}_e$. Furthermore, Ω_e may be arbitrary, ranging from very small solid angles to an entire sphere for a distant source providing uniform illumination from all directions[2].

Once we have obained the source flux through integration, we can determine the constant photon flux Φ_p (cf. equation 3.5):

$$\Phi_p = \frac{\hat{\Phi}_e}{N_e}. \tag{4.5}$$

Photons are emitted from each partition in turn. We choose a new origin within the current partition for each photon in order to break up clustering artifacts. The

[2]Such a source would be unrealistic, but can nevertheless be modeled in RADIANCE. The maximum solid angle encountered in practice would correspond to a hemisphere such as used for sky sources.

number of photons $N_{e,i}$ to emit from source partition i is then proportional to the partition's emitted flux:

$$N_{e,i} = N_e \frac{\Delta\overline{\Phi}_{e,i}}{\overline{\Phi}_e}, \tag{4.6}$$

where $\overline{\Phi}$ denotes the flux averaged over all colour channels.

Each photon is emitted according to an importance sampling PDF p_e proportional to the differential flux emitted along $\vec{\omega}_e$ (cf. equation 3.4):

$$p_e\left(\vec{x}_i, \vec{\omega}_e\right) \propto \Delta\hat{\Phi}_e\left(\vec{x}_i, \Delta\vec{\omega}_e\right). \tag{4.7}$$

The EDF can be specified as an arbitrary function or data in RADIANCE, but the associated CDF (cumulative density function) P is not readily invertible by analytical means for Monte Carlo sampling. A standard numeric inversion method consists of constructing a lookup table T in which each index j is associated with a differential solid angle $\Delta\vec{\omega}_{e,j}$ and $T(j) = P(\Delta\vec{\omega}_{e,j})$. A photon emission direction is then obtained by finding the array index j such that $T(j-1) < \xi \leq T(j)$ for a uniform random variable $\xi \in [0,1]$, and emitting into $\Delta\vec{\omega}_{e,j}$. See also [Jen95] for further details. This table lookup is inherently slow, but can be substantially accelerated by applying a binary search.

4.1.2 Local vs. Distant Sources

The photon emission from a local source is a straightforward application of the above, since the light sources are part of the scene geometry. Figure 4.1 is an example of emission from a partitioned local polygonal source.

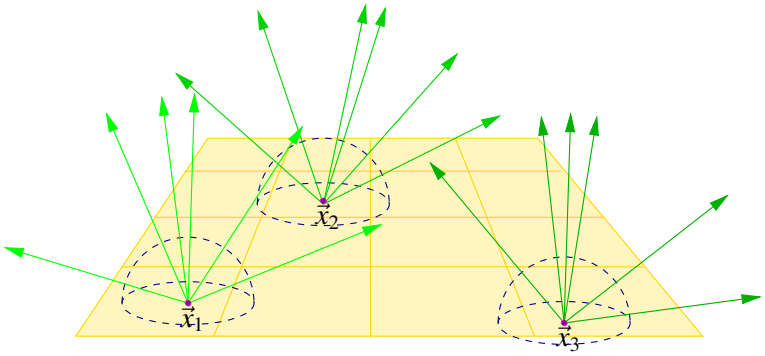


Figure 4.1: Photon emission from local light source

In contrast, distant sources such as those used for the solar and sky sources in daylight simulation (see figure 1.4) require a different treatment with the photon map. A distant source has no associated geometry contained in the scene, thus emission from its surface is impossible. A workaround is to reciprocate the above methodology and partition the faces of the bounding cube containing the entire scene and integrate the incident emission on them. Photons are then emitted into

the scene from these partitions within the source's solid angle [Jen01, p.58] (figure 4.2).

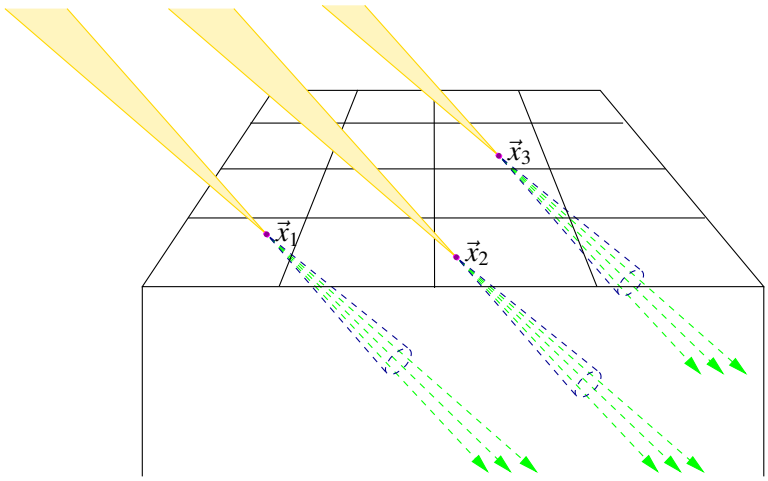


Figure 4.2: Photon emission from partitioned scene cube face for distant light source

The explicit flux integration need only be performed once for each scene cube face, since distant sources are invariant with respect to location. Thus all partitions on a scene cube face have the same incident flux.

For large solid angles, this brute force approach can become extremely inefficient. This comes as no surprise, since it creates a situation for which forward raytracing is fundamentally inappropriate. The emission overhead increases with the solid angle (worst case: the entire sphere) and the size of the scene cube, and can require emitting millions of photons, of which only a small fraction may actually contribute to the photon map. This situation may be further aggravated by confining the region of interest to an interior space in daylight simulation, which is only accessible to photons through windows and daylight systems.

This is a prime example of an application requiring a selective photon distribution mechanism such as the "importons" proposed by Peter and Pietrek [PP98]. A path mutation strategy similar to the Metropolis algorithm [VG97] is also conceivable. However, these extensions were not investigated in the scope of this work due to time constraints. Instead, a simpler alternative was devised, which directs the photons through user-specified "ports" into the interior.

4.1.3 Photon Ports

Conceptually, photon ports are apertures through which photons emitted from distant sources can enter the interior space containing the viewpoint. These can be windows, skylights, but also invisible polygons within the scene (figure 4.3). The user must define the port geometries for all apertures through which photons can

enter in order to obtain a physically valid simulation. The interior and exterior sides of a port are defined by its surface normal.

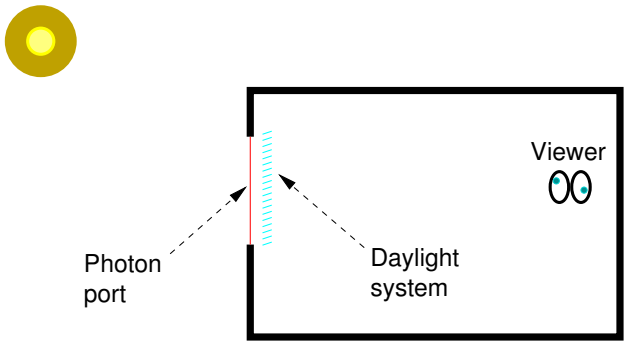


Figure 4.3: Typical daylighting geometry with window as photon port for distant sources.

Ports emit photons directly, bypassing the scene cube surfaces. Consequently the photon loss factor will be reduced dramatically, resulting in a substantial increase in performance over brute force emission from the scene cube. Photons originating from ports are scattered by their respective port as if they had passed through it. As with the brute force approach, the incident flux from each distant source is integrated on the partitioned port surfaces. Additionally, the sources are checked for occlusion. In the case of an occluded source one might consider performing an interreflection calculation (via backward raytracing) in order to account for exterior geometry, e.g. adjacent buildings. However, the resulting impact on performance would most likely forfeit the gain from using ports in the first place. This option was not investigated, and the current implementation does not account for external interreflection.

Though it caters to the majority of daylighting situations, this simple photon port concept lacks flexibility. Consider the atrium geometry in figure 4.4. The photons must pass through the skylights and subsequently find their way into one of the offices facing the atrium through its window. Essentially this scene consists of two interiors: the atrium, and the office nested within the atrium. This calls for an extension to the photon port concept which requires categorising ports according to the nesting level of the interiors they give access to. There can be several ports per level if an interior is accessible through multiple apertures, so that photons need only pass through one port within each level. The first ports through which photons must pass are termed primary ports and correspond to ports in the simple concept outlined above. Primary ports emit photons into the primary interior. The photons are then directed from the primary interior to the secondary interior(s) via secondary ports, from whence they may be further directed by tertiary ports, etc. The last port level finally directs the photons into the space containing the viewpoint. In figure 4.4, the primary ports are the skylights, which direct photons into the atrium. From there they are directed into the office via the secondary port, which is the office

window.

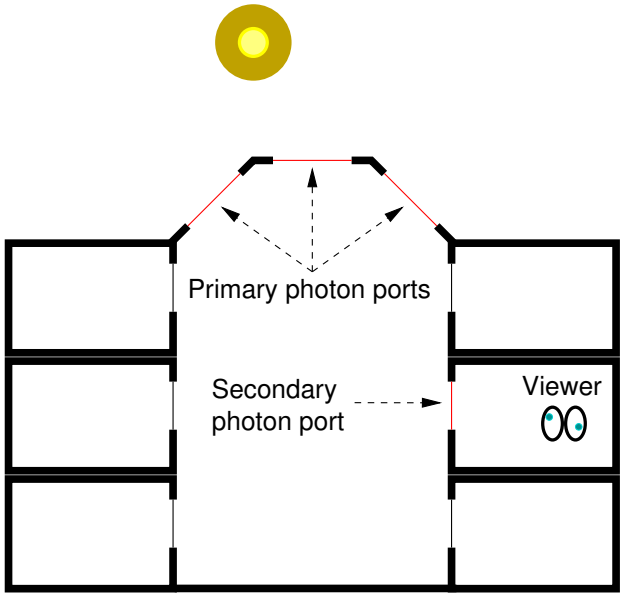


Figure 4.4: Atrium geometry with skylights as primary photon ports, and office window as secondary photon port.

In order to direct photons from one interior to next they must be scattered using importance information. Importons can supply this information, although the method requires adaptation to the port concept, since emitting importons from the viewpoint will inevitably clutter most importons within the space containing the viewpoint, while very few will pass through its ports and contribute importance for the outer ports. Instead, an importon is only stored if it has passed through all port levels, such that it contributes importance for an entire path from the primary port to the innermost space.

This extended port concept was not implemented in the course of this work due to time constraints, and is presented here as a proposal. Whatever the performance gains may be, there are caveats with importance based photon distribution which one must bear in mind:

- Photons must be scattered using numeric Monte Carlo inversion for all materials (even Lambertian), since the BRDFs must be combined with the importance for each scattering direction. This can impact performance dramatically.
- Like all importance sampling methods, the photon flux must be adjusted to compensate for the bias introduced by nonuniform sampling. This can increase noise and even produce outliers.
- The resulting photon maps are view dependent with respect to the spaces through which the photons are guided. Once the viewer leaves one of these

spaces, the photon map must be regenerated. This limits the use of extended ports in walkthroughs somewhat.

The photon port concept can accelerate the photon distribution step dramatically in difficult situations involving distant light sources. Figures regarding the speedup obtained in practice are given for the daylight simulation examples that follow.

4.1.4 Forward Pass Parametrisation

The photon distribution as described above is parametrised by the number of emitted photons N_e. However, this number bears no relation to the actual number of photons stored in the photon map after the forward pass is complete. The number of stored photons is a function of the photon absorbtion and leakage ratios, both of which in turn depend on the scene geometry, light sources, and surface characteristics. We won't know the outcome of the particle transport simulation a priori, and it is very difficult to predict. Even an experienced user will have trouble estimating how many photons must be emitted to generate a photon map of a given size, and it would be far more convenient to simply specify the approximate number of photons to be stored in the map.

The majority of photon map implementations do not consider arbitrary EDFs, and emit photons based on simple ones which can be sampled easily via analytical means. Photons are then emitted in a loop from a random source until the required number of photons is stored in the map. The EDFs can be sampled at little cost for each emitted photon, so the overhead per loop iteration is relatively small.

On the other hand, the necessary support for arbitrary EDFs becomes a burden for the RADIANCE photon map, because it can be expensive to construct the CDFs from the EDFs (both in terms of time and memory), and we cannot perform this process in a loop for every photon. It would also be prohibitive to keep all the CDFs in memory during distribution, particularly since this would be required for every surface partition to account for EDFs which vary over the light source surfaces (e.g. in applications requiring near-field photometry). Instead, we must commit ourselves to emitting a number of photons from each source partition in turn, and construct the CDF only once, without knowning how many photons are stored in the map until we are done with this partition.

We can, however, break the distribution process up into two passes, at the additional cost of repeating the CDF construction once. The first pass emits a fraction of the specified number of photons to store, and its outcome can be used to estimate how many remaining photons must be emitted to approximate the specifications. This exploits the linear scalability of the forward pass, provided that the set of source partitions and differential solid angles is the same for every source in both passes. While this is straightforward, we must tackle the problem of working with different photon map types, which will differ in density.

Given m user specified photon maps of different type requiring $n_1, \ldots, n_m$ stored photons, the two pass distribution algorithm selectively stores photons according to the relative photon map densities in order to approximate the specified photon count for each. The primary pass emits $N_{e,1} = k \min(n_1, \ldots, n_m)$ photons, with k being a constant affecting the accuracy of the estimate for the remaining number of photons to emit. Since the number of stored photons rarely exceeds the number emitted (except for closed scenes in which photons cannot terminate by leakage), this constant is usually less than 1.

The result of the primary pass is m partial photon maps containing $N_1, \ldots, N_m$ photons. If one of the maps is empty, we double k and repeat the primary pass, assuming that there is no anomaly in the scene which prevents the creation of a particular photon type (e.g. caustic photons specified for lambertian geometry). However, if there is an anomaly in the scene this will inevitably lead to an infinite loop, hence k is limited to a user specified maximum k_{max} before aborting.

After completion of the primary pass, the remaining number of photons to emit in the secondary pass is

$$N_{e,2} = N_{e,1} \max(r_1, \ldots, r_m), \tag{4.8}$$

where

$$r_j = \max\left(\frac{n_j}{N_j}, 1\right) - 1, \qquad j \in [1, m]. \tag{4.9}$$

The r_j ratios are clamped to 0 if the required number of stored photons is exceeded for a photon map in the primary pass. Overshooting the specifications is usually inconsequential, since it is more inconvenient to wind up with too few photons than with too many.

We must control the relative densities of each photon map in order to approximate the desired number of stored photons for each map in the secondary pass. The density control mechanism used by the RADIANCE photon map does not locally modify the photon densities based on local requirements throughout the scene geometry as is the case with Suykens' proposal [SW00], but merely the global relative densities of the individual photon maps.

The relative densities $d_j \in [0, 1]$ are found by normalising the r_j ratios:

$$d_j = \frac{r_j}{\max(r_1, \ldots, r_m)}, \qquad j \in [1, m]. \tag{4.10}$$

These are the probabilities of storing a photon in the corresponding photon map in the secondary pass. The photon map with the maximum r_j has a density of 1 and therefore is not affected. For the others this results in a density reduction in order to avoid exceeding the specifications. Since we are modifying the density we must compensate by increasing the photon flux used in the first pass separately for each map. Photons emitted in the primary pass have uniform flux

$$\Phi_{p,1} = \frac{\Phi_e}{N_{e,1} + N_{e,2}}. \tag{4.11}$$

Since $N_{e,2}$ is unknown during the primary pass, the flux must be scaled after the distribution is complete. Photons emitted in the secondary pass have the modified flux

$$\Phi_{p,2,j} = \frac{\Phi_{p,1}}{d_j}, \qquad j \in [1, m]\,. \tag{4.12}$$

Although this contradicts our philosophy of constant flux for all photons, it does not invalidate our approach, because we are effectively combining the partial photon maps from the primary and secondary passes. Since their respective flux to density ratios are the same, the reconstructed illumination would be similar to a single photon map with the same flux to density ratio [Chr01]. Using two different photon flux levels may however incur a slight increase in noise.

The two pass distribution is summarised in algorithms 4.1 and 4.2. Emitted photons which strike an object during distribution are scattered (and eventually terminated via russian roulette) by a generic procedure scatterPhoton(...), which calls addPhoton(...) to store incident photons. The former contains material specific code and scatters / absorbs according to the BRDF, while the latter performs the bookkeeping regarding the number of stored photons and modifying the photon flux according to the relative distribution ratios.

4.2 Daylight Simulation Examples

This section presents examples in which the photon map is applied to daylight simulation. The examples only encompass a qualitate analysis. Quantitative analysis can only be done with confidence if the simulation tool has been validated, a topic adressed in chapters 6 and 7. These examples are intended to demonstrate the photon map's ability to model the functional characteristics of daylight systems while revealing RADIANCE Classic's shortcomings in side-by-side comparisons.

The standard environment for the simulations is an office space with a window fitted with a daylight system. Glare is reduced at eye level from the lower windows with a diffuse screen (conventional blinds can also be used), while the respective daylight system is installed in the top windows.

A clear, sunny sky was generated for all examples for 12:00 pm, corresponding to a relatively high solar altitude of 49°. In order to demonstrate a system's response to different sun positions, a second sky at 5:00 pm was also generated for some systems, corresponding to a low solar altitude of 7°. Sun shading daylight systems are generally designed to block sunlight at high solar altitudes, while not affecting it at low solar altitudes, thus implementing an angular selectivity which reduces glare.

The examples rendered with the photon map used 250000 global photons with precomputed irradiance (see section 3.1.2.4), and 1000000 caustic photons, with the windows defined as photon ports, which typically accelerated the forward pass by a factor of 2. Examples rendered with RADIANCE Classic used 10 ambient bounces and zero ambient value.

```
procedure distribPhotons(n_1,...,n_m,k) {generate m photon maps containing
  n_1,...n_m photons}
  compute total light source flux Φ_e
  N_1 = ... = N_m = 0
  d_1 = ... = d_m = 1
  while N_j = 0, j ∈ [1,m] do {perform prepass}
    N_e,1 = k min(n_1,...,n_m)
    for l = 1 to N_e,1 do
      emit and trace photon p_l
      if object hit then
        scatterPhoton(p_l)
      end if
    end for
    k = 2k
    if k > k_max then {too many iterations, one or more maps still empty}
      abort
    end if
  end while
  for j ∈ [1,m] do {figure out how many more photons to emit}
    d_j = max(n_j/N_j, 1) − 1
  end for
  N_e,2 = N_e,1 max(d_1,...,d_m)
  for j ∈ [1,m] do {normalise relative densities}
    d_j = d_j N_e,1 / N_e,2
  end for
  for l = 1 to N_e,2 do {perform main pass}
    emit and trace photon p_l
    if object hit then
      scatterPhoton(p_l)
    end if
  end for
  for j ∈ [1,m] do
    for l = 1 to N_j do {scale photon flux}
      Φ(p_l) = Φ(p_l)/(N_e,1 + N_e,2)
    end for
  end for
return
```

Algorithm 4.1: Two pass photon distribution

```
procedure addPhoton(j, p) {add photon p to map j ∈ [1, m]}
  if random ξ ∈ [0, 1] < d_j then
    Φ(p) = Φ_e/d_j {compensate for nonuniform photon densities}
    store photon p in map j
    N_j = N_j + 1
  end if
return
procedure scatterPhoton(p)
  ...
  if photon of type j then
    addPhoton(j, p)
  end if
  ...
return
```

Algorithm 4.2: Two pass photon distribution (continued)

The office is rendered without daylight system in figure 4.5 with both RADIANCE Classic and photon map. Both algorithms required similar rendering times. As expected, the RADIANCE Classic and the photon map renderings are very similar. The crucial difference is that RADIANCE renders the sunlit area on the floor with direct illumination, whereas the photon map renders it as caustics arising from the transmission through glass.

4.2.1 Example 1: Y-Glass

The Y-glass system manufactured by INGLAS GmbH[3] consists of a plexiglass panel containing air filled lamella at regular intervals, and can be sandwiched between glazings for easy maintenance. The system redirects light via total internal reflection at the lamella boundaries, while light passing between the lamella is simply refracted, delineating a Y-shaped light path which lends its name (figure 4.6). The lamella are angled towards the interior by 7° and implement an angular selectivity by reducing the transmittance between lamella with increasing solar altitude (figure 4.7). The system's appeal lies in its simplicity, effectiveness, and ease of maintenance.

Figure 4.8 is a photon map rendering depicting Y-glass's redirecting properties for normal and 50° incidence. These images were obtained by inserting a polygon in the plane of incidence to visualize caustics, and using a spotlight as the incident beam. 250000 caustic photons were used here. At normal incidence, the majority of light is transmitted, with a small downward redirection as a result of reflection at the underside of the lamella. At 50° incidence, the incident beam is split into an intense beam which is reflected upwards by the lamella, and less intense beam which passes through the lamella, forming the Y-shaped light path. There are also a

[3] http://www.inglas.de/Home/Produkte/INGLAS_-_Y_/inglas_-_y_.html

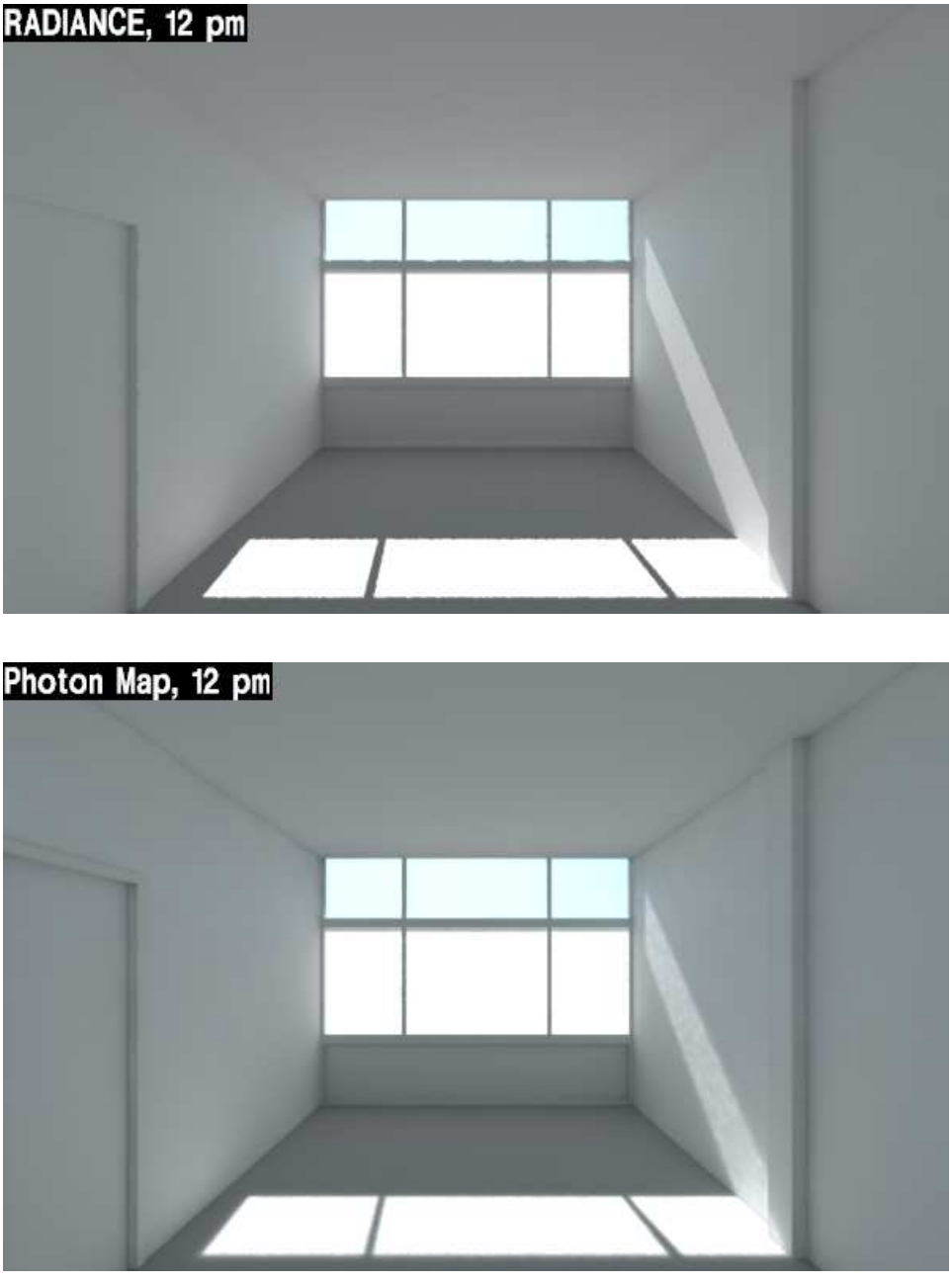


Figure 4.5: Office without daylight system rendered with RADIANCE Classic (top) and photon map (bottom). Both algorithms yield very similar results.

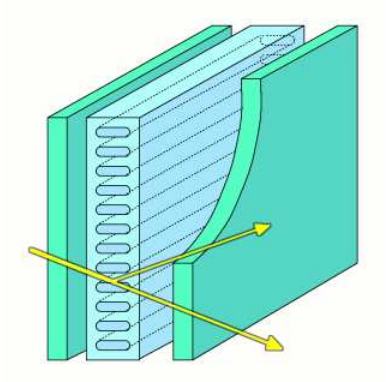


Figure 4.6: Schematic of light redirection in Y-glass panel. Courtesy of Georg Mischler, Schorsch.com.

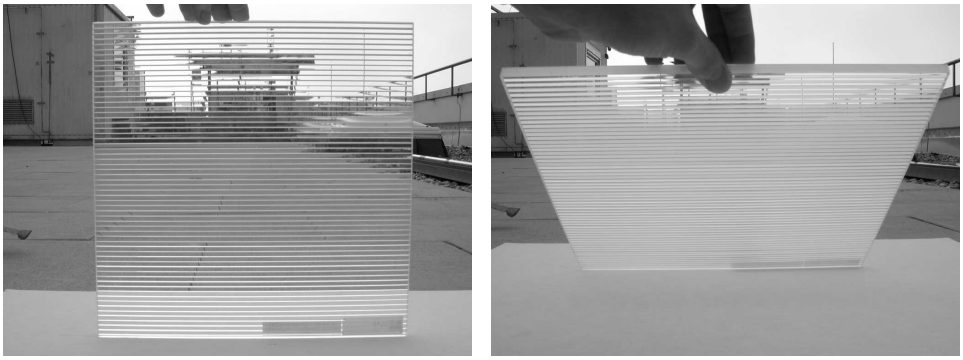


Figure 4.7: Photographs of a Y-glass sample. The air filled lamella reduce the transmittance for light incident from high angles, as demonstrated by inclining the sample.

number of weaker beams resulting from reflection while entering the dielectric and from multiple reflections inside the system, but these are exaggerated by the tone mapping due to the high dynamic range and do not actually contribute significantly.

Figure 4.9 depicts renderings of the office fitted with the Y-glass system at a high solar angle. As expected, there is an obvious caustic at the ceiling resulting from reflection off the lamella, which the photon map simulates faithfully. The ceiling reflection increases the ambient luminance in the room compared to RADIANCE Classic. The relative increase in luminance can be deduced from the falsecolour image in figure 4.10. It depicts the relative difference (in percent) between the RADIANCE Classic and photon map renderings. The ceiling caustic shows an increase of up to 400% compared to the RADIANCE Classic rendering, and the ambient luminance on the walls and floor outside the caustic is also raised by 40-60%. This constitutes a significant difference for daylight analysis and illustrates the necessity of forward raytracing in this application.

The photon map rendering in figure 4.9 also confirms that the majority of light

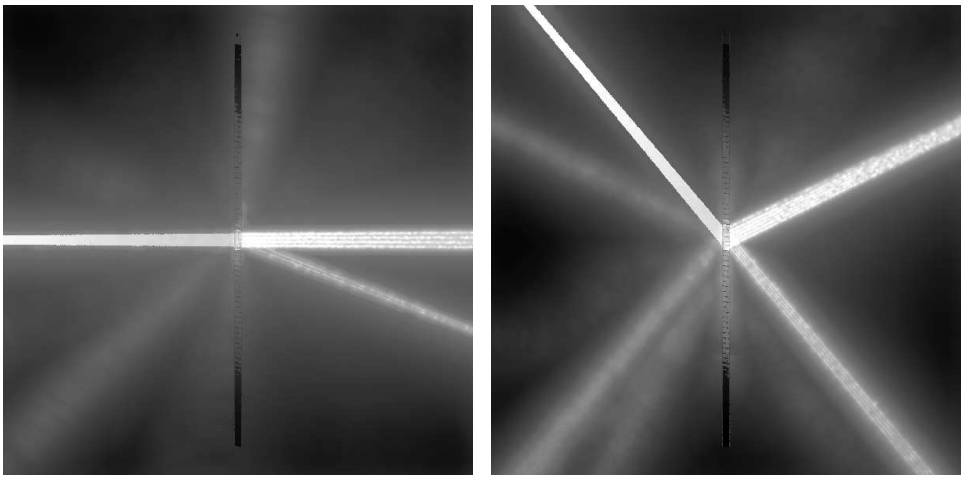


Figure 4.8: Y-glass redirection rendered with the photon map. At normal incidence (left), the beam is transmitted, while at 50° incidence (right) the beam is split into an intense beam redirected upward, and a less intense transmitted beam. The result is a Y-shaped light path as depicted in figure 4.6.

is redirected to the ceiling, thereby reducing glare at eye level. This is evident by comparing the luminance of the sunlit region on the floor resulting from light transmitted between the lamella with the luminance without daylight system in figure 4.5. The ratio of redirected to transmitted sunlight increases with the solar angle.

Figure 4.11 depicts the Y-glass installation at a low solar angle. The ceiling caustic is no longer present in the photon map rendering since the solar angle deviates only marginally from the lamella inclination, consequently most sunlight passes unaffected between the lamella. As a result, both renderings are similar. The rendering times for photon map and RADIANCE Classic were about the same in this example.

4.2.2 Example 2: Compound Parabolic Concentrator

Compound parabolic concentrators, or CPCs, consist of an array of specular (traditionally metallic) parabolic profiles designed to concentrate light entering the system [WW78]. The concentrating effect is accomplished via reflection at adjacent profiles in dependence of the angle of incidence, thereby implementing angular selectivity. Light entering the system within a solid angle around the normal (termed the *acceptance range*) is concentrated at the emitting aperture, while light incident from directions outside this range is reflected back out (rejected) after multiple bounces within the system (figure 4.12). The acceptance range is typically well defined, characterised by an abrupt transition from acceptance to rejection. CPCs are very versatile systems and can be employed in a number of ways, primarily to increase radiative gain in applications ranging from lasers to solar energy collectors.

A novel application of CPCs is in façades for sun shading [Kuc02]. This applica-

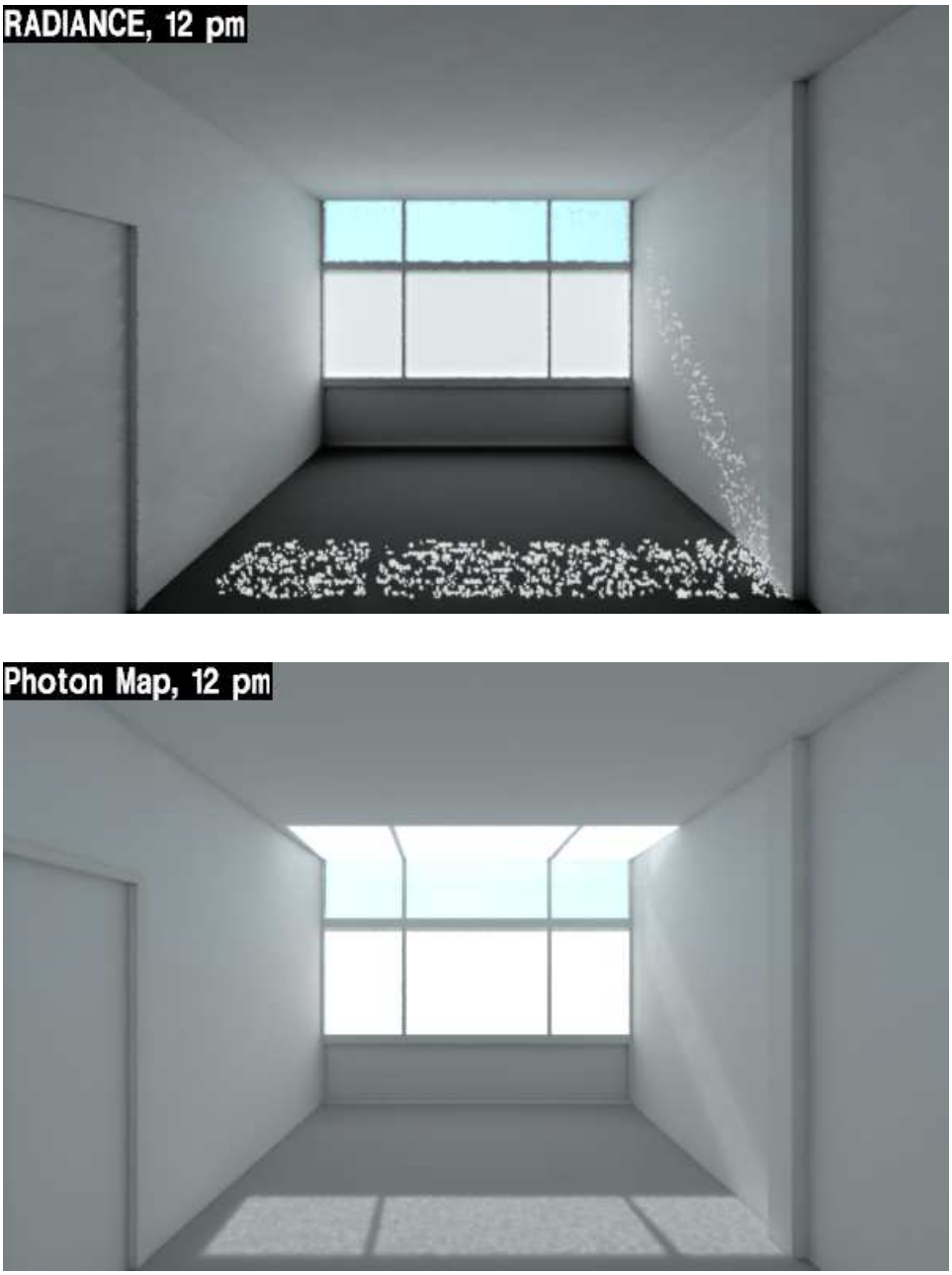


Figure 4.9: Y-glass installation rendered with RADIANCE Classic (top) and photon map (bottom) at high solar altitude. The photon map correctly models the behaviour of Y-glass, producing a caustic at the ceiling from reflection off the lamella and resulting in increased ambient luminance from ceiling reflection. As a consequence of redirection, transmitted light incident on the floor is reduced compared to figure 4.5.

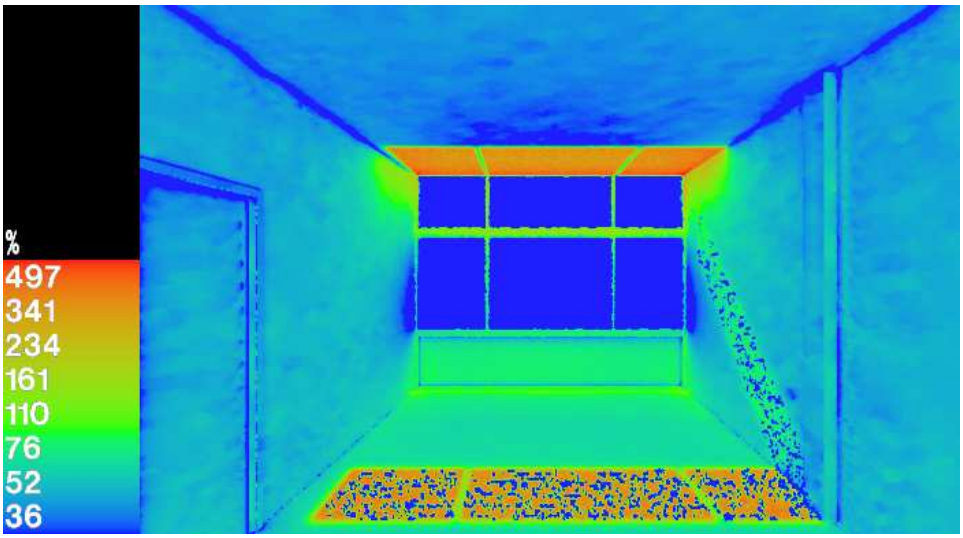


Figure 4.10: Relative difference between RADIANCE Classic and photon map renderings in figure 4.9. The increase in ambient illumination from the ceiling caustic raises the ambient luminance in the room by 40-60%, while the ceiling caustic itself accounts for a luminance increase of up to 400%.

tion is still under development and no manufacturer currently markets such a system. Sun shading CPCs differ in material and behaviour from conventional CPCs. These systems consist of a solid dielectric profile, and the concentrating effect is accomplished via total internal reflection (see figure 4.13). The emitting aperture is coated with an opaque reflecting layer, which may be specular or diffuse. In this way glare from direct sunlight which falls within the acceptance range is concentrated and reflected back out of the system by the opaque aperture. On the other hand, diffuse skylight which enters outside the acceptance range is not subject to total internal reflection and is instead transmitted at the profile walls, in stark contrast to conventional metallic CPCs (figure 4.14). Sun shading CPCs therefore exhibit complementary behaviour to conventional CPCs. A desirable property of sun shading CPCs is that rays transmitted by the system are fanned out due to the curved geometry of the concentrator walls, giving the illumination a diffuse quality.

The versatility of CPCs gives rise to a number of possible sun shading configurations [Kuc02]. The system can be installed as fixed horizontal or vertical panels with an acceptance range optimised to block direct sunlight throughout the summer months, while allowing it to pass through the system in winter, when the sun is low and less instense. The CPCs can then be mounted with a fixed angle, without the need to manually correct the alignment during seasonal changes. Alternatively, the CPCs can also be installed as manually adjustable lamella to fully utilise their redirecting characteristics to cover a broad range of shading or daylight utilisation needs.

Sun shading CPCs are manufactured from clear plastic or resin in an elaborate

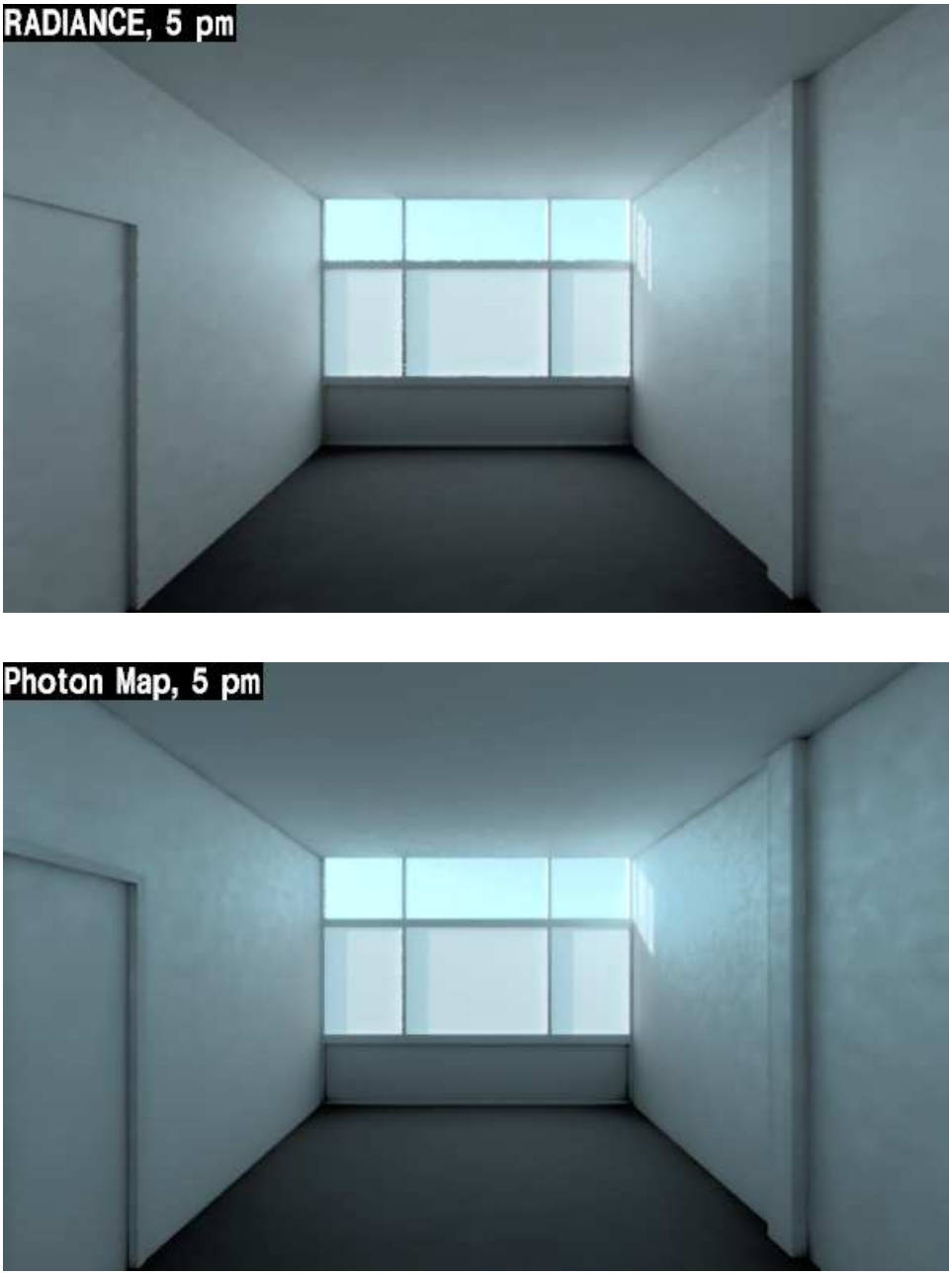


Figure 4.11: Y-glass installation rendered with RADIANCE Classic (top) and photon map (bottom) at low solar altitude. Sunlight mostly passes between the lamella, and no caustic is formed. Consequently, both algorithms yield similar results.

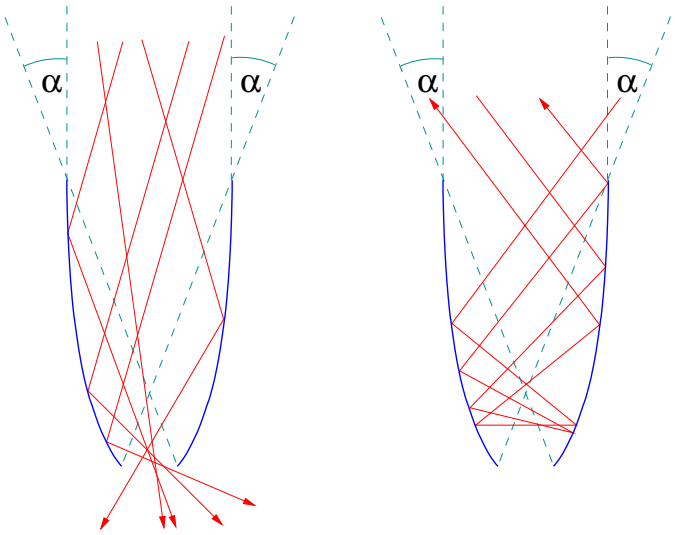


Figure 4.12: Schematic of light redirection in conventional CPC. Rays incident within the acceptance angle α (left) are concentrated via reflection at the metallic parabola walls and emitted below. Rays incident outside this range (right) are reflected back out.

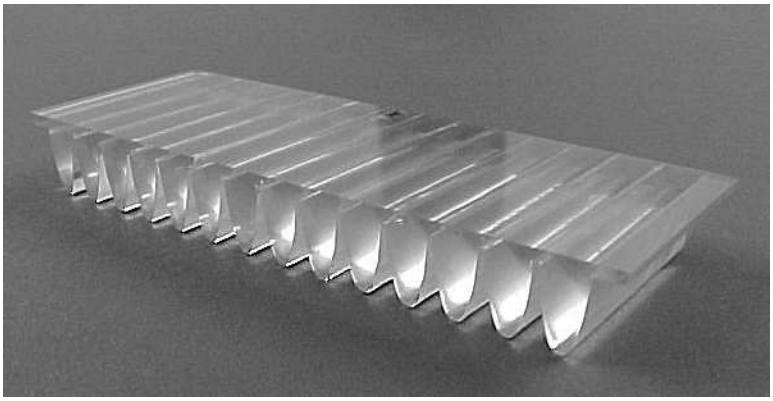


Figure 4.13: Photograph of dielectric sun shading CPC array. Reproduced from [Kuc02].

moulding or extrusion process. The difficulties in manufacture arise from the low tolerances; minute roughness and deformations in the CPC's walls can reduce the effect of total internal reflection and therefore degrade the system's performance. Glass has also been investigated as an alternative material since it is more resistant to long term UV exposure than plastics if the CPCs are mounted externally [Kuc02].

The photographs in figure 4.15 depict the relationship between viewing angle and the sun shading CPC's acceptance range, which serves to illustrate the system's behaviour. Viewing the sample within the acceptance range causes the system to appear opaque, since incident rays are blocked by the opaque layers on the

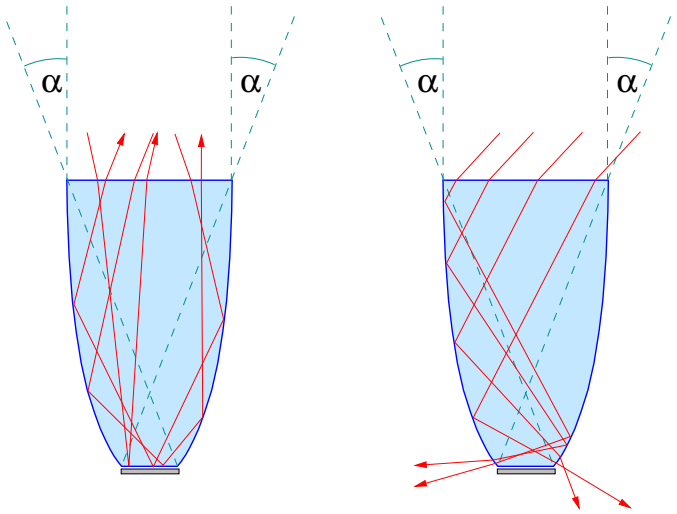


Figure 4.14: Schematic of light redirection in dielectric sun shading CPC. Rays incident within the acceptance angle α undergo total internal reflection at the walls and are concentrated at the opaque peak below, where they are reflected back out. Rays incident outside this range are transmitted and refracted at the parabola walls. This behaviour is complementary to that of conventional CPCs as seen in figure 4.12.

emitting apertures. Viewing the sample outside the acceptance range causes the system to appear translucent, since the system is now in a transmitting configuration.

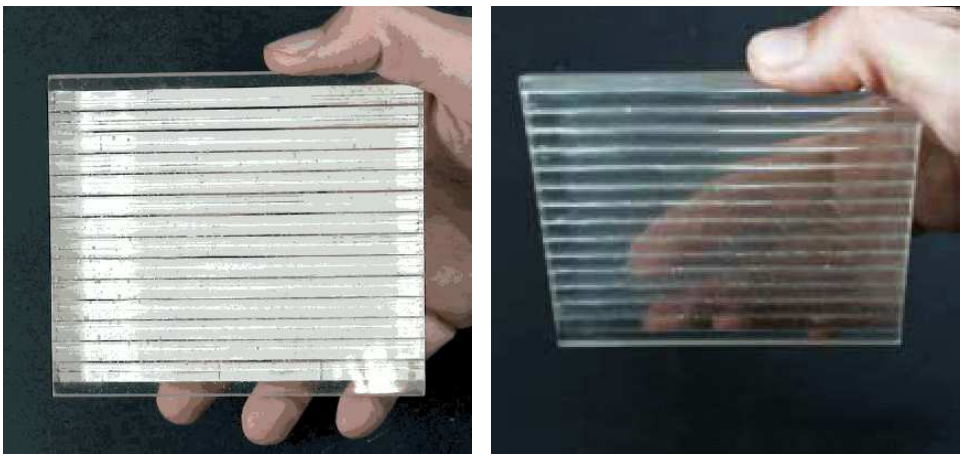


Figure 4.15: Sun shading CPC acceptance range vs. viewing angle. Left: viewing angle within acceptance range, system is opaque. Right: viewing angle outside acceptance range, system is translucent. Photographs reproduced from [Kuc02].

Figure 4.16 succinctly illustrates the redirecting behaviour of a sun shading

CPC. These photographs were taken in a gaseous environment with a laser. At normal incidence, the beam is incident within the acceptance range and reflected back. At incidence outside the acceptance range the beam is transmitted by the system and redirected.

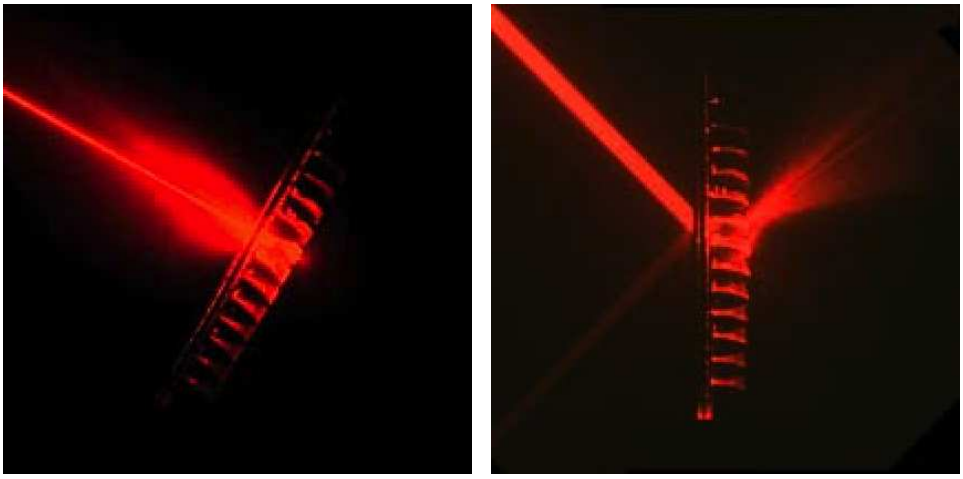


Figure 4.16: Photographs of laser redirection in sun shading CPC. Left: beam incident within acceptance range, the beam is reflected back out. Right: beam incident outside acceptance range, beam is transmitted and redirected. Reproduced from [Kuc02].

The redirecting properties of a sun shading CPC are also rendered with the photon map in figure 4.17. These renderings are inspired by figure 4.16 and serve as comparison. As in the photographs, the beam is reflected back if it is incident within the acceptance range, while at incidence outside the acceptance range, the system transmits and fans out the emitted light. The breakup into stripes visible in the fan-out is an artifact resulting from the faceted model used for the parabola walls. The renderings are similar to those in figure 4.16, although the geometries of the rendered and photographed CPCs are not identical. The absence of the downward redirection in the photographs is explained by the smaller coverage of the laser compared to that of the beam in the rendering. This is corroborated by figure 4.14, since a narrow incident beam partially covering the CPC width would only account for a portion of the rays shown. With the incident beam covering only the right portion of the CPC, the system will only emit on the right side, corresponding to the configuration seen in the photographs.

Figures 4.18 and 4.19 depict renderings of the office fitted with an external sun shading CPC lamella system at high solar angle in accepting and redirecting settings, respectively. In the accepting setting the CPCs are inclined by the solar angle, such that sunlight is incident within the acceptance range and reflected out of the system. Consequently the photon map and RADIANCE Classic renderings are similar, except for a slight increase (up to 10%) in ambient luminance in the photon map rendering. This is contributed by a small fraction of caustic photons which were

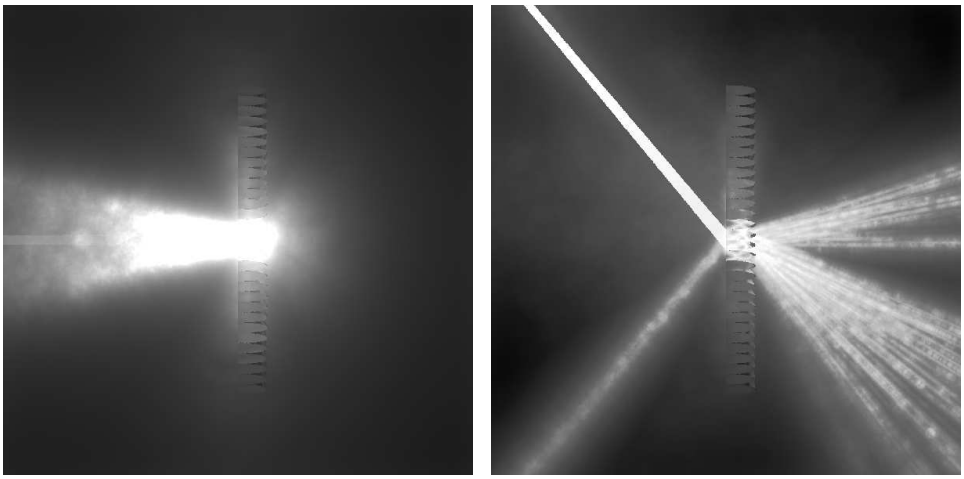


Figure 4.17: Sun shading CPC redirection rendered with the photon map. At normal incidence (left), the beam is concentrated in the system and reflected back out at the opaque apertures, whereas at 50° incidence (right) the beam is transmitted and fanned out. Note also the reflection while entering the dielectric. Compare these images to figures 4.14 and 4.16.

not subject to total internal reflection and therefore transmitted into the room.

In the redirecting setting the CPCs are vertical, such that sunlight is incident outside the acceptance range and transmitted towards the floor and ceiling. This effect is reproduced by the photon map, but not by RADIANCE Classic, as expected. Consequently, there is a dramatic increase in ambient luminance contributed by caustics. From figure 4.20 we can see that this increase amounts to 100% or more on the walls and ceiling. The caustics on the floor are even more intense, resulting in an increase of almost 300%. Again, this is a point in case for the application of forward raytracing in daylight simulation.

Unlike the previous examples, in which the systems were mounted internally, the photon ports for the top windows were replaced by an invisible alcove encasing the CPCs outside the window. The ambient noise in the RADIANCE Classic renderings is typical for situations involving high gradients in the ambient irradiance (conceivably due to the occasional ambient ray which found a caustic path), and can only be effectively suppressed with denser stratified samples, leading to exorbitant rendering times. The rendering times for RADIANCE Classic in this example already exceeded those of the photon map by a factor of 2.5-3.

4.2.3 Example 3: Lightpipe

The lightpipe [EGW98] is representative of those systems which cannot be adequately simulated with RADIANCE Classic. Daylight is conducted into building interiors by the pipe, and can be decoupled where necessary to reduce the need for artificial lighting. This is achieved by lining the pipe's inner surface either with a

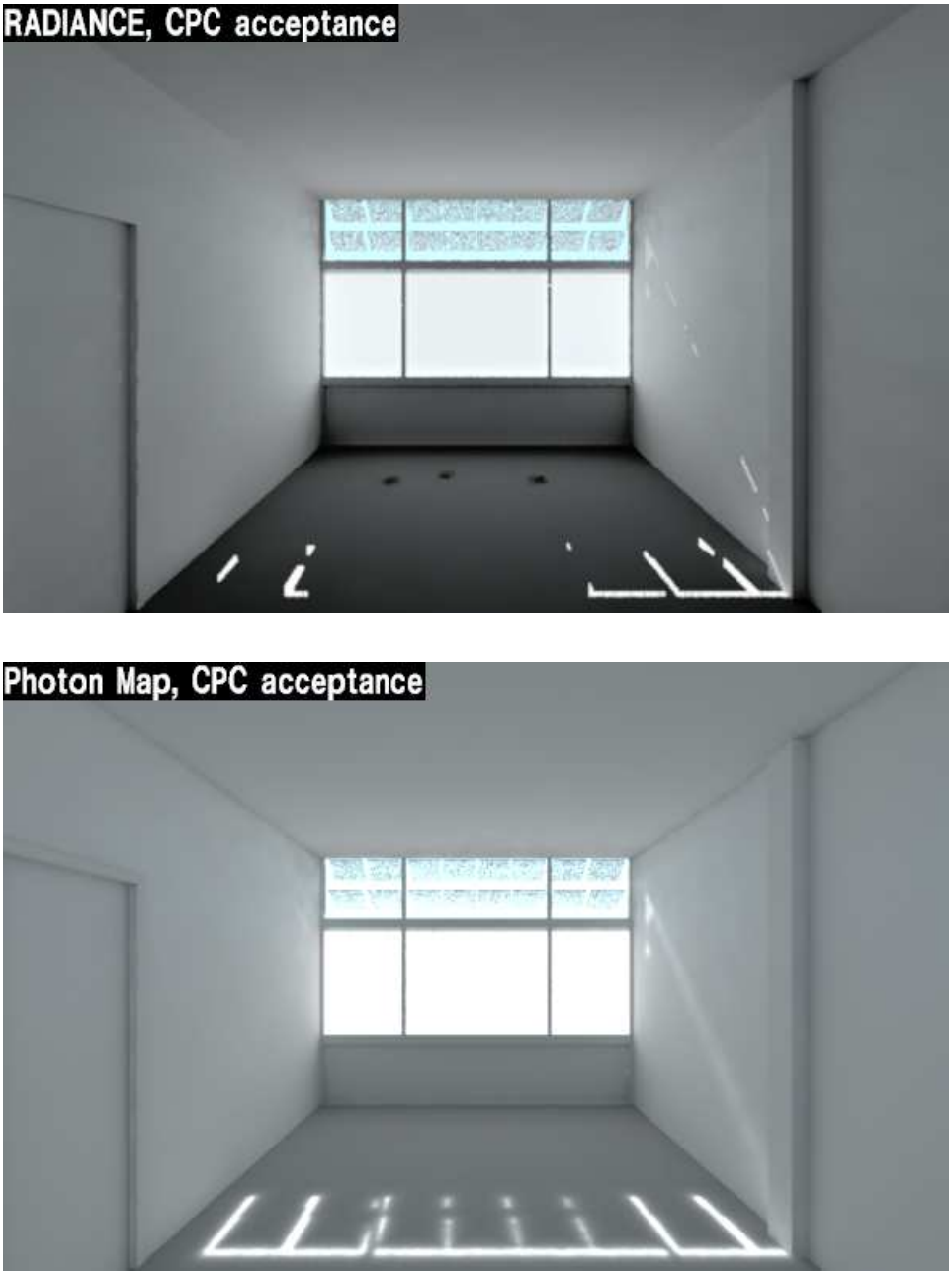


Figure 4.18: Sun shading CPC installation rendered with RADIANCE Classic (top) and photon map (bottom) at high solar altitude with CPCs in accepting setting. The photon map correctly models the acceptance range, reflecting sunlight out of the CPC panels towards the exterior, and thus not giving rise to any noticeable caustics. There is however a slight increase in ambient luminance due to a small fraction of caustic photons which were not subject to total internal reflection, and admitted into the room. This minor difference appears somewhat exaggerated by the tone mapping applied to the renderings.

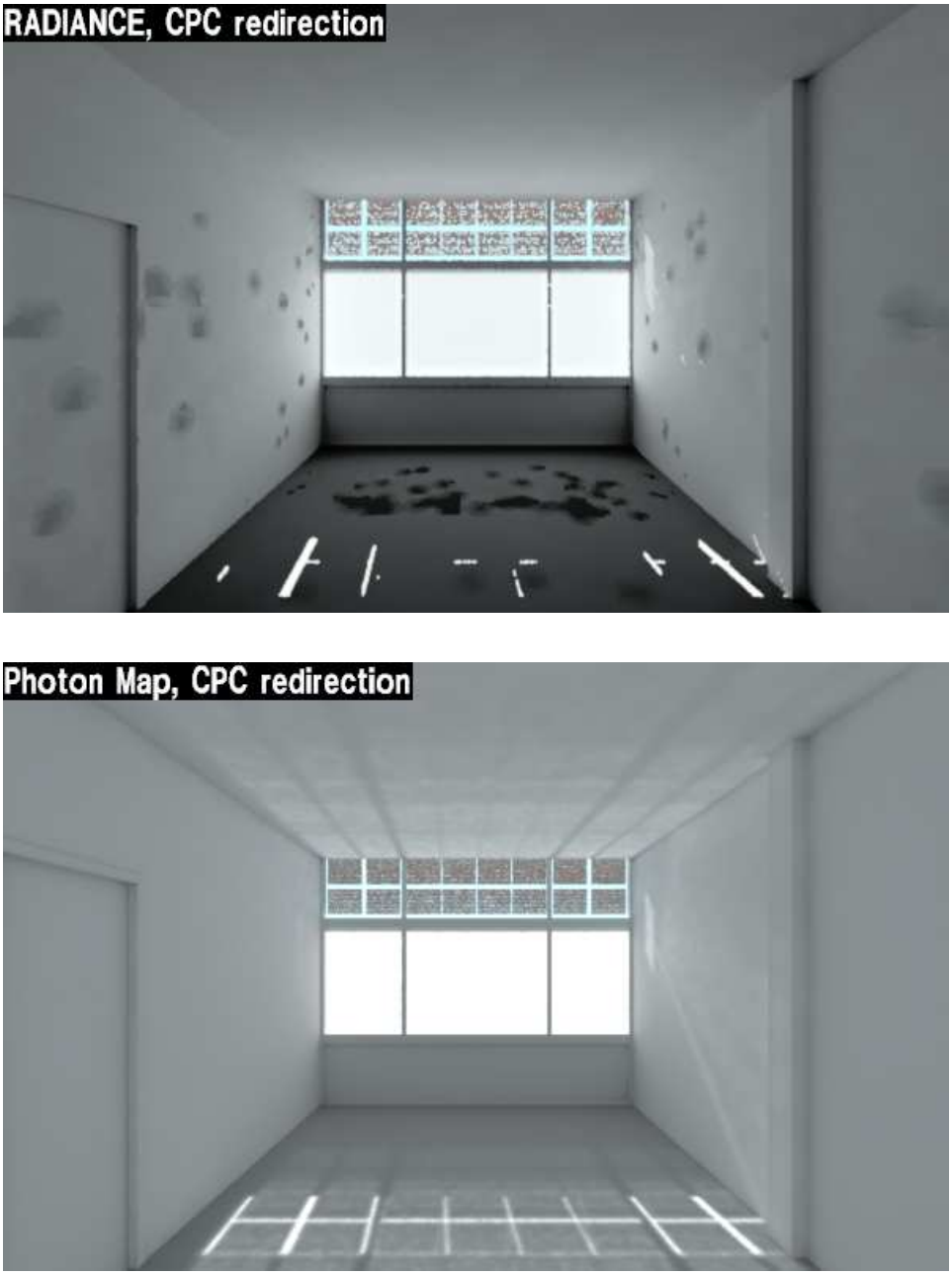


Figure 4.19: Sun shading CPC installation rendered with RADIANCE Classic (top) and photon map (bottom) at high solar altitude with CPCs in redirecting setting. The photon map correctly models the redirection of sunlight as caustics towards the floor and ceiling, resulting in a dramatic increase in luminance compared to RADIANCE Classic. Note the high level of noise in the RADIANCE Classic rendering, in spite of higher rendering times than required with the photon map.

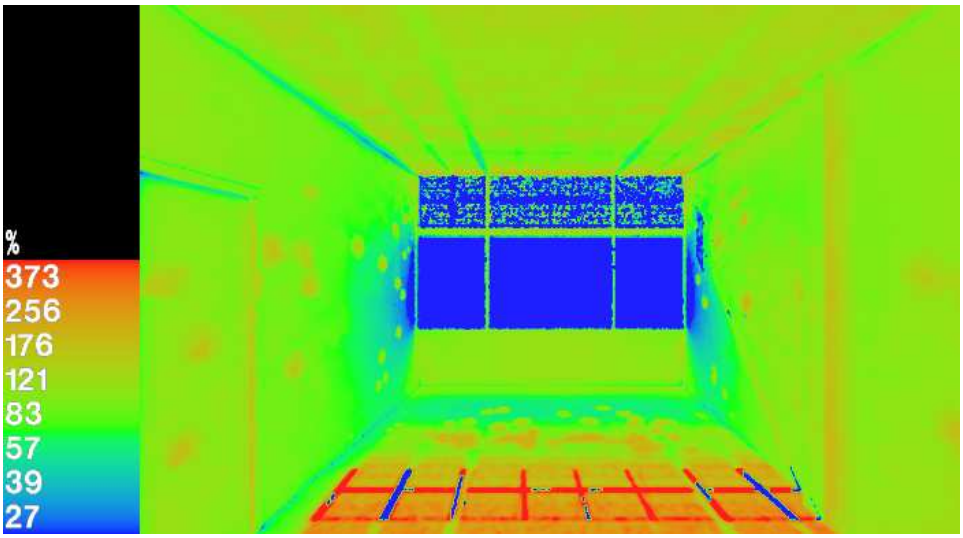


Figure 4.20: Relative difference between RADIANCE Classic and photon map renderings in figure 4.19. There is a substantial increase in luminance on the ceiling and walls due to caustics in excess of 100%. The caustics on the floor constitute an increase of nearly 300%.

mirror finish or with a prismatic foil, the latter conducting the majority of the light via total internal reflection, while still transmitting a small fraction which lends the system an appealing self-luminous appearance. Sunlight is usually directed into the pipe's aperture with a heliostat and concentrated with a Fresnel lens in order to increase the system's gain.

Figure 4.21 shows an example of a lightpipe installation in the 1st floor corridor of Fraunhofer ISE. An externally mounted heliostat which tracks the sun's path directs light onto the centre of a fixed mirror. This mirror directs sunlight downward onto a second fixed mirror, which in turn directs it horizontally onto a Fresnel lens mounted in the corridor window before entering the lightpipe. The lightpipe runs the entire length of the building (26 metres), partly illuminating the corridor through its prismatic foil and transporting the majority of light to a kitchenette and a small computer room at the end of the corridor, both of which lack windows. Lightpipes can also be employed in subterranean environments such as basements, underground parking lots, and subways. Figure 4.22 is an example of lightpipes installed at Potsdamer Platz subway station, Berlin. In this case the pipes are mounted vertically and conduct daylight to the subway platforms below. Each pipe is fitted with its own heliostat.

Figure 4.23 depicts renderings of a lightpipe installed in a corridor. The lightpipe has a mirror finish, and extends through several doorways towards the outside above the window in the far wall. A heliostat was modeled and placed at the pipe aperture outside. It is aligned towards the sun, reflecting sunlight directly into the pipe, which is then conveyed through the pipe towards the viewer. A mirror inclined

Figure 4.21: Lightpipe installation in 1st floor corridor at Fraunhofer ISE. Sunlight is directed into the pipe via a heliostat and two fixed mirrors (top). Daylight is then conducted by the pipe's prismatic foil through the corridor (left) and directed into rooms lacking windows (right). Images reproduced from [EGW98].

at 45° (seen from the back) decouples light from the lightpipe and directs it to the ceiling, illuminating the room indirectly. Figure 4.25 is an overview of the scene, depicting a cross-sectional rendering using the photon map. The corridor has been shortened for clarity.

The difference between the photon map and RADIANCE Classic renderings is striking, yet not surprising. The photon map clearly simulates the principle behind the system, directing caustic photons into the pipe via the heliostat, conveying them through the pipe and up to the ceiling, finally illuminating the room via diffuse reflection. Photon ports were defined for the window, pipe aperture, and the heliostat. Because the pipe represents such a severe bottleneck for the photon distribution, the use of ports in this scene has a far greater impact than in the previous examples, yielding a dramatic speedup by a factor of 175. 500000 global photons and 1000000 caustic photons were used for this example.

Figure 4.22: Lightpipes in vertical configuration at Potsdamer Platz subway station, Berlin. The pipes conduct daylight to the subway platforms below (left). At night they conduct artificial light from lamps mounted within (right).

RADIANCE Classic has serious problems here, and only manages to produce a vestige of the expected caustic, since few ambient rays find their way into the pipe via the mirror, and towards the solar source via the heliostat at the far end of the pipe. To quantify the extreme divergence between the two algorithms in this example, we note that the photon map increases the ambient illuminance on the walls at the near end of the corridor by a hefty 3000-6000%, while deviations in the ceiling caustic even exceed a staggering 100000%!

The futility of capturing caustics via backward raytracing comes fully to bear in this example, particularly when considering that RADIANCE Classic required the sixfold rendering time of the photon map to render this scene. This is the direct result of extreme parameter settings in order to suppress noise in the ambient luminance from what little caustics is sampled.

4.3 Conclusions

The photon map is capable of producing a complete, accurate global illumination solution suitable for daylighting applications. The daylight simulation examples have shown that it is capable of modelling the functional charateristics of some representative daylighting systems. Furthermore, it outperforms RADIANCE Classic in most situations.

However, the distant light sources typically used in daylighting applications pose fundamental problems for a forward raytracer. In these cases, the gain in performance over RADIANCE Classic is achieved primarily by applying the proposed photon port mechanism, which significantly improves the efficiency of the forward pass at the expense of some user intervention.

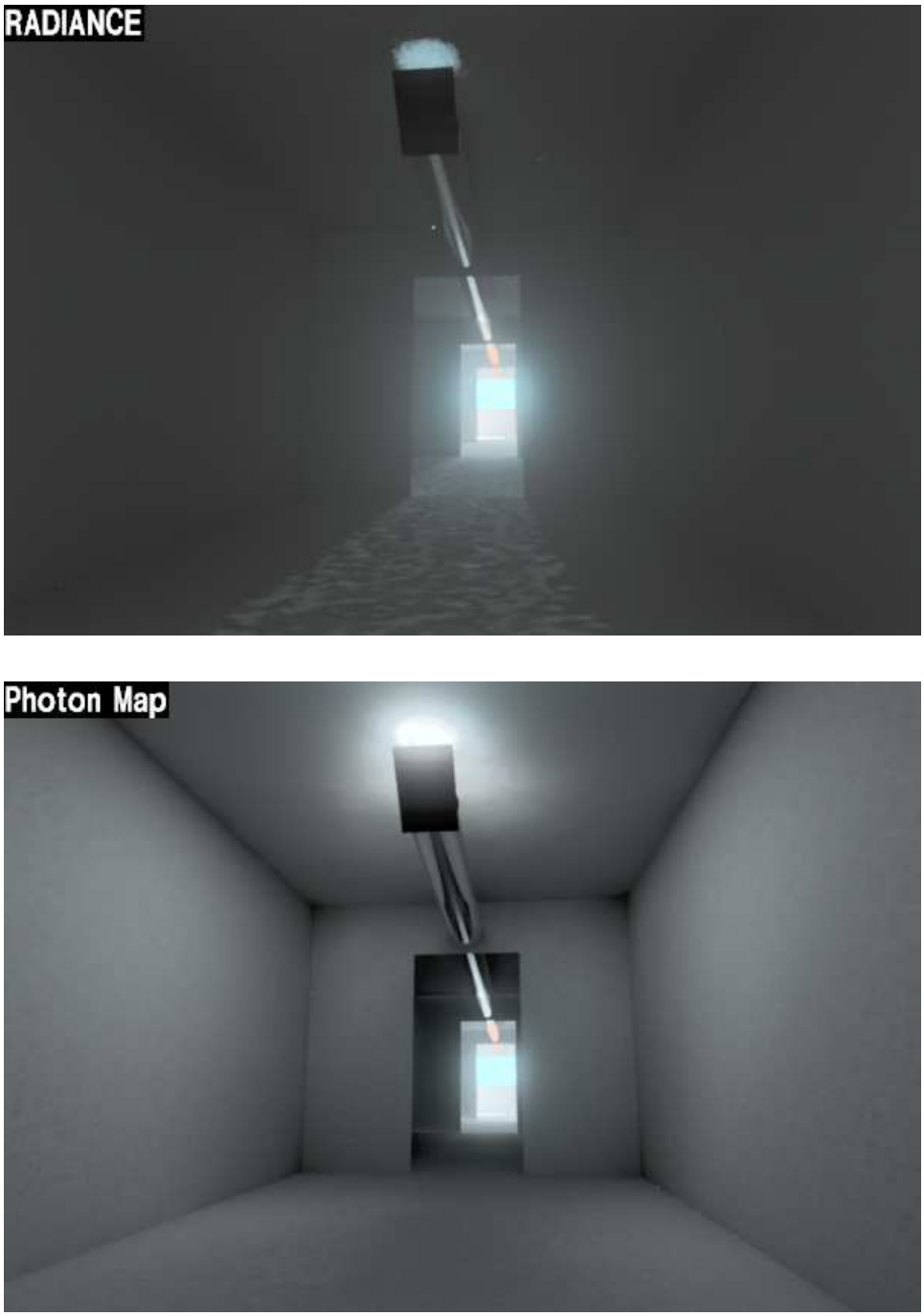


Figure 4.23: Corridor with lightpipe rendered with RADIANCE Classic (top) and photon map (bottom). The photon map correctly simulates the principle behind the system, conducting caustic photons from the far end of the corridor to the near end, where they illuminate the room via ceiling reflection. RADIANCE Classic only manages a vague hint of the caustic, leaving the room in darkness.

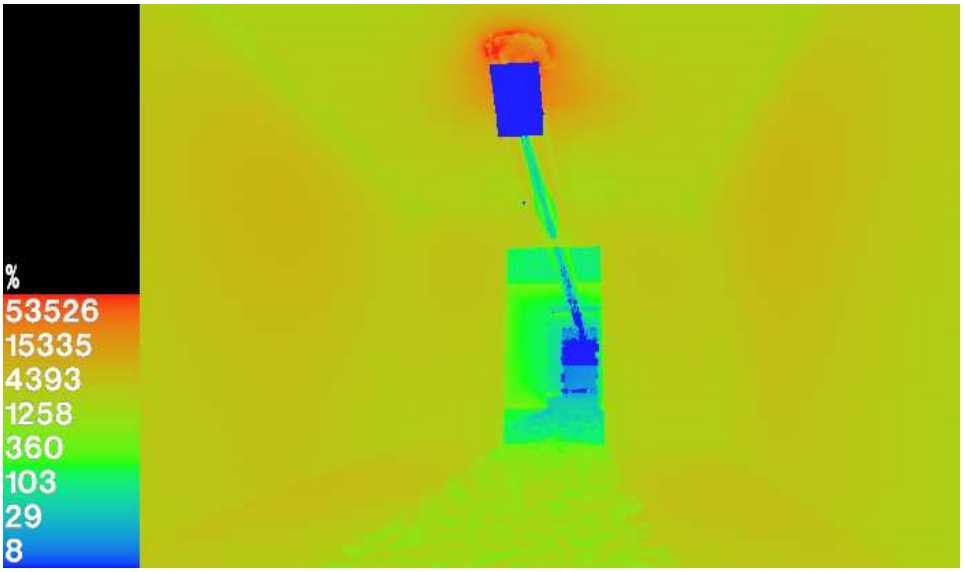


Figure 4.24: Relative difference between RADIANCE Classic and photon map renderings in figure 4.23. The deviations in luminance due to the caustic are staggering, amounting to 3000-6000% on the walls and even 100000% in the caustic itself.

Figure 4.25: Cross-section of lightpipe scene rendered with photon map. The viewpoint in figure 4.23 is positioned at the right end of the corridor, facing the doorway.

Chapter 5

Bias Compensation

Few publications have subjected the photon map to an error analysis. Indeed, most of the relevant literature is scant on the subject and rarely questions the validity of the method, merely pointing out that errors in the reconstructed indirect illumination can be reduced by simply increasing the number of photons. While this is true in general (as for all Monte Carlo methods), we would like to characterise these errors and make them more concrete. Furthermore, since daylighting simulation is an application which demands accuracy, we must concern ourselves with a means of estimating and reducing this error. Little research has been done here in conjunction with the photon map, and is the subject of this chapter. After introducing the bias/noise tradeoff and its relation to density estimation, we develop a proposal for a bias compensating operator specifically for radiometric applications which dynamically adapts the bandwidth for nearest neighbour density estimates.

5.1 Density Estimation Error

Given a photon map, we are interested in finding the error in density estimates based on nearest neighbour search in dependence of the bandwidth N_p. As described in section 2.6.1 for Monte Carlo methods in general, this error can be broken down into two components: variance and bias.

Variance is the fundamental random error common to all Monte Carlo methods and is visible as noise in the renderings. This error drops as N_p increases, but typically at the modest rate proportional to $1/\sqrt{N_p}$ derived in equation 2.32. The variance σ^2 of a set of density estimates $\hat{f}(\vec{x},N_p)$ is defined as:

$$\sigma^2\left[\hat{f}\left(\vec{x},N_p\right)\right] = E\left[\hat{f}^2\left(\vec{x},N_p\right)\right] - E^2\left[\hat{f}\left(\vec{x},N_p\right)\right]. \tag{5.1}$$

Unlike noise, bias is the systematic error in the reconstructed illumination. It is defined as the deviation of the expected value of the estimator $\hat{f}$ from the actual illumination f:

$$\beta\left[\hat{f}\left(\vec{x},N_p\right)\right] = E\left[\hat{f}\left(\vec{x},N_p\right)\right] - f\left(\vec{x}\right). \tag{5.2}$$

This implies that bias cannot be quantified without knowing the actual illumination, which is generally not available.

Bias resulting from nearest neighbour density estimation can be classified into three categories based on the source of the error. In this document, these are referred to as *proximity, boundary,* and *topological bias*.

Proximity bias is a fundamental problem with density estimation arising from the fact that photons from nearby regions are drawn upon for the density estimate. The expected value of a density estimate $\hat{f}(\vec{x}, N_p)$ is

$$E\left[\hat{f}(\vec{x}, N_p)\right] = \int_{\{\vec{y}:\|\vec{x}-\vec{y}\|\leq r_p\}} K(\|\vec{x}-\vec{y}\|) f(\vec{y})\, d\vec{y}, \tag{5.3}$$

where $\vec{y}$ are nearby points within the kernel bandwidth defined by the search radius r_p. This means that the estimate converges to the actual illumination function f *convolved with the kernel* K used to weight the photons, rather than to f proper [WJ95]. This is visible as smearing of details in the renderings, particularly caustics. Effectively, proximity bias displaces illumination from nearby regions to areas where it should not contribute, thus locally violating energy conservation [Suy02, p.113]. Proximity bias is typically introduced in those regions where the bandwidth crosses some illumination feature boundary.

Boundary bias is visible as a darkening of density estimates at polygon boundaries and occurs when the bandwidth extends beyond the polygon boundary, where no photons are located [Suy02, p.112]. This results in a spurious decrease in density.

Topological bias is the error introduced by the density estimate under the assumption that the area occupied by the photons is planar. Any curvature will generally lead to underestimation of this area and thus overestimation of the illumination [Suy02, p.111].

Regardless of its source, bias increases with the bandwidth N_p. We will be primarily concerned with proximity bias in this chapter, which is the most common form, and unless specified, bias is due to proximity.

Myszkowski [Mys97] and Walter [Wal98] are among the few who have investigated bias in density estimation. Myszkowski reported substantial bias with caustics (in excess of 10%) using nearest neighbour methods, clearly implying that a naive approach to density estimation cannot satisfy the requirements for daylighting simulation.

5.2 A Bias Case Study

A simple case study to illustrate the behaviour of noise and bias is shown as photon distribution plot in figure 5.1. The 50000 photons are distributed directly in the

plane rather than being distributed from a light source in order to rule out any bias inherent in the particle tracing step. The circular highlight in the centre has the tenfold irradiance (10 W/m^2) of the surrounding area (1 W/m^2) and represents a situation analogous to a caustic.

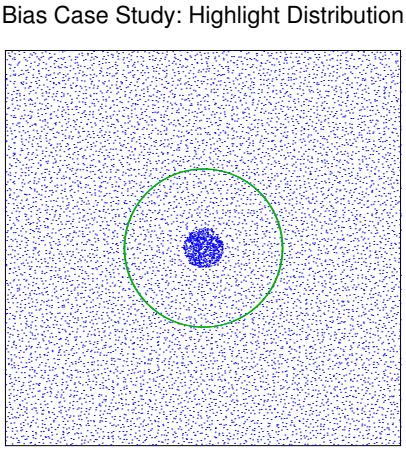


Figure 5.1: Photon distribution for highlight case study. The dense highlight region in the centre of the plane has the tenfold irradiance of the surroundings. Density estimates for the case study are taken at points on the green circumference.

Figure 5.2 is a plot of the RMS (root mean square) noise and mean bias for 1000 density estimates taken outside the highlight on the circumference shown in figure 5.1. The density estimates are graphed for the uniform, cone, Epanechnikov, and Silverman kernels.

For all kernels, the noise drops slowly as expected for Monte Carlo, while the bias hugs the zero axis until ca. 3300 photons per estimate. Bias sets in once the bandwidth crosses the highlight boundary and photons are gathered from its interior. For the uniformly weighted density estimates, the bias rises dramatically, while those using nonuniform weights fare better and suppress bias considerably, as well as slightly reducing noise. The Silverman kernel appears to be particularly effective. However, none of the kernels actually eliminate the bias, but merely delay its onset as the bandwidth increases.

A visual impression of extreme bias can be seen in figure 5.3, which is a false-colour rendering using the uniform kernel with a bandwidth of 5000 photons. Note that the highlight barely exceeds an irradiance of 4 W/m^2, less than half its actual value, and boundary bias leads to darkening well into the interior.

Bias and noise are antithetic; while increasing the bandwidth reduces noise, it can increase bias, and vice versa. This leads to the conflicting requirement of a large bandwidth to reduce noise, yet a small bandwidth to preserve details and reduce bias. Instead of using kernels, what is needed is an adaptive density estimate

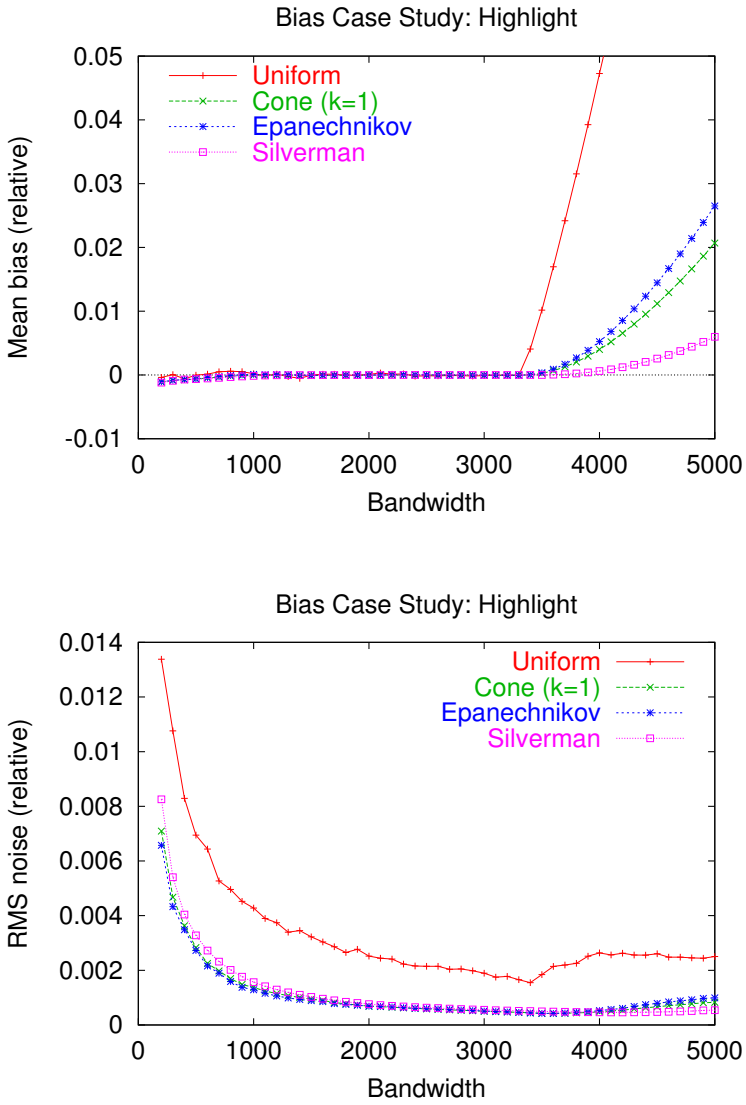


Figure 5.2: Mean bias (top) and RMS noise (bottom) for highlight case study using 1000 density estimates with uniform, cone, Epanechnikov, and Silverman kernels.

that adjusts the bandwidth to the illumination.

Wann Jensen recognised the problems inherent in his photon map early on and proposed a *differential checking* method [JC95, Jen01]. This limits the bandwidth during a photon gathering step if the irradiance monotonically increases or decreases as more photons are used. While simple and efficient, this approach is sensitive to noise and difficult to control.

Shirley et al [SWH$^+$95] used density estimation to obtain caustics on a polygonal mesh suitable for viewing with an interactive walkthrough. Their method uses

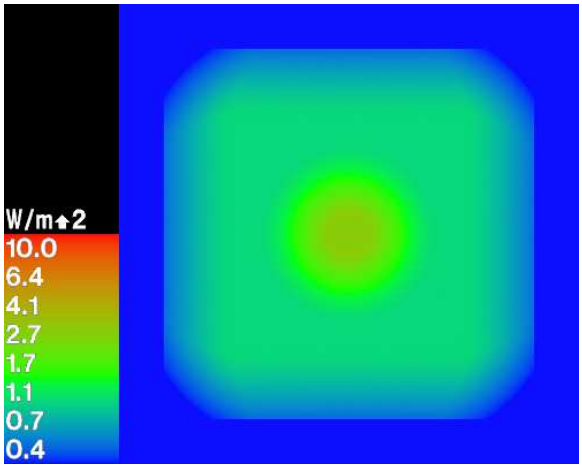


Figure 5.3: Falsecolour rendering of highlight case study using uniform kernel and a bandwidth of 5000 photons.

the Silverman kernel and counters boundary bias by reflecting samples near boundaries. The bandwidth is adapted to the particle density of each polygon, but is constant over the polygon surface.

Myszkowski [Mys97] was one of the first to thoroughly investigate bias in density estimation. He reported substantial bias with caustics (in excess of 10%) using nearest neighbour methods. He developed an enhanced nearest neighbour (ENN) method which uses a number of density estimates around the point under consideration to derive a combined bias/noise error estimate for the density estimate at the centre. The method evaluates this error for a number of bandwidths and selects one that minimises the error.

Walter et al [WHSG97] extended Shirley's polygonal mesh approach with more sophisticated compensation methods. They reduce boundary bias with an elaborate polynomial regression using local weights based on the Epanechnikov kernel. This requires transforming regression into a density estimation problem. In his PhD thesis [Wal98], Walter augments this with a bandwidth selection strategy based on perceptual error metrics for the noise and bias. He uses the Central Limit theorem to derive a target standard deviation for the initial bandwidth of a density estimate such that the expected fluctuations will not be visible to a human observer.

More recently, Hey and Purgathofer [HP02] have proposed an advanced density estimation method, which departs from the assumption that the area containing the photons is planar, thus compensating for topological bias. The actual area is derived by tesselating the geometry contained in the density estimate bandwidth. However, this conflicts with the philosophy of generality behind Wann Jensen's original proposal: the independence of photons from geometry. This constrains Hey's method somewhat, since some geometries cannot be readily tesselated, and tesselation may also introduce discontinuities in the reconstructed irradiance.

The bias compensating operator proposed here combines aspects from Wann

Jensen's and Walter's approaches. The operator also evaluates density estimates with increasing bandwidth for the same point under consideration (i.e. concentrically, in contrast to Myszwoski's approach). The crucial difference is the application of a binary search within a user-specified bandwidth range. Furthermore, the operator is more robust with respect to noise compared to Wann Jensen's method since it estimates the noise in the reconstructed irradiance. We apply the Central Limit Theorem as Walter has done, but in a different manner. Walter uses it to obtain an initial bandwidth and detects bias by thresholding deviations in the density estimates against the estimated noise using a user-specified scaling factor. By contrast, the proposed method uses the theorem to obtain the likelihood the deviations in the density estimates are due to noise, and use this as a heuristic to drive a binary search with the goal of minimising probable bias. Because our operator is designed specifically for more rigorous radiometric applications, we do not apply any perceptually driven rationale as Walter has done. Unlike Hey and Purgathofer's method, the operator knowns nothing about the geometry the photons reside on. Consequently, it does not compensate for topological bias.

5.3 A Bias Compensating Operator

The general idea of the proposed operator is to perform a binary search for an optimal bandwidth N_p during density estimation based on the likelihood that deviations in the irradiance are due to noise or bias. In order to determine this probability, we use a running average from which we estimate the variance.

Given an initial minimum and maximum bandwidth N_{min} and N_{max}, the operator first determines the expected value, or mean μ of the irradiance from density estimates $\hat{f}(\vec{x},1),\ldots,\hat{f}(\vec{x},N_{min})$ using 1, 2, ..., N_{min} photons for point $\vec{x}$:

$$\mu\left[\hat{f}\left(\vec{x},N_p\right)\right] = \sum_{j=1}^{N_{min}} \frac{w(j)}{N_w}\hat{f}\left(\vec{x},j\right), \tag{5.4}$$

$$N_w = \sum_{j=1}^{N_{min}} w(j), \tag{5.5}$$

where $w(j)$ is a normalised weighting function for each sample based on its bandwidth, which increases monotonically and ensures that density estimates with high bandwidth outweigh less reliable ones with low bandwidth. N_w is the weight normalisation factor. The choice of weights is investigated later in section 5.4. Good results are achieved with quadratic weights of the form $w(j) = j^2$.

μ is our first estimate for the irradiance; any bias present in the estimates up to N_{min} photons cannot be detected. We therefore assume that the irradiance within N_{min} is uniform and contains no bias.

Next, the operator estimates the sample variance $\hat{\sigma}^2$ based on the mean:

$$\hat{\sigma}^2\left[\hat{f}\left(\vec{x},N_p\right)\right] = \mu\left[\hat{f}^2\left(\vec{x},N_p\right)\right] - \mu^2\left[\hat{f}\left(\vec{x},N_p\right)\right]. \tag{5.6}$$

The weighting function $w(j)$ used to evaluate μ enforces a reduction in $\hat{\sigma}^2$ with increasing bandwidth. We consider $\hat{\sigma}^2$ an estimator because we make the assumption that the photons are independent and identically distributed. Strictly speaking, photon distributions are generally not independent and identically distributed, since there is a correlation along photon paths during distribution as well as in the resulting hit points on surfaces. However, we make the simplifying assumption that the photons within the bandwidth N_{min} are identically distributed, based in turn on our assumption that the irradiance within this bandwidth is uniform and unbiased.

The effectiveness of the bias compensation hinges on μ and $\hat{\sigma}^2$, we therefore try to extract a maximum of information to start with. Once initialised, these variables drive the binary search that follows.

Each recursion of the binary search consists of splitting the interval $[N_{min}, N_{max}]$ at $N_{mid} = (N_{min} + N_{max})/2$, and deciding which subinterval to recurse in based on a density estimate $\hat{f}(\vec{x}, N_{mid})$ using N_{mid} photons. As for the initial density estimates, we assume that the irradiance within the current lower bound for the bandwidth N_{min} is uniform and contains no bias. The choice of subinterval hinges on identifying the potential bias introduced by the estimate $\hat{f}(\vec{x}, N_{mid})$. We cannot evaluate the bias exactly because this requires the actual irradiance f, but we can assume that our average μ is a reliable estimate of f. From this we can estimate a combined error ε consisting of noise and an unknown amount of bias. This is simply the deviation of the new density estimate from the average μ:

$$\varepsilon\left[\hat{f}(\vec{x}, N_{mid})\right] = \hat{f}(\vec{x}, N_{mid}) - \mu\left[\hat{f}(\vec{x}, N_p)\right]. \tag{5.7}$$

Since we cannot separate the two components without knowning the exact bias, we must estimate the noise in ε. We can do this by determining the likelihood that the deviation is due to noise using the *Central Limit Theorem*.

Principally, this theorem states that the noise in a set of independent and identically distributed samples converges to a Gaussian distribution. Many unknown distributions in nature tend to be Gaussian, and we intend to apply these findings to the noise distribution in density estimates. Here we make the same assumption as in the evaluation of the variance estimator (equation 5.6), and assume the photons are independent and identically distributed.

We verify the applicability of the Central Limit Theorem emprically based on a series of density estimates in regions of uniform density. Figure 5.4 is a plot of the noise distribution relative to the mean of 50000 (unbiased) density estimates in a uniform photon distribution using 1000 photons each. The noise distribution matches the superimposed Gaussian, which is parametrised by the variance of the density estimates.

Thus, the probability p that ε is attributed to noise is:

$$p = e^{-\varepsilon^2\left[\hat{f}(\vec{x}, N_{mid})\right] / 2\hat{\sigma}^2\left[\hat{f}(\vec{x}, N_p)\right]}. \tag{5.8}$$

It is normalised to lie in the interval $[0, 1]$, and we can apply a rejection sampling scheme to decide whether we should treat ε as a noise artifact or as genuine bias.

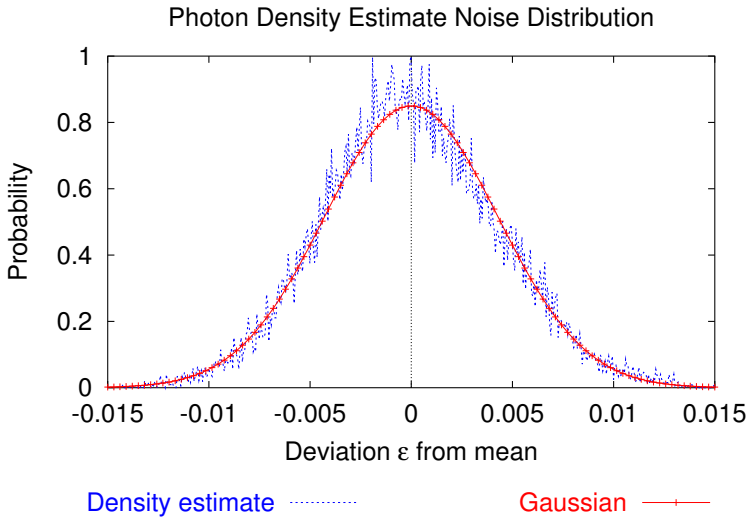


Figure 5.4: Noise distribution using 50000 density estimates of 1000 photons each. The distribution matches the superimposed Gaussian based on the variance of the density estimates, implying that the Central Limit Theorem is applicable to density estimation in areas of uniform density.

We do this by thresholding p against a uniform random variable $\xi \in [0,1]$, and accepting ε as noise if ξ lies within the Gaussian. This effectively amounts to applying fuzzy logic to the Gaussian boundary.

If $\xi < p$ we recurse in the subinterval $[N_{mid}, N_{max}]$ in order to increase the bandwidth. In this case the density estimate $\hat{f}(\vec{x}, N_{mid})$ is deemed unbiased and included in the average μ and estimated variance $\hat{\sigma}^2$ (equations 5.4 and 5.6), which are updated for the next iteration.

On the other hand, when p is low, recursion will likely take place in the subinterval $[N_{min}, N_{mid}]$ in order to reduce the bandwidth. The density estimate $\hat{f}(\vec{x}, N_{mid})$ is assumed to contain bias and μ and $\hat{\sigma}^2$ remain unchanged.

The operator thus effectively increases the bandwidth until a probable bias threshold is reached, in other words, when a systematic error protrudes from the noise "blanket". As with all binary methods, the recursion terminates when $N_{min} = N_{mid} = N_{max}$, with $\hat{f}(\vec{x}, N_{mid})$ being the final density estimate.

Since it is difficult to predict how many photons will actually be used by the operator, the gathering step will need to retrieve all N_{max} photons prior to applying bias compensation. An exception to this rule may be if there is some coherence between subsequent gathering steps (i.e. neighbouring points), but this is not the case in general. The photons must be sorted by distance in accordance with increasing bandwidth. In most implementations, the photon gathering routine does not fully sort the photons for reasons of efficiency, but merely keeps track of the furthest photon by maintaining a maxheap [Jen01, Sed92]. The entire set of gathered photons could be quicksorted prior to the binary search, but it is more efficient to

partition each recursive subset such that all photons in $[N_{min}, N_{mid} - 1]$ are closer than those in $[N_{mid} + 1, N_{max}]$, with the median at N_{mid}. This way only the photons are sorted in the recursed intervals, effectively implementing a partial quicksort.

We can tailour the operator to also handle chromatic bias by evaluating equations 5.4, 5.6, 5.7 and 5.8 in separate colour channels (this assumes the scalar representation of the irradiance $\hat{f}$ is an unweighted average over all colour channels, as would be the case for radiometric applications). No additional recursion of the binary search is required; the evaluations can be performed for each colour channel per recursion, based on the spectral irradiance. When deciding upon the interval to recurse in, we average p over all colour channels and threshold the result against ξ.

Since the operator is detached from the geometry, it can also be applied to density estimates with volume photon maps in participating media [JC98]. Volume caustics would benefit from this, although these phenomena are rarely encountered in practice.

The initial bandwidth range $[N_{min}, N_{max}]$ is subject to caveats which the user should be aware of. Setting N_{min} too low may bring out more detail in the illumination, but also reduce the accuracy of the initial expected value and variance estimates for bandwidths up to N_{min}. Consequently, the operator may mistake noise for bias upon evaluating density estimates at higher bandwidths and terminate the binary search with too low a bandwidth, resulting in excessive noise. On the other hand, setting N_{min} too high will stabilize the binary search in terms of reliably detecting bias, but any bias in the initial density estimates up to N_{min} will not be detected, blurring any detail therein. Setting N_{max} too high will impact performance during photon lookups. The user should therefore choose bandwidths in a reasonable range, typically where N_{max} is roughly 5–20 times N_{min}.

The bias compensating operator is summarised as pseudocode in algorithm 5.1. The routine *partition(i, j, k)* performs the partitioning of the photon subsets relative to the median at *j* as described above.

5.4 Results

Applying bias compensation to our highlight case study with initial $N_{min} = 50$ and N_{max} ranging from 50 to 5000 yields the error plots in figure 5.5. The effect of the sample weighting function $w(j)$ is shown for uniform ($w(j) = 1$), linear ($w(j) = j$), quadratic ($w(j) = j^2$), cubic ($w(j) = j^3$), and exponential ($w(j) = e^{0.003j}$) cases. Uniform weights perform almost as poorly as uniform kernel density estimates, while linear weights yield only a modest improvement. These functions are not steep enough with respect to the bandwidth, resulting in a large influence of noisy low-bandwidth density estimates on the expected value μ. This in turn raises the variance estimator $\hat{\sigma}^2$ and the probability of deviation p. Consequently, bias is often mistaken for noise and a high bandwidth results. Exponential weights also perform poorly because their initial gradient is too low, rising dramatically once the band-

procedure biascomp($\vec{x}, min, max$)
  $N_{min} = min$
  $N_{max} = max$
  gather N_{max} photons
  for $j = 1$ to N_{min} **do**
    partition($j, j+1, N_{max}$)
    get irradiance estimate $\hat{f}(\vec{x}, j)$ for j closest photons
    include $\hat{f}(\vec{x}, j)$ in average μ
  end for
  evaluate $\hat{\sigma}^2$
  while $N_{min} < N_{max}$ **do**
    $N_{mid} = (N_{min} + N_{max})/2$
    partition($N_{min}, N_{mid}, N_{max}$)
    get irradiance estimate $\hat{f}(\vec{x}, N_{mid})$ for N_{mid} closest photons
    $\varepsilon = \hat{f}(\vec{x}, N_{mid}) - \mu$
    $p = exp(-\varepsilon^2/2\sigma^2)$
    if random $\xi \in [0,1] < p$ **then** {ε probably noise, recurse in $[N_{mid}, N_{max}]$}
      include $\hat{f}(\vec{x}, N_{mid})$ in average μ
      update σ^2
      $N_{min} = N_{mid}$
    else {ε probably bias, recurse in $[N_{min}, N_{mid}]$}
      $N_{max} = N_{mid}$
    end if
  end while
return $\hat{f}(\vec{x}, N_{mid})$

Algorithm 5.1: Bias compensating operator

width already extends well beyond the bias threshold. On the other hand, quadratic and cubic weights yield consistenly lower bias and noise than the other functions. The results that follow all use a quadratic weighting scheme.

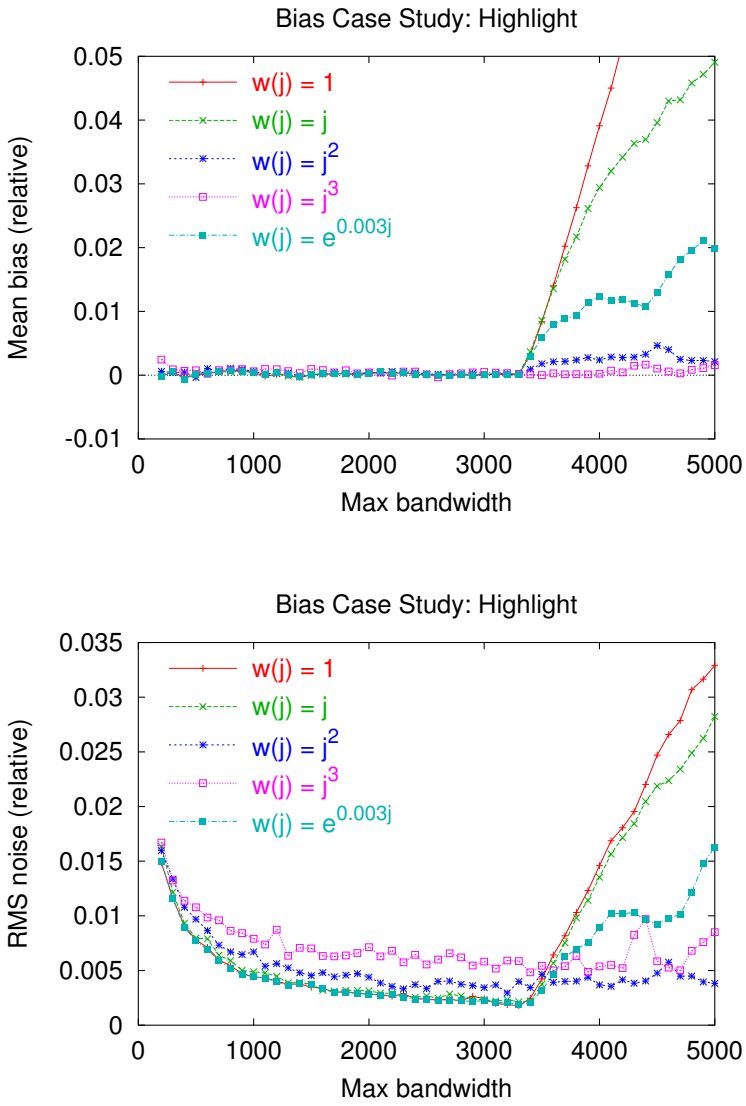


Figure 5.5: Mean bias (top) and RMS noise (bottom) for the highlight case study using bias compensation with uniform, linear, quadratic, cubic, and exponential sample weights $w(j)$.

Comparing figure 5.5 to the kernel density estimates in figure 5.2, we can see that the operator clearly reduces bias with quadratic weights at the expense of a slight increase in noise. Due to the stochastic nature of the bias compensation, there remains a small residual bias of ca. 0.2% from the occasional density estimate

in which bias is mistaken for noise. By contrast, the bias continues to rise with kernel based estimates.

Figure 5.6 is a falsecolour representation of the highlight case study rendered with bias compensation, and the bandwidths used in the rendering. Boundary bias is substantially reduced, and the highlight contours are preserved. The irradiance values agree favourably with the reference (10 W/m^2 in the highlight, 1 W/m^2 outside). Predictably, some noise is visible around boundaries. The bandwidth image confirms that the operator is consistent in using a low bandwidth at boundaries and a high bandwidth where the irradiance is uniform. The average bandwidth was about 1030.

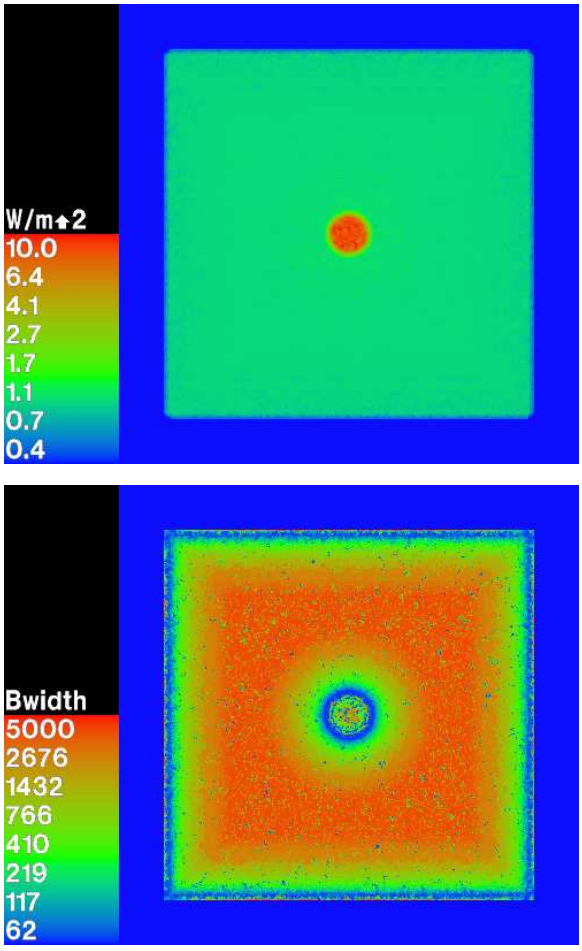


Figure 5.6: Falsecolour renderings of highlight case study with bias compensation using a bandwidth of 50–5000 photons (top), and resulting bandwidth (bottom). Bias compensation reduces the bandwidth around boundaries, while increasing it in regions of uniform irradiance.

Figure 5.7 is a cross-sectional irradiance plot of the highlight case study using a fixed bandwidth and bias compensation. Bias causes the flanks of the highlight

to flare and the plateau to taper. With bias compensation, the flanks are steep and the plateau's width is preserved to a large extent. The price to pay is the obvious noise in the highlight. The plot also exhibits a notable characteristic of the bias compensating operator: it overpredicts the bandwidth on the rising (left) flank as it makes the transition into the highlight, and underpredicts the bandwidth on the falling (right) flank as it leaves the highlight region. This explains the asymmetry of the bias compensated plot. This behaviour is directly influenced by the sample weighting scheme.

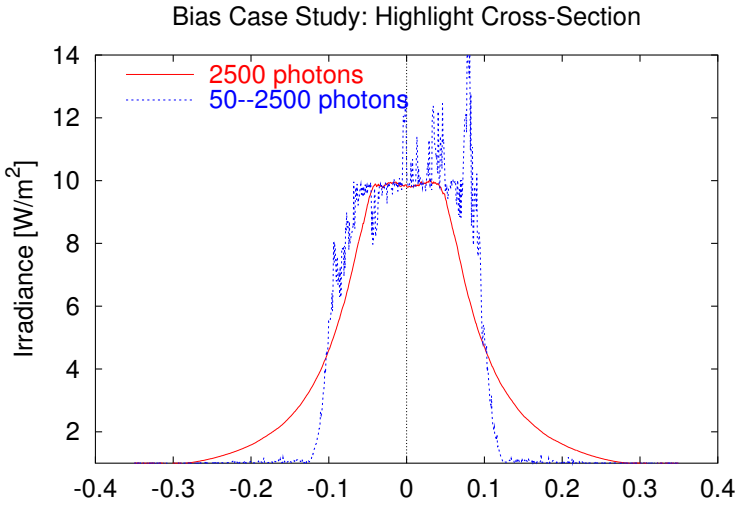
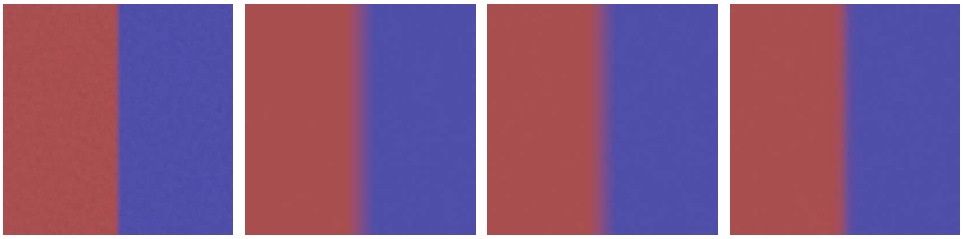


Figure 5.7: Cross-sectional irradiance plot of highlight case study using a fixed bandwidth of 2500 photons, and bias compensation with a variable bandwidth of 50–2500 photons. Bias affects the flanks of the highlight, which bias compensation strives to preserve. The resulting tradeoff in noise is evident.

Figure 5.8 shows the results of applying chromatic bias detection. The chroma case study consists of uniformly distributed photons of different colours divided into red and blue sections. The photon flux averaged over all colour channels is identical. Treating μ, $\hat{\sigma}^2$, ε, and p as scalars derived from the averaged (monochromatic) photon flux leads to chromatic bias; the boundary between the red and blue sections goes undetected, and a high bandwidth is used throughout. By evaluating these variables in separate colour channels, the operator can detect the boundary and modulate the bandwidth accordingly.

Having tested and developed the operator primarily to handle caustic contours, we now apply the operator to a smooth gradient. The photons for this case study are distributed in a linear density gradient to yield an irradiance ranging from 0 to 1 W/m^2 as shown in figure 5.9. Density estimates are performed on the green line passing through the plane. The resulting irradiance plot can be seen in figure 5.10 for density estimates using a fixed bandwidth of 50 photons and a variable



(a) (b) (c) (d)

Figure 5.8: Chromatic bias case study. Fixed bandwidths of (a) 50 and (b) 1000 photons, (c) monochrome and (d) chromatic bias compensation with bandwidth of 50–1000 photons.

bandwidth of 50–1000 photons. There is considerable noise in both cases, indicating that bias compensation detects the gradient and favours low bandwidths. More importantly though, there is no apparent bias with respect to the actual irradiance. However, because the operator is nondeterministic, it does introduce noisy patches as seen in the falsecolour renderings shown in figure 5.11. These arise from instances in which the gradient causes the operator to curtail the bandwidth dramatically. The basic problem here is that there is no information about the bandwidths chosen in the vicinity of these regions in order to enforce some degree of coherence.

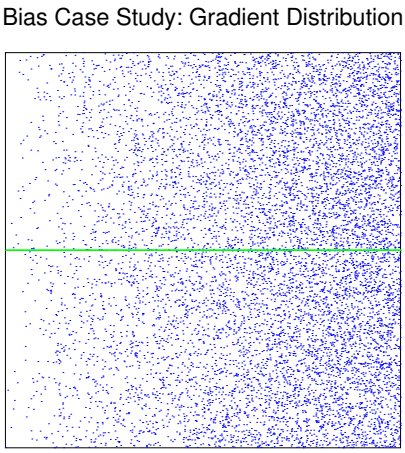


Figure 5.9: Photon distribution for gradient case study. The irradiance ranges from 0 to 1 W/m^2. Density estimates are taken at uniform intervals on the green line bisecting the plane.

An example of bias compensation applied to real caustics can be seen in figure 5.12. The metal ring caustic was rendered using ca. 500000 caustic photons. The renderings with fixed bandwidths used 50 and 2000 photons. While 50 photons

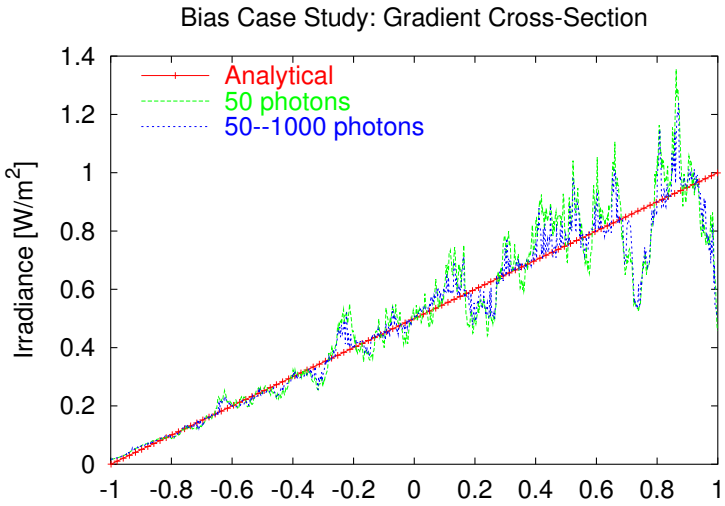


Figure 5.10: Cross-sectional irradiance plot of gradient case study comparing density estimates with a fixed bandwidth of 50 photons, density estimates using bias compensation with a bandwidth of 50–1000 photons, and the actual irradiance per analytical solution.

produce well defined contours, there is some noise in the interior of the caustic. Using 2000 photons drastically reduces the noise, but also blurs the contours which we would like to preserve. Applying bias compensation to this scene with a bandwidth of 50–2000 photons yields a satisfactory trade-off between the two extremes. Noise is reduced in the interior of the caustic by raising the bandwidth, yet contours are preserved by lowering it. The average bandwidth in the rendering is ca. 270 photons, indicating that low bandwidths to preserve details dominate in this scene.

Figure 5.13 is a more complex situation involving caustics. The venerable Cornell box is flooded with water, resulting in a caustic ripple pattern in the lower portion of the box from 250000 photons. Since the water surface is specular, the submerged portion of the box is entirely illuminated by caustics. Faint caustics are also visible on the ceiling from reflections on the water surface. These caustics are problematic because they are smooth and lack well-defined contours, and may therefore be mistaken for low frequency noise. As a consequence, bias compensation using 50–500 photons exhibits noisy patches in the submerged portion from inconsistencies similar to those in the gradient case study. Rendering with a fixed bandwidth of 50 photons produces obvious noise in the caustics, particularly on the floor. With a fixed bandwidth of 500 photons, the caustics are noticeably darker and lack sparkle as a result of proximity bias. There is also some minor boundary bias at the floor edges.

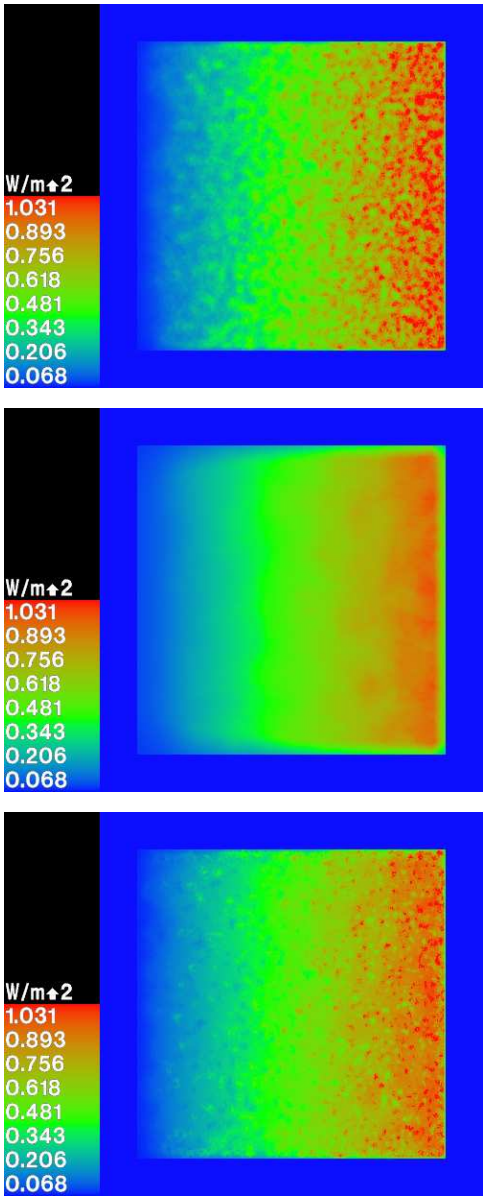


Figure 5.11: Falsecolour renderings of gradient case study with fixed bandwidths of 50 and 1000 photons (top and centre), and bias compensation using a bandwidth of 50–1000 photons (bottom). While bias compensation favours low bandwidths in order to reduce bias, the lack of coherence among neighbouring density estimates leads to noisy patches.

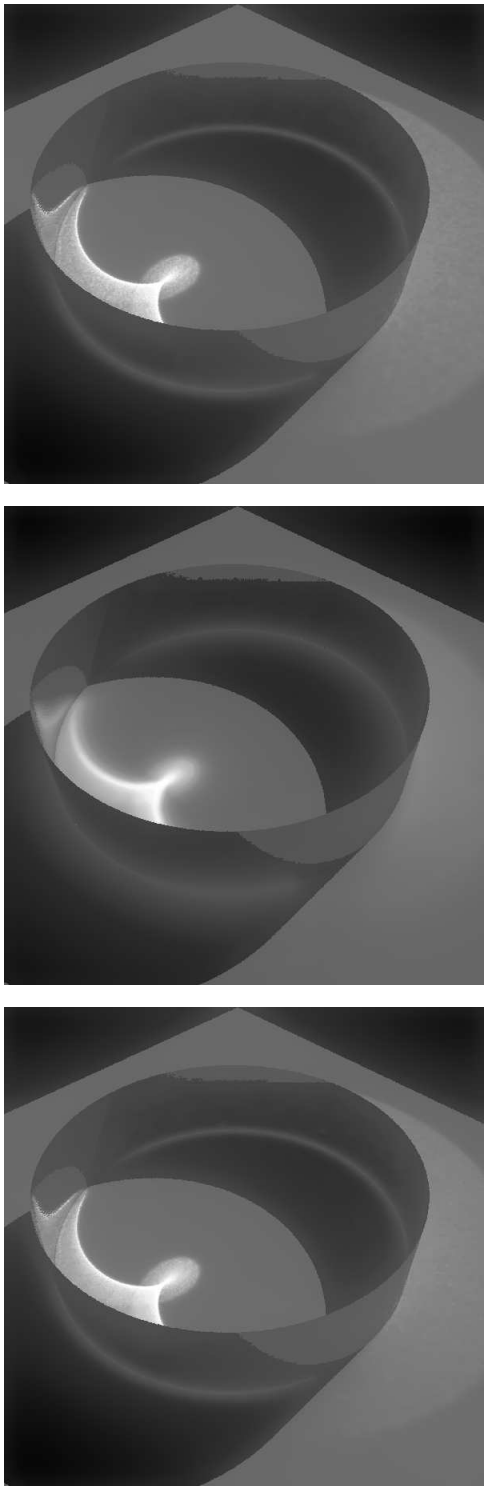


Figure 5.12: Metal ring caustic rendered with fixed bandwidths of 50 and 2000 photons (top and centre), and with bias compensation using 50–2000 photons (bottom).

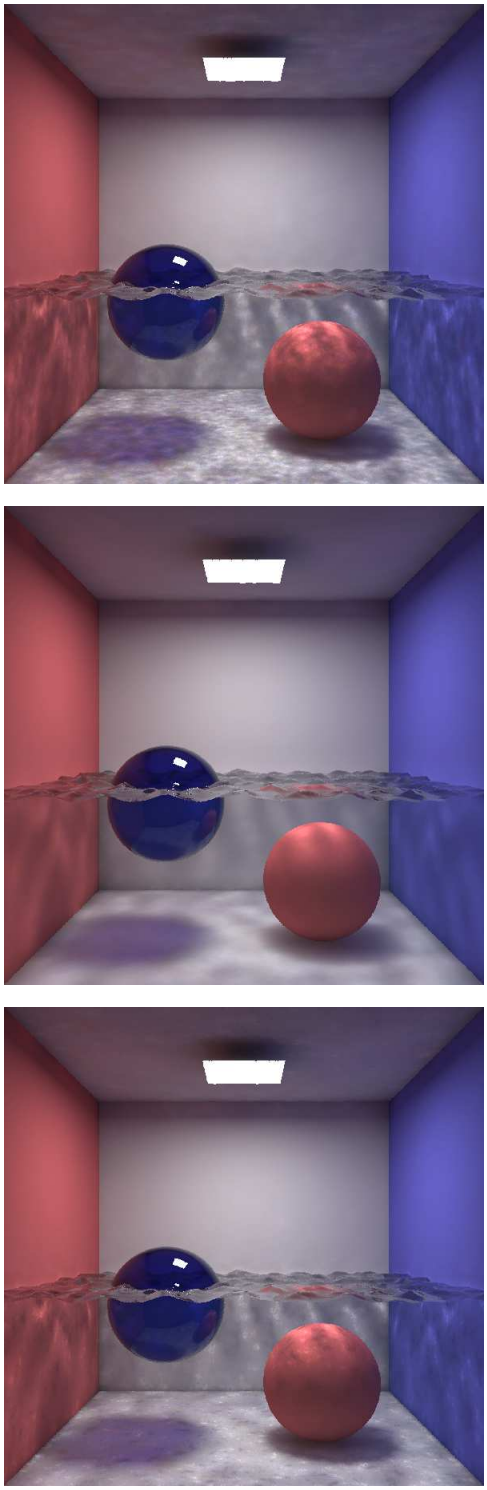


Figure 5.13: Flooded Cornell box with fixed bandwidths of 50 and 500 photons (top and centre), and with bias compensation using 50–500 photons (bottom).

5.5 Performance

Given the initial bandwidth range $[N_{min}, N_{max}]$, the operator performs N_{min} initial density estimates which requires sorting the N_{min} closest photons according to their distance. The subsequent binary search performs $\lceil \log_2(N_{max} - N_{min} + 1) \rceil$ iterations, during each of which the subintervals of size $N_{max} - N_{min} + 1, (N_{max} - N_{min} + 1)/2, \cdots, 2$ are partitioned. The partitioning routine is responsible for the bulk of the additional overhead incurred by bias compensation, since it must recurse within each subinterval until the median is in place (this interval subdivision can be easily implemented iteratively), which is tantamount to quicksorting a single element [Sed92]. Since the found photons are stored in a maxheap during the nearest neighbour search, they may infact be partially sorted. For this reason we minimise the average number of recursions while partitioning by choosing our median in the middle of the interval. In our results, we observed an average of only 1 recursion in the partitioning routine when choosing the median from the middle. By contrast, choosing a median at the interval extremes increased the average number of recursions to 4.

Another aspect affecting performance is the number of photons to search for prior to subjecting them to bias compensation. Conservativism could backfire as we may gather too few photons and would have to restart the search with a higher bandwidth. Ideally, we should find as many photons as will actually be used by the operator. This would require a prediction for the bandwidth, which is difficult to achieve reliably. Depending on the raytracing algorithm, a prediction may be made based on a density estimate history if there is some coherence between subsequent gathering steps (i.e. neighbouring points)[1]. In general, however, this is not the case, and the gathering step will need to retrieve all N_{max} photons to start with, as is done in the current implementation.

Performing the bias compensation during photon gathering is not an option, either. Wann Jensen's analogy that the nearest neighbour search entails "expanding a sphere around the intersection point until it contains enough photons" [Jen01] is intuitive, but misleading. In actuality, nearest neighbour search *constrains* the search volume until it contains the required number of photons. Bias compensation, on the other hand, would effectively require expanding the search volume during the binary search. Consequently, bias compensation cannot be performed "on the fly" during nearest neighbour search.

Table 5.1 lists rendering times for the examples in section 5.4 using the bias compensating operator compared to renderings using the fixed minimum and maximum bandwidths N_{min} and N_{max}. The location and number of density estimates is identical for fixed and variable bandwidths per scene. Total times refer to the time spent rendering photon contributions, and are expressed relative to N_{min} for each scene. Each total is broken down into the percentage spent in the density estimate

[1] Coherence could also be enforced through a caching and interpolation scheme of bandwidths at neighbouring points.

(summing of photon flux, also binary search and evaluation of μ and $\hat{\sigma}^2$ with bias compensation), photon lookup in the k-d tree, and photon partitioning (with bias compensation). We can see that bias compensation incurs an additional overhead of 20–30% compared to renderings using N_{max}. Most of this is attributed to photon partitioning, which typically takes up 10–20% of the total rendering time. The time spent in the density estimate is also increased by a factor of 1.5–4 by the binary search, although its contribution to the total is small. Not surprisingly, the bulk of the rendering time in all cases is taken up by photon lookups. These are particularly dominant with the metal ring, where bias compensation affects performance only marginally. This is because the renderings were done with an increased initial search radius for photon lookups in order to bring out the faint caustic from reflection off the outer rim.

Scene	Bandwidth	Time			
		Total	Density estimate [%]	Photon lookup [%]	Partitioning [%]
Highlight	50	1.0	5	95	0
	5000	107.3	7	93	0
	50/5000/1030	131.0	7	75	18
Chroma	50	1.0	5	95	0
	1000	22.0	5	95	0
	50/1000/655	29.3	13	71	16
Gradient	50	1.0	4	96	0
	1000	21.0	5	95	0
	50/1000/378	27.5	11	72	17
Metal ring	50	1.0	1	99	0
	2000	3.9	1	99	0
	50/2000/269	4.0	2	95	3
Cornell box	50	1.0	2	98	0
	500	9.6	1	99	0
	50/500/247	11.7	8	81	11

Table 5.1: Performance of fixed bandwidth vs. bias compensation. Total rendering times are relative to the minimum bandwidth for each scene. Each total is broken down into the percentage spent in the density estimate, photon lookup, and photon partitioning (when using bias compensation). Bandwidths for bias compensation are given as minimum/maximum/average.

5.6 Conclusions

This chapter has presented a breakdown of errors inherent in the photon map, consisting of noise and various forms of bias. These errors are inversely related to each other and subject to the density estimate bandwidth. In situations involving

caustics, low bias is preferred in order to preserve detail. On the other hand, in situations involving uniform irradiance, low noise is preferred. This implies that an optimal bandwidth must be dynamically adjustable to the illumination, motivating the development of a bias compensating operator.

The operator uses a binary search within a specified range for the optimum bandwidth. This search is governed by error estimates extracted from the reconstructed irradiance in order to identify probable bias using the Central Limit Theorem. Unlike previous work, the operator is specifically geared toward quantitative analysis such as applications in lighting design. It is conceptually simple and general enough to be used in most density estimation frameworks because it does not rely on additional information, but rather makes use of what can be deduced from the reconstructed irradiance.

Chapter 6

Analytical Validation

The validation of a global illumination algorithm and its implementation is a prerequisite to demanding practical applications such as lighting simulation. It serves as an indication of the error incurred in the simulation by comparing the results with a reference in a series of representative case studies to which the algorithm is applied.

The primary reference for such a validation are analytical case studies. These involve applying radiometric theory to very simple geometry for which a solution is tractable and can be computed by hand. The scope of such a validation is quite limited, but its results are most reliable. Consequently, there is an emerging demand for a standardised set of analytical test scenes in the global illumination research community. Szirmay-Kalos et al. [SKKA01] proposed such a set, and more recently, technical committee TC 3.33 of the Commission Internationale de l'Éclairage (CIE) has been formed specifically to develop a set of test scenes based on analytical solutions and measurements [MF03].

6.1 Validation Case Study

A very simple analytical solution based on a diffuse unit sphere containing an isotropic point light source at its origin can serve as a validation case study. This setup results in a constant irradiance E and radiance L at every point on the inner surface of the sphere. This type of scene is commonly referred to as a *furnace* [SJ00, SKKA01, MF03][1], since it is based on radiation transfer in a uniform closed environment. The sphere is a special case of the furnace which is particularly amenable to analytical solution, as demonstrated by Szirmay-Kalos [SKKA01] for diffuse and ideal specular surfaces when using arbitrary isotropic EDFs.

Since the illumination within the sphere is uniform, there is no proximity bias. There is also no boundary bias for density estimates on the sphere's surface. This setup effectively simulates an integrating sphere, originally used to measure

[1] The author was initially unaware of these references and developed the spherical case study independently. The references were subsequently added for completeness.

a lamp's emitted flux. Integrating spheres can also be applied as reflectometers to measure the reflectance of material samples.

Given the light source intensity I_l, the irradiance from the light source at a point $\vec{x}$ on the surface of the unit sphere is

$$E_l(\vec{x}) = I_l \frac{\cos\theta}{|\vec{x}|^2} = I_l. \tag{6.1}$$

RADIANCE does not support point sources, but this can be approximated by a sufficiently small spherical source so as to minimise obstruction of reflected photons. For the direct component, spherical sources are not partitioned and thus behave like point sources. For a spherical source of radius $r_l \ll 1$ and radiance L_l, the irradiance is then

$$E_l(\vec{x}) = L_l \cos\theta d\omega = L_l \pi \left(\frac{r_l}{|\vec{x}|}\right)^2 = L_l \pi r_l^2. \tag{6.2}$$

Given the sphere reflectance ρ, the irradiance on the inner sphere surface can be found by expanding the diffuse rendering equation into a geometric series, and taking the limit as the number of reflections approaches infinity:

$$\begin{aligned}
E &= E_l + \int_{\vec{\omega}_i \in \Omega} L(\vec{\omega}_i) \cos\theta d\omega && (6.3)\\
&= E_l + L \int_{\Omega} \cos\theta d\omega && (6.4)\\
&= E_l + \pi L && (6.5)\\
&= E_l + \pi \left[\frac{\rho}{\pi}\left[E_l + E\right]\right] && (6.6)\\
&= E_l + \rho E_l + \rho^2 E_l + \rho^3 E_l + \cdots && (6.7)\\
&= E_l \sum_{i=0}^{\infty} \rho^i && (6.8)\\
&= \frac{E_l}{1-\rho}. && (6.9)
\end{aligned}$$

Paradox though it may seem, the spherical case study also has the advantage of being exempt from topological bias[2]. To verify this, consider the geometry in figure 6.1. During a density estimate, the photons found on the inner surface of the sphere occupy a spherical cap, or zone, extending to the furthest photon with distance r_p. The estimated area populated by the photons (included in the kernel normalisation factor) will be $\hat{A} = \pi r_p^2$. The actual area A is that of the cap, and can be found from its solid angle ω:

$$A = \omega |\vec{x}|^2 = 2\pi(1 - \cos\theta). \tag{6.10}$$

[2]This holds not only for the unit sphere, but also in the general case.

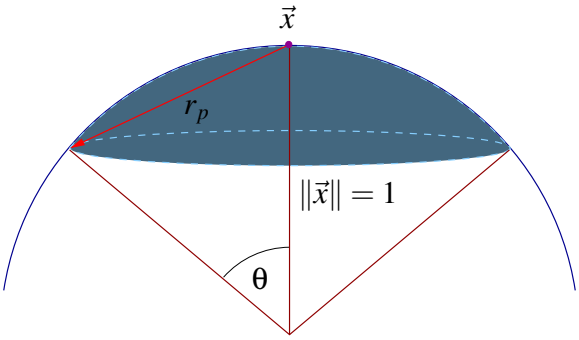


Figure 6.1: Geometry for photon gathering in spherical case study. The area of the cap occupied by the photons and the assumed (planar) area are equal, hence the case study is not subject to topological bias.

The angle θ subtended by the cap is related to r_p by

$$\sin\frac{\theta}{2} = \frac{r_p}{2\,|\vec{x}|} = \frac{r_p}{2}. \tag{6.11}$$

Substituting this into the area confirms the estimated and actual areas are the same:

$$\begin{aligned} A &= 2\pi\left(1-\cos\left(2\sin^{-1}\frac{r_p}{2}\right)\right) && (6.12)\\ &= 2\pi\left(1-\left(1-2\sin^2\left(\sin^{-1}\frac{r_p}{2}\right)\right)\right) && (6.13)\\ &= \pi r_p^2 = \hat{A}. && (6.14) \end{aligned}$$

6.2 Results

The spherical case study was used to validate the diffuse indirect illumination reconstructed from the global photon map. The reflectance of the sphere was chosen to be high ($\rho = 0.9$) in order to bloat the error propagated during particle tracing. The direct component E_l computed analytically by the stock RADIANCE code was subtracted, since we are only interested in the photon map's contribution. The validation consisted of constructing photon maps of $N_p = 50000, \ldots, 1000000$ photons, and using each map in an `rtrace` run with a bandwidth of $0.05N_p$ photons. Each `rtrace` run evaluated the irradiance at 1000 uniformly distributed points on the inner surface of the sphere. The irradiance was evaluated directly via density estimation at the points under consideration (no ambient rays were spawned for final gathering).

As with all Monte Carlo simulations, the quality of the photon distribution step hinges on the underlying random number generator (RNG). An RNG exhibiting strong correlation or spectral shifts in the sequences it produces can be a source of bias in the simulation. Three RNGs available in standard libraries on UNIX platforms

were used for the validation: SystemV-style `drand48()` and `erand48()`, and BSD-style `random()`. The former two are standard linear congruential generators and deemed obsolete according to the documentation, while the latter is (theoretically) a more sophisticated additive feedback generator. The RNG seeds were identical for each test run. In the case of `erand48()` separate generator states were maintained for photon emission, scattering, and russian roulette.

The flux stored with each photon was represented as a `float` triple in order to eliminate the bias introduced by the popular RGBE representation [War91b]. This format represents RGB flux as 3 mantissa bytes and a common exponent byte, resulting in rounding errors and therefore underestimation. The mantissas are only accurate to the second digit for saturated colours, yet considering the memory savings (almost 30% for this implementation), the minor bias incurred is a small price to pay. In this case study, the RGBE representation was found to introduce a bias of ca. -0.5%.

In the course of the validation it was also discovered that bias in low bandwidth density estimates is reduced by gathering an extra photon and using the averaged distances of the two furthest photons for the maximum photon distance r_p in equation 3.17 (the extra photon will be beyond this averaged distance and must be discarded). This confirms the validity of this optimisation suggested by Christensen [Chr01]. Without this measure, overprediction of 1% or more was observed for bandwidths below 50 photons. While such low bandwidths are not uncommon in practice (particularly for caustics), the merits of this optimisation are debatable given the small reduction in bias.

The results of the analytical validation are plotted as relative bias in figure 6.2. The particle transport simulation is carried to completion in the forward pass, hence the photon map is already converged, although there is the inevitable noise in the data. The noise tends to decrease as more photons are distributed and gathered. Increasing the number of photons only affects the noise, not the bias (recall that the spherical topology does not introduce bias of any kind with uniform irradiance). The graph also reveals erratic behaviour resulting from the different RNGs. The `random()` generator, which is generally accepted as a superior RNG, performed disappointingly with several seeds. The resulting illumination strays considerably above and below the zero axis. The simpler `drand48()` delivers slightly better results, but with a slight positive bias. The `erand48()` generator with separate states gives the most consistent results and appears to introduce less noise, presumably due to stronger decorrelation compared to using a single state. With this RNG, the bias typically lies within the $\pm$ 0.5% margin, which is more than adequate for even the most demanding applications.

The disparities of the tested RNGs are quite surprising, particularly given that `random()` is well tested and documented. Yet research in practical application of RNGs has shown that generators which perform well in statistical tests can fail in Monte Carlo applications [Hel98, Cod93]. While `random()` has a very large period (on the order of $2^{31} - 1$), this does not guarantee good spectral properties or the absence of correlation. Furthermore, unlike `drand48()` and `erand48()`, the se-

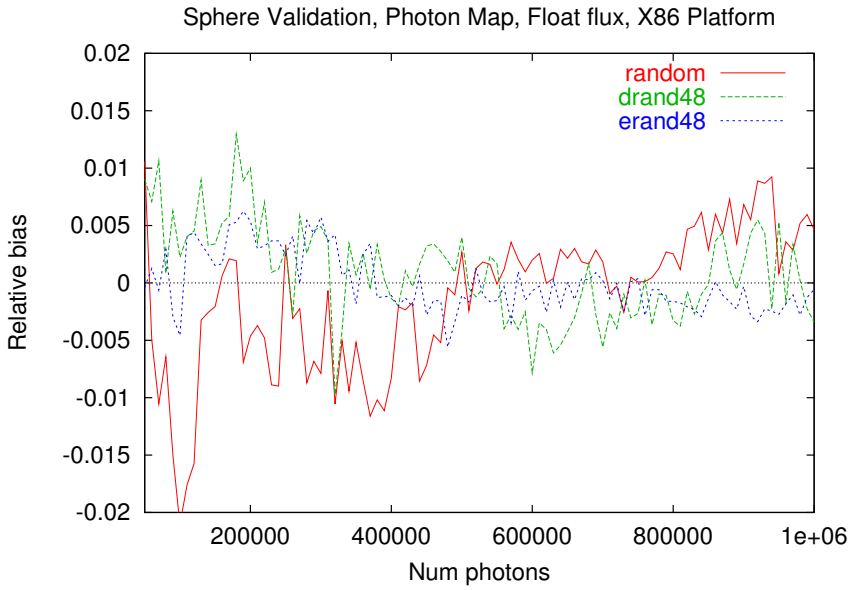


Figure 6.2: Relative bias for spherical case study applied to photon maps of 50k-1M photons using the `drand48()`, `erand48()`, and `random()` generators.

quences produced by `random()` are not portable across different platforms when using identical seeds. This appears to arise from differences in the size and initialisation of the state array used by the RNG. Portable sequences are a desirable property in rendering, particularly when correlation is needed between frames in animations [Chr01].

Applying the case study to RADIANCE Classic allows comparison with the photon map. While the photon map is already converged, we are interested in the convergence rate of RADIANCE Classic. The two governing parameters in this test are the number of ambient bounces (-ab) and the number of ambient divisions (stratified samples, -ad). The former affects the bias of the solution, while the latter influences the noise. As with the photon map, we subtract the direct component.

RADIANCE Classic's ambient illumination code is optimised to terminate rays based not only on their recursion depth, but also on their weights, which are thresholded against a weight limit (specified with the -lw option). Rays below this limit are truncated and an ambient value (specified with -av) is used instead. This ambient value can be moderated with the computed ambient illumination to form a moving average for scenes in which the indirect illumination fluctuates little. More accurate results can then be achieved with fewer ambient bounces by making an educated guess for a suitable ambient value.

The ambient illumination code attenuates the ray weights based on an assumed average reflectance of 0.5 for all surfaces. For scenes involving high reflectance such as the spherical case study, this clearly leads to premature truncation of am-

bient rays and thus substantial negative bias. To avoid this when applying the case study, ray weights were not thresholded (-lw 0). Furthermore, the direct threshold was also set to zero (-dt 0), which disables RADIANCE's optimisation of estimating the direct illumination from light sources based on statistical visibility estimates, and forces it to perform an explicit visibility test via shadow rays instead. Together, these measures force RADIANCE Classic to consider all relevant higher order ambient bounces and direct illumination components to attain maximum accuracy. We have thus tailoured RADIANCE Classic to suit the validation, whereas under standard conditions the described optimisations are in effect and can compromise the results.

A factor which must generally be taken into consideration for a validation is RADIANCE's irradiance cache. Its interpolatory behaviour results in smearing and proximity bias very similar to the photon map. However, proximity bias is not an issue in this case study and should only affect the noise, not the bias. Disabling the irradiance cache is not an option anyway, as it results in prohibitive computation times.

Figure 6.3 is a plot of the relative bias for `rtrace` runs using up to 40 ambient bounces with the number of ambient samples set to 1000. The plots used different settings of -av in order to assess its effect on the bias. With an ambient value of zero, RADIANCE Classic converges from below, starting with an enormous bias of -90%, which is of course expected with such a high reflectance. The bias drops below 1% after 33 ambient bounces. In practical application, RADIANCE Classic is rarely used with more than 7 ambient bounces. With an overestimated ambient value of $4.5E_l$, RADIANCE starts with a positive bias of 30% and already drops below 1% after 25 bounces. Note that the convergence rate itself does not improve, but rather the error introduced by the truncated ambient rays. Specifying the exact indirect illumination (uniform incident radiance of $9E_l/\pi$) for the ambient value, RADIANCE Classic is immediately converged with no bias to speak of. This is however a contrived situation, since in practice the exact ambient illumination is unknown and must be either estimated or determined via trial and error. While scenes with a high reflectance as in this case study are rarely encountered in practice with RADIANCE Classic, the validation shows that there is a potential for bias with reflectances over 0.5 when using conservative settings of -ab, or if -av is poorly chosen or not used at all.

Figure 6.4 compares the relative performance of the photon map and RADIANCE Classic with respect to computation time. For the photon map this also includes the time spent in the forward pass. Times are for a Pentium II 400 MHz system running Linux, and are presented on a logarithmic scale. RADIANCE's slow convergence is evident, while the photon map delivers results with sufficient accuracy in a fraction of the computation time. RADIANCE is only on par with the photon map if the optimum ambient value is used, although it is up to the user to find this setting. Estimating reasonable ambient values can be difficult and time consuming; the practical implications therefore speak in favour of the photon map.

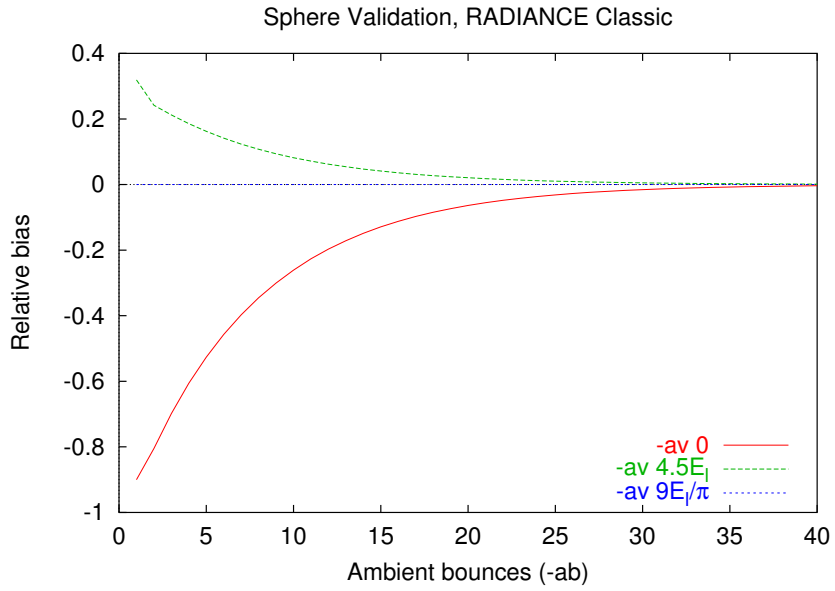


Figure 6.3: Relative bias for spherical case study with RADIANCE Classic.

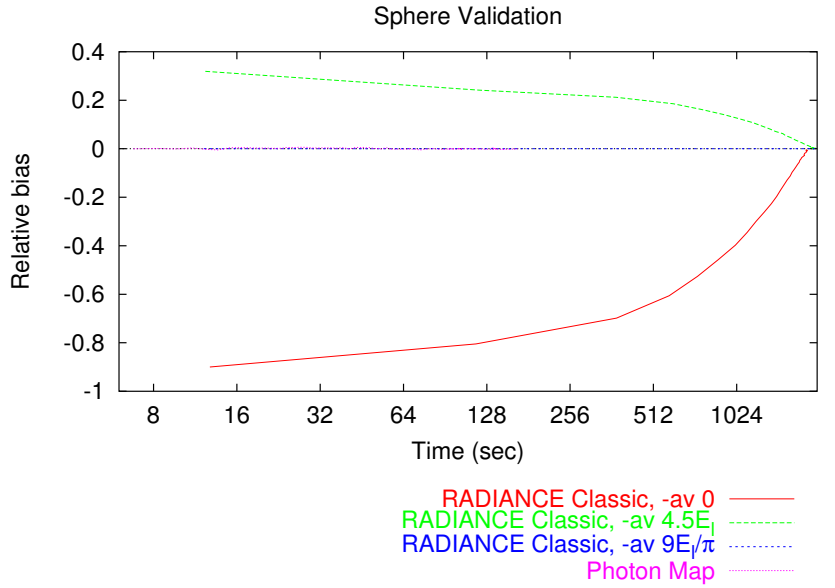


Figure 6.4: Computation times for spherical case study.

6.3 Validation with Lafortune BRDF

Because the radiance and irradiance is uniform for every point in the spherical case study, the setup cannot reveal errors in the sample ray distribution. In this section we take this paradigm further, and validate the photon map based on specular

reflection. Rather than using RADIANCE's built in Gaussian model, we use the popular multilobed Lafortune BRDF model as an example. This model will become significant in the next chapter due to its application in the experimental validation. As a prerequisite it must therefore be validated in conjunction with the photon map. Details on the Lafortune model can be found in appendix B.2.

The validation BRDF consists of a constant diffuse term $f_{r,d}$ and a single cosine lobe with coefficients C_x, C_y, C_z and exponent n. The derivation of the solution is greatly simplified by setting $C_x = C_y = 0$, whereby the cosine lobe is aligned with the surface normal. This results in rotational symmetry with respect to the incident and exitant directions, reducing them to angles θ_i and θ_r relative to the normal.

$$f_r(\theta_i, \theta_r) = f_{r,d} + (C_z \cos\theta_i \cos\theta_r)^n . \tag{6.15}$$

Such a material is obviously highly unrealistic, but still fulfills the requirements for a physically valid BRDF. In particular, Helmholtz reciprocity is satisfied if θ_i and θ_r are exchanged in equation 6.15.

The solution is based on a series expansion similar to equation 6.9, but burdened by the directional component of the BRDF. As a consequence, the irradiance is constant for any point on the inner sphere surface, but the radiance varies with the incident angle. The derived expression appears to defy a closed form representation, and is instead distilled into a recursive schema. Details of the full derivation can be found in appendix D.

Using the simplified Lafortune BRDF defined in equation 6.15, the solution for the irradiance E at any point on the sphere's inner surface is:

$$E = E_l [1 + a] , \tag{6.16}$$

where

$$a = 2\pi \left[\frac{f_{r,d}}{2} [1 + a] + \frac{C_z^n}{n+2} [1 + b] \right] , \tag{6.17}$$

$$b = 2\pi \left[\frac{f_{r,d}}{n+2} [1 + a] + \frac{C_z^n}{2n+2} [1 + b] \right] . \tag{6.18}$$

Strictly speaking, this solution is not analytical, since the expression must be evaluated numerically. It is nevertheless a means to obtain a theoretical reference value, and this is the essence of analytical validation in the broad sense.

The results of the validation are shown for the photon map in figure 6.5. As in the diffuse case, three RNGs were tried (`drand48()`, `erand48()`, and `random()`). As above, the photon flux was represented as `float` triples, and photon irradiance was visualised directly.

The RNGs exhibit similar behaviour as in the diffuse case, with `erand48()` once again producing the most consistent results in terms of proximity to the zero axis, with a minimal bias also within the $\pm$ 0.5% margin.

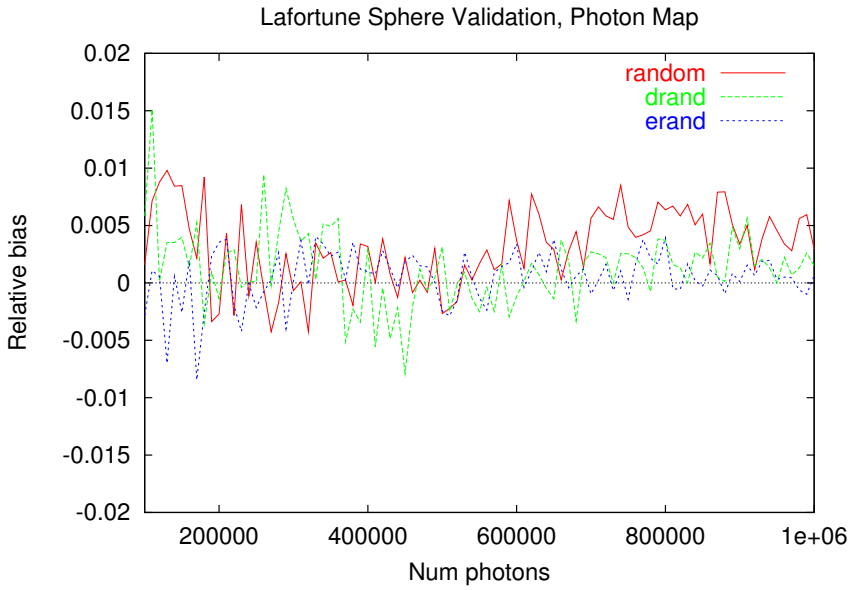


Figure 6.5: Relative bias for spherical case study with photon map using Lafortune BRDF.

6.4 Conclusions

Analytical validation is an effective means of ascertaining the accuracy and fundamental soundness of a global illumination algorithm. Though necessarily simple and constrained in scope, this approach defines a controlled environment which is generally more tractable than even the simplest setup used in a physical validation. The spherical case study has shown both for diffuse reflection and for a special case of the Lafortune BRDF model that the photon map produces minimal deviations within $\pm$ 0.5%. These deviations are predominantly attributed to noise, which gradually diminishes with increasing bandwidth. However, we have also seen that the deviations can vary significantly depending on the random number generator used, and that a well behaved generator is essential for unbiased Monte Carlo applications.

The photon map's primary advantage is that it is a complete solution to the indirect illumination in the scene. RADIANCE Classic, on the other hand, depends on the recursion governed by the number of ambient bounces (-ab). If this parameter is set too low, underprediction results. This is particularly important with high reflectances as is the case with our spherical case study. Furthermore, optimisations such as thresholding based on ray weights can also result in underprediction in such scenes due to the assumption that the average reflectance is 0.5, and modifying the ray weights accordingly. In order to obtain an unbiased solution comparable to that of the photon map, RADIANCE Classic requires over 30 ambient bounces for our case study – a value almost never used in practice. Naturally, the performance suffers and computation times are inordinately higher compared to photon map. All

things considered, the findings of the analytical case study speak well in favour of the photon map.

Chapter 7

Experimental Validation

The preceding chapter covered the analytical validation of the RADIANCE photon map. Where the complexity of the case studies precludes analytical solution, one may be tempted to resort to a numeric solution – only to realise that this is exactly what a global illumination algorithm effectively implements. In the case of novel algorithms like the photon map, however, comparison with numerical solutions provided by a tried and proven algorithm like RADIANCE Classic may be feasible, provided the latter has itself been validated.

The ultimate validation reference lies in experiments involving radiometric or photometric measurements of physical models. This is the most flexible validation method, with the scope only limited by the constraints of the experimental setup. However, it involves innumerable problems that can compromise the accuracy of the measurements, which is in general difficult to quantify. A large number of parameters can introduce errors, and the bulk of the time and effort invested in the validation is geared toward reducing this error in order to obtain a reliable reference to compare with the simulation. For this reason, even the simplest physical validation turns out to be a formidable, highly complex, and time consuming undertaking.

Experimental validations are very rare in computer graphics, primarily because they are difficult and generally go beyond the scope of what's required in mainstream applications – mainly to produce visually pleasing and convincing imagery. Not so with lighting simulation, where physical accuracy is paramount. With the advent of physically based rendering, validations have become increasingly important as a benchmark for physically based algorithms.

To date, there are very few documented instances of experimental validation applied to density estimation algorithms, hence part of the work in developing the RADIANCE photon map involved the challenge of validating the module with measurements. Drago and Myszkowski [DM01] are among the few who undertook a validation of a geometry-bound density estimation particle tracing algorithm similar to Walter's [Wal98] using lighting measurements of the Aizu University atrium, as well as perceptual comparison with the actual scene.

RADIANCE Classic, on the other hand, has been subject to more analysis.

Grynberg undertook one of the first attempts at an experimental validation [Gry89], while Mardaljevic [Mar99] undertook a more recent validation under daylight conditions using a variety of measured skies and importing these into RADIANCE using sky models. Reinhart and Walkenhorst [RW01] also undertook a validation of dynamic daylight coefficients using RADIANCE in a full scale test office fitted with external venetian blinds based on over 10000 measured sky conditions.

The purpose of the validation presented here was to test the principle behind the photon map, rather than to concentrate on the daylighting aspects. The validation should therefore be considered exemplary, and by no means comprehensive. Furthermore, the emphasis lies in methodology and error handling rather than the case studies themselves. Consequently, the case studies are kept simple for the sake of tractability.

7.1 Validation Methodology

The primary qualities one strives for in an experimental validation are:

Accuracy: the measurements should be a reliable reference for comparison

Reproducibility: the measurements are carried out in a controlled environment such that identical setups produce identical results for different measurements and with a minimum of preparation

Tractablity: the complexity of the setup is limited to the extent where it can be broken down in such a way that sources of error can be isolated both on the physical and simulation side

The guiding principle behind the methodology chosen for the photon map is to break the validation up into tractable *component* case studies which can be checked individually. Where possible, the accuracy of the measurements can be assessed by comparing them to analytical estimates. Once the accuracy of the components has been ascertained, they can be combined to form more complex *compound* case studies which would be too complex and intractable to validate as a whole. It is assumed that the validity of the component case studies ensures the validity of the compound case studies.

7.2 Validation Setup

The validation setup is shown in figure 7.1. It consists of the following components:

- A test box with a window. A very simple geometry to study light transport with.

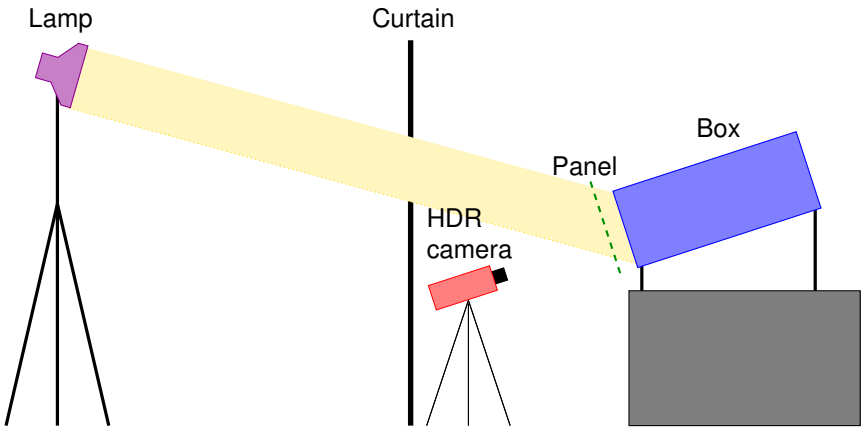


Figure 7.1: Experimental validation setup

- Absorbing and reflecting materials on the inner walls of the box. Absorbing materials allow isolation of individual light transport modes in component case studies, thereby contributing to tractability.
- Illuminance sensors in the box. These are the instruments which deliver the data for the validation.
- An artificial light source. This is preferable to daylight in terms of stability and contributes to reproducibility.
- A diffuse panel which is placed over the front of the box to determine the lamp's emission distribution.
- An HDR (high dynamic range) camera to acquire intensity distribution images on the panel.
- An absorbing curtain with a cutout to constrain the incident light to the solid angle subtended by the front of the box in order to suppress stray reflections.

These components are described in detail in the following subsections.

7.2.1 Test Box

The simple test box was constructed with the proportions of a typical office (0.6 $\times$ 0.5 $\times$ 1.2 m). It consists of an aluminium frame and panels held in place by clamps for easy assembly. The box's interior is accessible via a removable lid. A window admits light and can be supplemented with a daylight system. Apertures at the front and back faces can accomodate an HDR camera lens to allow visual monitoring of the lighting levels within. These apertures are covered with the wall material during illuminance measurements. Figure 7.2 is an HDR image of the test box.

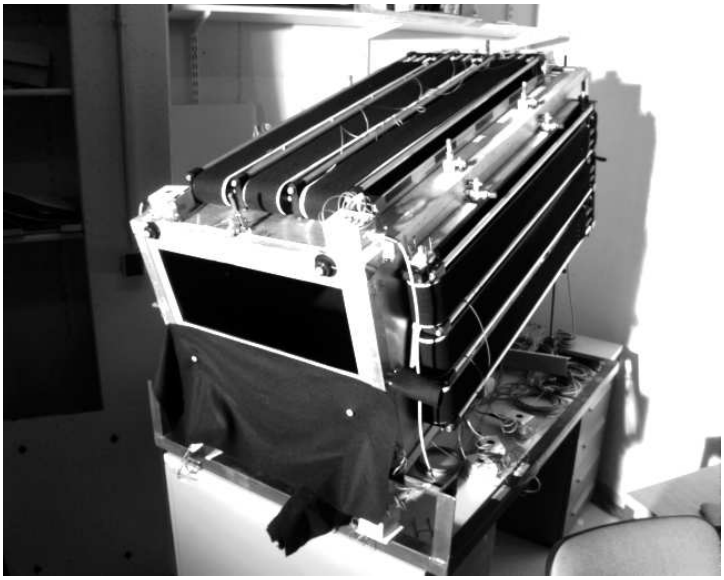


Figure 7.2: Experimental validation test box

7.2.2 Illuminance Sensors

The top, bottom, and sides of the box have slots spanning the entire length for three sensors which measure illuminance in the box interior and can be positioned at arbitrary distances to the window in 1 cm increments. It was assumed that the influence of these interior sensors on the interreflection within the box was negligible, given the small sensor area.

The sensors are guided along belts. Each belt is covered with the box's current wall material in order to prevent light leakage through the slot. The material strip is fastened to the belt with a spring mechanism. This ensures that the strip moves with the belt, yet can be exchanged with little effort. The three belts per side are driven in parallel by a shaft, which is operated manually via a handwheel with detents. This setup can also be upgraded with a stepper motor, which could perform the measurements much faster under software control. The sensor guidance mechanism is shown in figure 7.3. Positioning errors for this mechanism are estimated at $\pm$2mm, and considered negligible in their effect on the measurements.

The sensors are equipped with a $V(\lambda)$ filter which defines a photometric spectral response. They cover a dynamic range of 100000 lx which is divided into a low and high range with transition at 10000 lx. The sensor error is a combination of absolute and relative components relating to its electrical and photometric characteristics. The absolute error is mainly incurred by its circuitry (voltage to current converter, amplifier) and is specified as $\pm$0.05% of the maximum for each range, i.e. $\pm$5 and $\pm$50 lx for the low and high ranges, respectively. The relative error accounts for deviations in the $V(\lambda)$ filter, the cosine factor correction, and the linearity of the signal, specified as $\pm$4%. Each sensor is individually calibrated against a

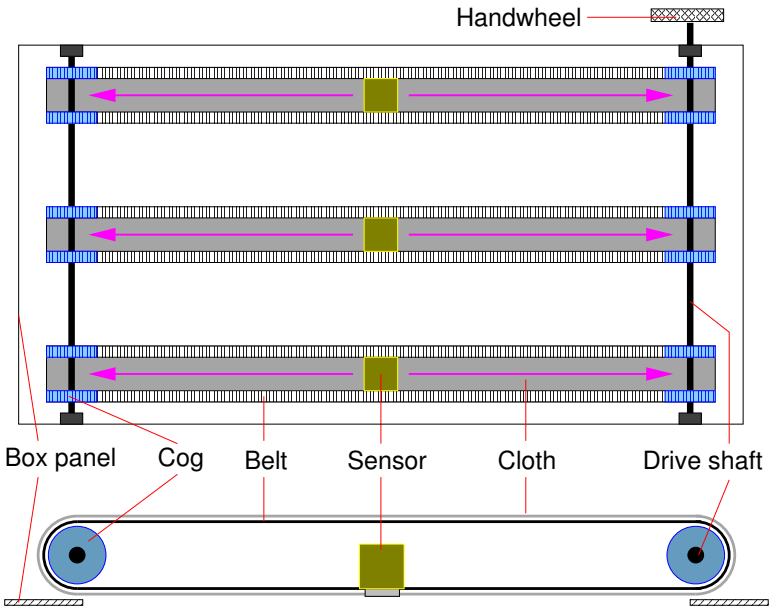


Figure 7.3: Schematic of sensor guidance mechanism, top and side views.

reference sensor prior to the validation.

Data from the illuminance sensors is collected with an Agilent multimeter connected via RS-232 interface to a garden variety PC, running in-house software under Linux. The error incurred by the multimeter is on the order of 0.1% and therefore negligible. Five measurements are made per sensor position in order to obtain averaged values.

7.2.3 Light Source

Rather than using sunlight as a light source, which is subject to fluctuation and requires the use of empirical sky models, it was decided to perform the validation with a powerful floodlight with known intensity distribution. This ensures that the measurements are carried out in a controlled and reproducible environment. The floodlight (Philips ArenaVision) is equipped with a 1 kW HID (high intensity discharge) bulb and mounted on a fully extended tripod. Due to limited ceiling clearance, the box is also inclined to obtain an effective incident angle of ca. 60° to the box window. Stray light is blocked by a curtain between the lamp and box, with a cutout large enough to just illuminate the box's front face.

Four additional exterior sensors were mounted on the front face of the box near the window corners in order to monitor the direct illuminance from the light source. This avoids having to depend on specifications from the lighting manufacturer, since the emission is subject to a *maintenance factor* accounting for depreciation and aging. This factor can vary dramatically [DM01] and linearly influences the results of

the simulation. The actual direct illuminance can thus only be determined accurately by measurement, which is what the exterior sensors accomplish.

These sensors helped establish the warm-up duration for the lamp at the beginning of each validation measurement (ca. 30 minutes). The sensors also revealed slight temporal fluctuations in the lamp illuminance during the course of measurements (figure 7.4). These fluctuations directly affect the illuminance of the interior sensors. To cancel them, the data from the interior sensors is taken relative to the average illuminance of the four front sensors, which is measured simultaneously with the interior sensors for each sensor positioning cycle. This dimensionless relative illuminance data can be compared and plotted coherently. An additional benefit is the fact that the lamp model used in the simulation can have an arbitrary emitted luminance.

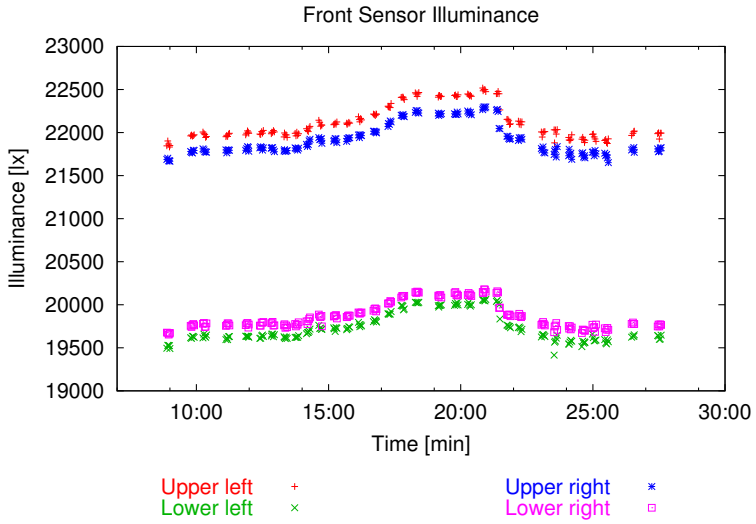


Figure 7.4: Illuminance plot of the four exterior sensors at the box window corners monitoring the light source over a period of ca. 19 minutes. The illuminance for each sensor fluctuates by up to $\pm$2%. The light source EDF (covered in section 7.3.6.2) is responsible for the higher illuminance at the upper sensors, and is not subject to fluctuation.

7.2.4 Materials

The inner surfaces of the test box were intended to be lambertian, and the component case studies required absorbing (black) and reflecting (gray) walls. Unfortunately, real materials don't work this way; no material is perfectly diffuse or 100% absorbing. Candidate materials had to be selected based on their approximation of our ideal, and simulated accordingly. Therein lay the biggest challenge of the entire validation. One candidate was barium sulphate, a highly reflective substance commonly used to coat the interior of integrating sphere reflectometers. Although this

would have come close to our ideal, it exhibits some degree of specularity and was rejected on grounds of cost and the compound's fragility, which would have made it difficult to reassemble the test box without chipping the coating. The BRDFs of candidate materials were analysed with a goniophotometer, and the choice was narrowed down to textiles. The wall material chosen was heavy molleton in black and light gray. The molleton was attached to the aluminium panels with velcro, which facilitated exchanging the wall material.

The validation also called for a case study involving caustics. However, many daylight systems are difficult to model realistically. Particularly systems composed of plastics such as those discussed in section 4.2 have manufacturing tolerances which can vastly affect the resulting caustics observed in reality. These result from deformation of the material during the moulding or extrusion process, often in the form of warped surfaces or rounded edges. While they can be measured, manufacturing defects are difficult to model, and the simulation will deviate significantly if these are not fully accounted for.

In the light of these difficulties, validation case studies involving problematic systems like Y-glass and CPCs were foregone in favour of the most primitive daylight system: the light shelf. The light shelf is simply a large metal profile designed to reflect light towards the ceiling. While real light shelves are often curved, the light shelf used in the validation consisted of a flat sandblasted aluminium profile. Sandblasting was necessary to eliminate any anisotropy from grooves which accumulate on the surface during the rolling process in manufacture. It also has the desirable side effect of lending the surface a glossy finish, which is less critical to sample than highly peaked specular reflection and facilitates comparison between forward and backward raytracing.

7.3 Validation Procedure

Figure 7.5 gives an overview of the validation procedure. It is composed of modules pertaining to specific problems in the validation. The arrows in the diagram indicate dependencies, where intersections should be interpreted as branches; a path from A to B denotes that B requires A and depends on it as input. By this token, the diagram conveys the complexity of the validation. The modules and their constituents are as follows:

BRDF Acquisition: the BRDFs of the materials used in the validation must be obtained in order to be taken up in the simulation. This requires two tasks:

- **Goniophotometer measurement:** the BRDF is measured from material samples using a goniophotometer
- **Artifact removal:** spurious shadowing artifacts incurred by the goniophotometer during acquisition must be removed because they adversely affect the simulation.

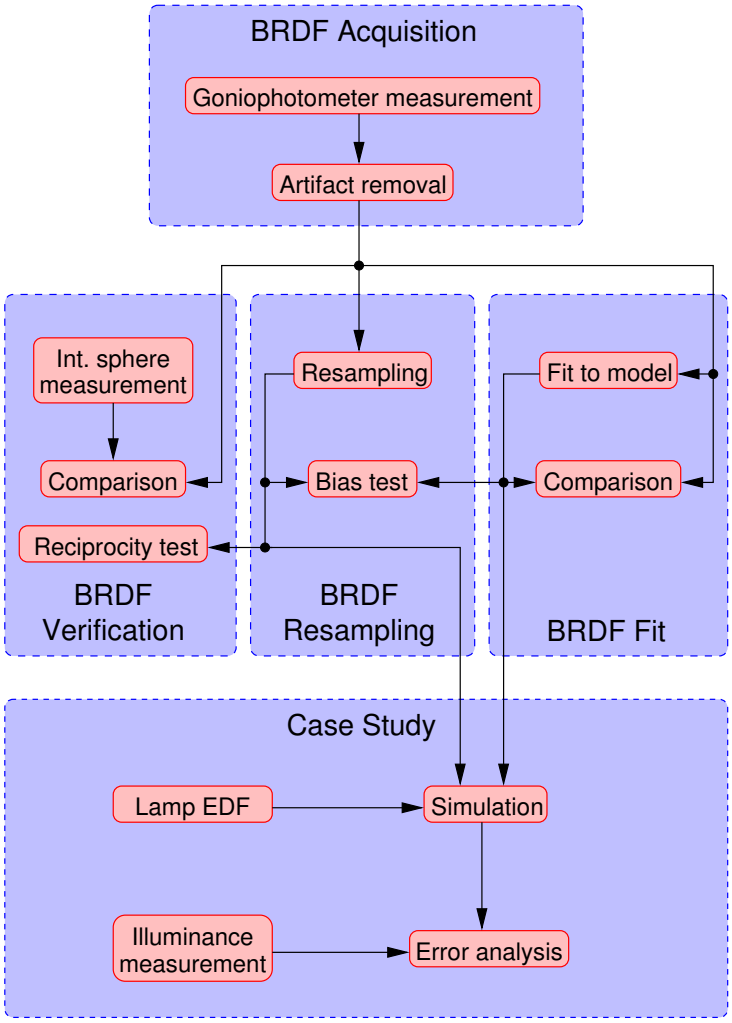


Figure 7.5: Experimental validation procedure overview. The validation is composed of modules encompassing tasks aimed at handling a specific problem in the validation. The arrows denote dependencies.

BRDF Fit: a BRDF model is fitted to the measured BRDF data subsequent to artifact removal. The fitted model is required to sample the measured BRDF data in the simulation. Here we perform the following tasks:

Fit to model: we fit a suitable BRDF model to the measured BRDF using a standard fitting algorithm.

Comparison: We assess the quality of the fit by comparing the fitted model with the measured BRDF.

BRDF Resampling: the measured BRDFs must be evaluated for arbitrary incident and outgoing directions in the simulation. Since the goniophotometer mea-

sures at a finite number of non-uniformly distributed locations, the BRDF must be resampled in the simulation. The resampling requires two tasks:

Resampling: development of the resampling method. This requires the measured BRDF data as input after artifacts have been removed.

Bias test: verification of bias inherent in the resampling method. Here we resample and compare the fitted BRDF model.

BRDF Verification: here we assess the accuracy of the goniophotometer measurement. Though a preliminary task with high priority, the BRDF resampling is partly a prerequisite. We verify the BRDF as follows:

Comparison with integrating sphere: we obtain the reflectance from an integrating sphere reflectometer and compare it to the reflectance computed from the goniophotometer data. This can be done immediately after BRDF acquisition.

Reciprocity test: we verify that the goniophotometer data complies with Helmholtz reciprocity. This requires the resampled BRDF in order to evaluate it at arbitrary reciprocated directions. Furthermore, the potential bias incurred by the resampling must be known before this test can be applied with confidence.

Case Study: the preliminary work done in the above modules culminates in the individual case studies. Each case study involves the following tasks:

Lamp EDF: we acquire the emission distribution of our light source via a high dynamic range camera. This must be done per case study since the box is modified and therefore moved.

Simulation: the case study simulation requires the EDF, measured BRDF (subject to resampling), and fitted BRDF model as input (used as sample ray distribution). The output is simulated illuminance data sampled at the sensor positions.

Illuminance measurement: measured illuminance data is acquired from the sensors in the validation setup.

Comparison and derivation of error: we compare the measured and simulated illuminance and derive error bounds which take the constituent errors of the validation setup into account.

Each of these procedural components is described in detail in the following subsections.

7.3.1 BRDF Acquisition

7.3.1.1 Goniophotometer Measurement

The BRDFs of the molleton and aluminium were obtained from Fraunhofer ISE's goniophotometer [AB94, ABH98]. The device has four degrees of freedom as shown in figure 7.6: two for the sample disc, and two for the detector arm. The incident beam is obtained from a fixed lamp (usually xenon) with condensor and diaphram. The orientation of the sample disc relative to the incident beam accounts for the incident direction, while the orientation of the detector relative to the disc accounts for the exitant direction.

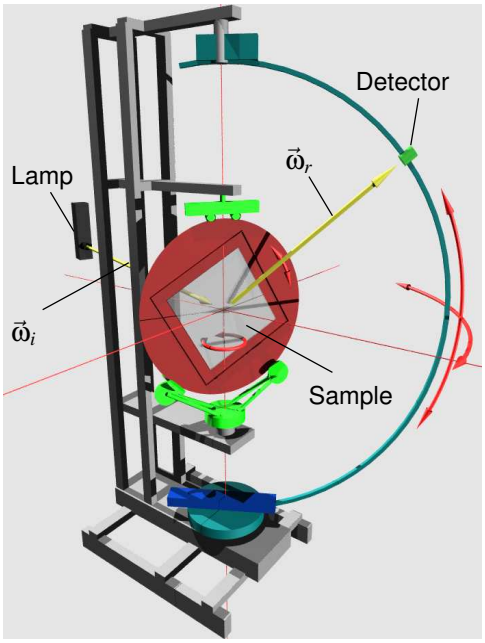


Figure 7.6: Diagram of Fraunhofer ISE's goniophotometer used for BRDF acquisition. The device has four degrees of freedom: two for the sample disc, and two for the goniophotometer detector. The incident beam is obtained from a fixed lamp.

Due to optical and mechanical limitations of the device, the maximum angle measured was 78°, yielding only a partial BRDF. These limitations result from light source occlusion towards grazing angles by the clamp holding the sample in place. Furthermore, the illuminated sample area elongates to the point where it extends beyond the sample boundaries, causing a falloff in measurement. Lastly, the raw goniophotometer data must be divided by $\cos\theta_r$ per definition in order to obtain the BRDF, which introduces numerical instabilities at grazing angles. These limitations preclude obtaining data for grazing angles (where BRDFs really exhibit interesting behaviour) and are common in goniometric BRDF acquisition.

BRDF measurement is a laborious task with numerous problems pertaining to

calibration and physical limitations. Installations for BRDF acquisition are few and far between, and the results can vary dramatically depending on design and calibration. To illustrate this, Leonard and Pantoliano [LP88, LPR89] conducted a round robin exercise in which BRDF measurements of identical samples were collected from a number of facilities. The outcome of the experiment was devastating, revealing massive deviations. These difficulties are of course carried over in any physical validations which use these measurements, further complicating the validation process. As a consequence, the data should be verified with other references (e.g. integrating sphere reflectometer) before its application in the validation.

The BRDF plots obtained from the goniophotometer (figure E.1) reveal the light grey molleton is not entirely diffuse; they exhibit highly diffuse reflection at normal incidence, but with an emerging specular component towards grazing angles. The position of the peak is also shifted well beyond the direction of specular reflection. This phenomenon is termed *off-specular* reflection, and is typical of very rough materials; off-specularity increases with roughness, with the peak position tending towards grazing angles for very rough surfaces [TS65, TS67]. This characteristic results from a decrease in apparent roughness along the incident and exitant directions, giving rise to specular reflection. The specular component rises dramatically at grazing angles to the point where it significantly affects the illuminance measurements in the box. This is a challenge because it has to be taken into account, but is difficult to simulate with most BRDF models. It should be noted that of all the diffuse candidate materials, this was the most benign – foiled by reality once more.

At first glance, the aluminium's glossy reflection as evidenced from the BRDF measurements (figure E.2) exhibits more benign behaviour. One would therefore expect this material to be readily amenable to BRDF modelling. However, closer examination of the BRDF plots reveals two phenomena which may pose problems for the BRDF model. Firstly, the peak is slightly off-specular. Secondly, it is asymmetric, elongating along θ_r towards grazing angles. These characteristics are also difficult to account for in a BRDF model, and constrained the choice of model to use in the validation.

7.3.1.2 BRDF Artifact Removal

A problem deserving attention during BRDF acquisition were the goniophotometer shadowing artifacts visible particularly in the molleton BRDF plots (figure 7.7). These come in two forms: shadows cast by the goniophotometer arm, giving rise to rifts across the interior of the plots, and shadows cast by the sample clamp, giving rise to an abrupt falloff at the periphery of the plots.

The goniophotometer arm shadow dominates at normal incidence and wanders towards the periphery of the plots with increasing incident angle. It arises from an obstruction of the fixed incident beam by the goniophotometer arm bearing the detector, which pivots about the sample (figure 7.8). The arm casts a penumbra which tends to be quite broad, with an estimated coverage of roughly 20° in the plane of exitance, corresponding to the angle α_s between the planes delineating

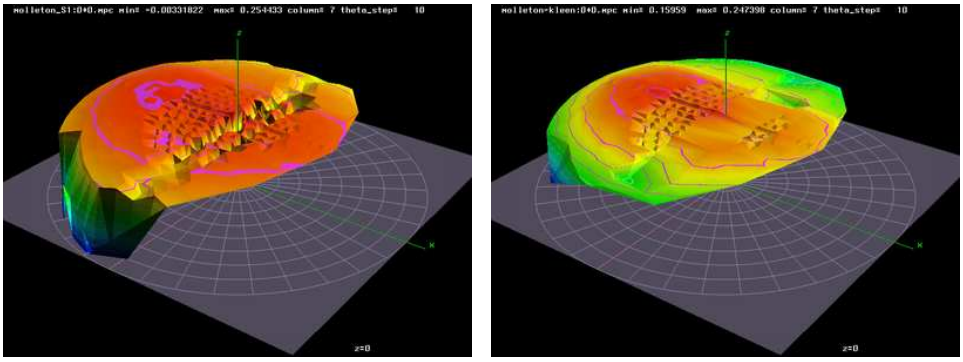


Figure 7.7: Molleton BRDF at normal incidence before (left) and after (right) artifact removal. Artifacts arise from shadows cast by the goniophotometer arm in the interior of the plots, and shadows cast by the sample clamp at the boundary.

the penumbra limits in figure 7.8.

The goniophotometer arm shadow is usually not a problem with specular materials, since these artifacts only arise if the BRDF reflects a significant component in the vicinity of the incident direction, which is typically the case with diffuse or retroreflecting materials. For this reason, the specular aluminium BRDF is largely devoid of these artifacts, whereas the mostly diffuse molleton BRDF suffers considerably. This warrants correcting these artifacts to some degree.

To this end, the BRDF data was subjected to an algorithmic artifact removal which simply deletes samples which are affected by shadows cast by the sample clamp or the goniophotometer arm (the resulting "gaps" are filled in during BRDF lookups by the nearest neighbour resampling described in section 7.3.4). The algorithm is extremely simple and relies on heuristics and manual specification of shadow parameters rather than autodetection.

A BRDF sample $f_r(\vec{\omega}_i, \vec{\omega}_r)$ is removed if:

- the sample lies in the goniophotometer arm's penumbra defined by the user specified angle α_s, i.e. if

$$\frac{\alpha}{\alpha_s} \leq \frac{1}{2}, \tag{7.1}$$

  where α is the angle between the incident direction $\vec{\omega}_i$ and the goniophotometer arm plane defined by the exitant direction $\vec{\omega}_r$.

- the sample is below a user specified fraction of the BRDF average for the measured angle of incidence, and its angle of exitance θ_r is above a user specified θ_{max}:

$$f_r(\vec{\omega}_i, \vec{\omega}_r) < c_b \overline{f_r}(\vec{\omega}_i), \qquad \theta_r \geq \theta_{max}. \tag{7.2}$$

  This treats the boundary falloff caused by the shadows cast by the sample clamp as outliers. Imposing a limit on the exitant angle prevents removal

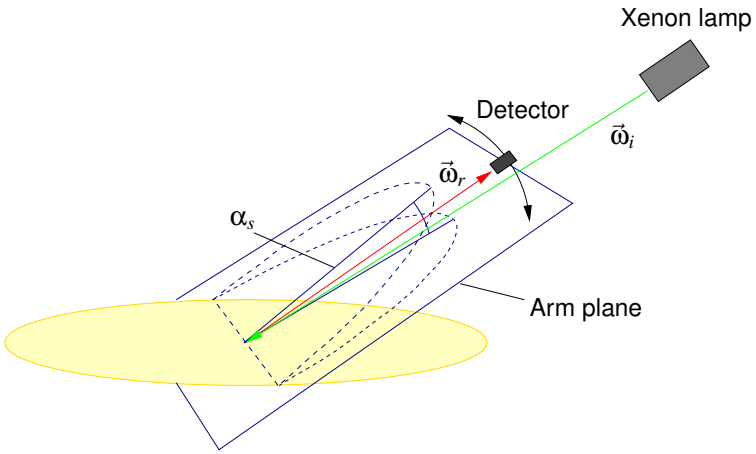


Figure 7.8: Goniophotometer shadowing artifacts during BRDF acquisition. The goniophotometer arm (not shown) bearing the detector pivots about the sample, and can obstruct the incident beam, attenuating the measured BRDF. The shadow covers an angle α_s in the plane of exitance. Rays incident within this range are obscured.

of samples in the interior. This heuristic can be improved for mostly diffuse BRDFs such as molleton by using a cosine weighted average. This reduces the impact of specular components at grazing angles on the average.

Figure 7.9 shows statistics of the artifact removal. For both molleton and aluminium, the percentage of samples removed drops with increasing incident angle. This is explained by the concentration of samples at the specular components by the goniophotometer's adaptive scanning algorithm. Since these wander outside the influence of the goniophotometer arm shadow with increasing incident angle, the number of samples removed decreases. This effect is particularly evident with the aluminium BRDF, while it is more gradual with the molleton due to the emergence of the specular component towards grazing angles.

Figure E.3 shows the plots of the molleton BRDF after artifact removal. The rifts visible across the plots in figure E.1 arising from the goniophotometer arm shadow have been largely removed, and the falloff at the boundaries arising from the sample clamp shadow has been clipped. Figure E.4 shows plots of the aluminium BRDF after artifact removal. Here the goniophotometer arm shadow is only significant up to an incident angle of ca. 15°, beyond which it lies outside the specular peak. Due to the negligible diffuse component, there are no significant boundary artifacts. The goniophotometer arm shadow was specified for both materials with an angle of $\alpha_s = 20^\circ$, while boundary artifacts beyond $\theta_{max} = 70^\circ$ were thresholded with a factor of $c_b = 0.7$ for the molleton, and $c_b = 0$ for the aluminium.

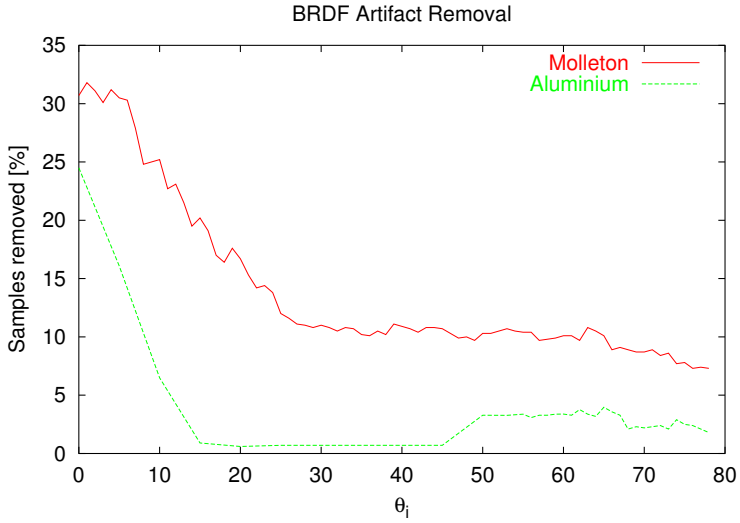


Figure 7.9: Percentage of BRDF samples removed during artifact removal as a function of incident angle. The percentage drops with increasing incident angle as the specular component wanders outside the influence of the goniophotometer arm. Because the molleton BRDF's specular component emerges towards grazing angles, the effect is more gradual with this material.

7.3.2 BRDF Verification: Integrating Sphere

We verify the goniophotometer measurement by comparing its reflectance to the results obtained from Fraunhofer ISE's integrating sphere reflectometer. The photometric directional-hemispherical reflectances of the black and light gray molleton at normal incidence were 0.0164 and 0.7094, respectively. The light gray molleton BRDF data obtained from the goniophotometer correlates with the findings obtained with the integrating sphere; the reflectance increases towards grazing incident angles due to the off-specular component (figure 7.10). Here we also see the quantitative impact of the shadowing artifacts incurred by the goniophotometer during BRDF acquisition by comparing the reflectance of the BRDF before and after artifact removal. The drop in reflectance due to shadowing artifacts in the original BRDF accounts for an average deviation of ca. 4% and a maximum of nearly 10%. These deviations would become significant in case studies involving interreflection, which justifies the task of artifact removal.

The reflectance from the goniophotometer BRDF is obtained by integrating over a Voronoi diagram of the measured points [ABH98, Har95]. Due to the limited measurement range of the goniophotometer, data will be missing at the hemisphere boundary. The reflectance is therefore extracted in the form of an upper and lower bound based on assumptions of what lies beyond the boundary. The upper bound is based on the assumption that the BRDF is constant beyond the measurement

limit, while the lower bound is based on the assumption of an abrupt falloff to zero beyond the boundary. Given the molleton's off-specularity, clipping the BRDF to zero beyond the boundary would clearly underestimate the reflectance. On the other hand, the good agreement of the upper bound with the reflectance obtained from the integrating sphere confirms the validity of the measurements conducted with the goniophotometer.

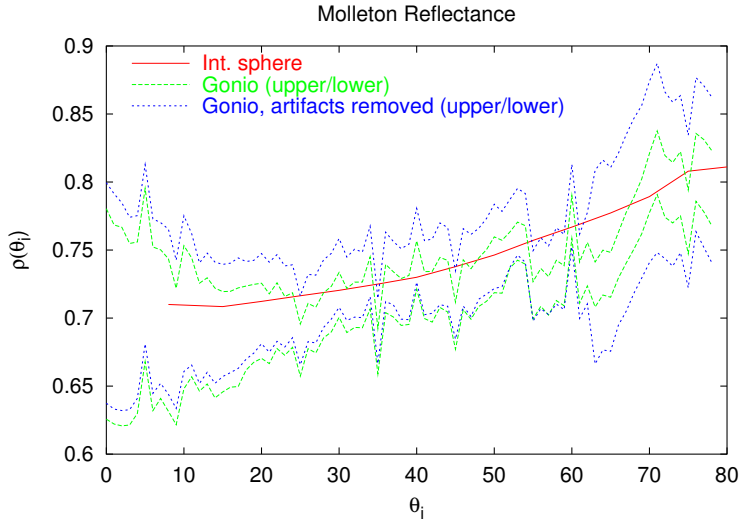


Figure 7.10: Molleton reflectance obtained with integrating sphere reflectometer and goniophotometer. The goniophotometer reflectance is shown before and after artifact removal. The upper and lower bounds for the goniophotometer reflectance arise from treating BRDF data beyond the measurement limits as constant and zero, respectively.

7.3.3 Fit to BRDF Model

The parameters for the BRDF model used in the validation simulations are extracted by *fitting* the model to the goniophotometer data. Fitting describes the class of numerical optimisation methods by which parameters for a mathematical model approximating a data set are found such as to minise the deviations based on some error metric [PTVF92]. Depending on the specific algorithm, model, initial parameters, and size of data set, this can be a very slow process. To ensure a good fit, the unknown parameters must be estimated with reasonable accuracy, which requires some experience with the specific model used. The BRDF model was fitted simultaneously for all measured incident angles after artifact removal, yielding one parameter set. These parameters are subject to the caveat that, since they were obtained from a partially measured BRDF due to goniophotometer constraints, there are unknown deviations towards grazing angles.

RADIANCE's built in gaussian model (defined in appendix B.1) cannot account for all the characteristics of the BRDFs used in the validation (notably off-specularity), and this is obvious in the BRDF plots of the fitted models (figures E.5 and E.6). The alumium BRDF is approximated with an average difference of 0.153 between the fitted model and the measured BRDF. The fit to the molleton BRDF deviates as a result of the off-specularity with an average difference of 0.035. For both materials, the fit tends towards physically implausible values, producing a specularity greater than 1 for the aluminium, and a reflectance greater than 1 for the molleton. These values were clamped to 1 in figures E.5 and E.6, although this is obviously unrealistic and indicates the model's unsuitability to off-specular materials.

In the light of the problems with the molleton BRDF and the principal role the material plays in the validation, the choice of BRDF model was reconsidered, leading to the investigation of the Lafortune model (defined in appendix B.2) as an alternative. Because it is more general, it can account for off-specularity, although the price to pay is the larger number of parameters, which affects the complexity and stability of the fit.

The popular *Levenberg-Marquardt* nonlinear least squares fitting method [PTVF92, AB95] proved unsuitable for the Lafortune fit. This method is based on a gradient descent within the parameter space and involves inverting a matrix consisting of partial derivatives for each parameter. With the Lafortune model, this matrix tends to exhibit singularities due to the tight interdependence of the parameters, causing the matrix inversion and consequently the algorithm to fail. In contrast to the Ward model, the Lafortune model suffers from the fundamental problem that modifying one parameter requires modifications to other parameters during each iteration of the fit.

A more robust but also less efficient alternative is the *downhill simplex* method [PTVF92, NM65]. This method does not compute derivatives, but merely requires evaluations of the BRDF model. For n parameters, the algorithm operates on an n-dimensional geometric simplex in the parameter space consisting of $n+1$ vertices. During each iteration, it attempts to contract the simplex towards a minimum vertex. If the parameter space contains a large number of minima, the algorithm must repeatedly restart with new parameters in order to find a global minimum. The Lafortune model's parameter interdependence and the relatively high dimensionality of the fit (four parameters per lobe) gives rise to such a situation, dramatically impacting the fit's performance. No fitting algorithm can be expected to find a global minimum in such a parameter space in reasonable time. Lacking viable alternative BRDF models, an optimimum was therefore chosen from a limited number of minima.

The error metric used in the fit was a straight sum of differences over all points. Although other metrics such as reflectance and cosine weighting were tried, the results of the fit did not differ remarkably.

The Lafortune model was fitted for multiple lobes in conjunction with a Lambertian term. The effect of the number of cosine lobes on fit deviation is shown in figure

7.11. The graph indicates that 3 lobes and a Lambertian term are sufficient for the validation simulations, as there is negligible improvement with additional lobes. BRDF plots of the fitted Lafortune model can be found in figures E.7 and E.8.

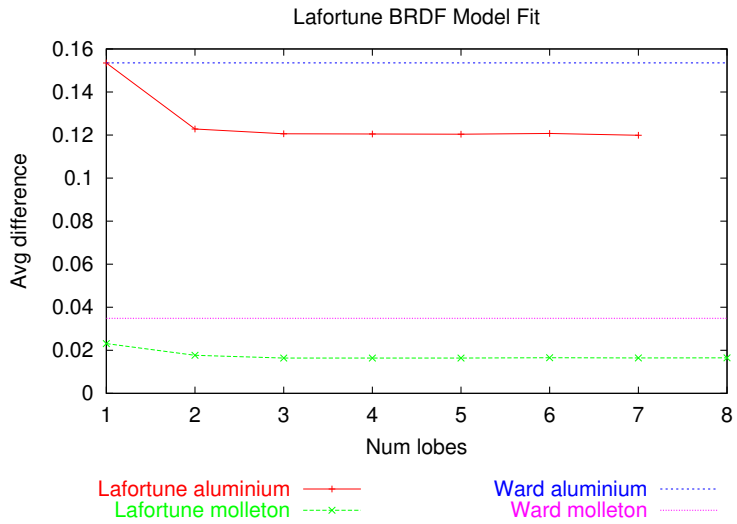


Figure 7.11: Number of cosine lobes in the fitted Lafortune BRDF model versus fit error compared to the fitted Ward BRDF model. The error metric is the average difference between measured and fitted data points

With the aluminium BRDF, the Ward and Lafortune models exhibit somewhat complementary behaviour; while off-specularity appears to be the dominant source of error with the Ward model, the Lafortune model's errors arise mainly from the peak's elongation. It is apparent from the BRDF plots that the Ward model actually elongates the peak towards grazing angles, whereas the Lafortune model does not, giving rise to two lobes flanking the location of the measured peak in the difference plots. While the aluminium BRDF still poses problems for both the Ward and Lafortune models, the molleton BRDF is substantially better approximated by the latter, reducing deviations by 50% compared to the Ward model. Disregarding residual boundary artifacts, the corresponding BRDF difference plots are relatively flat compared to those of the Ward model.

The average differences between the fitted Lafortune model and the measured BRDFs were 0.12 for the aluminium and 0.016 for the molleton. This compares favourably to the respective fit deviations of 0.15 and 0.035 obtained with the Ward model. In both cases there is a residual error resulting from noise and smaller artifacts in the measured BRDF. There remain, however, some deviations from BRDF features which neither the Ward nor Lafortune models can account for by principle.

We could use the fitted BRDF model directly in the validation simulations, in which case we would expect the fit deviations to be the principal source of error. We could obtain a spatially distributed estimate for this error which is propagated along with the ray luminance in the simulation, but we could do even better by sam-

pling the actual measured BRDF in the simulation, thereby eliminating this error. Instead of using the fitted BRDF model in the simulation directly, we use it as a PDF to sample the measured BRDF. The justification for doing so is that the Lafortune BRDF is fitted to the measured BRDF, making it a suitable PDF for applying importance sampling to the latter. Although this approach demotes the fitted BRDF model to an ancillary role, deviations from the measured data will introduce noise to the sampling, which is why a suitable choice of model and goodness of fit are still important. In what follows, all validation simulations use the fitted Lafortune BRDF model as PDF.

7.3.4 BRDF Resampling

Integrating the measured BRDF into the simulation requires storing it in a suitable data structure which allows quick retrieval for a given pair of incident and outgoing directions. Furthermore, the data must be available for arbitrary directions, not just those which were actually measured. This is particularly important because the goniophotometer does not sample the reflecting hemisphere at regular intervals, but rather adaptively, concentrating samples where high gradients (specular components) are detected [AB94]. Retrieving the measured BRDF at arbitrary points therefore effectively constitutes resampling the data, which is not a trivial matter and can itself contribute some deviations.

Point resampling is particularly relevant in 3D mesh reduction and specific algorithms have been developed in this field. A simpler approach was used to resample the measured BRDF by using the already available photon map code and storing the measured data in a bilevel kd-tree as shown in figure 7.12. The top level kd-tree contains 3D keys corresponding to the cartesian[1] incident directions $\vec{\omega}_{i,j}$ from the set of measured incident directions $\Omega_{i,m}$. Each incident node with direction $\vec{\omega}_{i,j}$ points to a kd-subtree containing the set of corresponding measured exitant directions $\Omega_{r,m}(\vec{\omega}_{i,j})$ (recall that the goniophotometer scans adaptively, thus $\Omega_{r,m}$ is a function of $\vec{\omega}_{i,j}$). Each such exitant direction $\vec{\omega}_{r,k}$ has a node in the subtree containing a measured BRDF sample $f_{r,m}(\vec{\omega}_{i,j}, \vec{\omega}_{r,k})$.

The resampled BRDF $\hat{f}_{r,m}(\vec{\omega}_i, \vec{\omega}_r)$ can be evaluated with a bilevel nearest neighbour lookup for the closest matches to the incident and outgoing directions under consideration (figure 7.13). The lookup consists of finding the N closest incident directions to $\vec{\omega}_i$ in the set of incident samples $\Omega_{i,m}$. For each such direction $\vec{\omega}_{i,j}$, we perform a lookup in its corresponding subtree for the closest exitant directions $\vec{\omega}_{r,k}$ from the set of exitant samples $\Omega_{r,m}(\vec{\omega}_{i,j})$. We then sum these BRDF samples weighted by a function w based on the sample distance to obtain a subsample $\hat{f}_{r,m}(\vec{\omega}_{i,j}, \vec{\omega}_r)$ for each incident direction. These subsamples are, in turn, weighted

[1] 2D polar coordinates would yield a more compact representation at the expense of cumbersome handling of the $0^\circ/360^\circ$ wraparound.

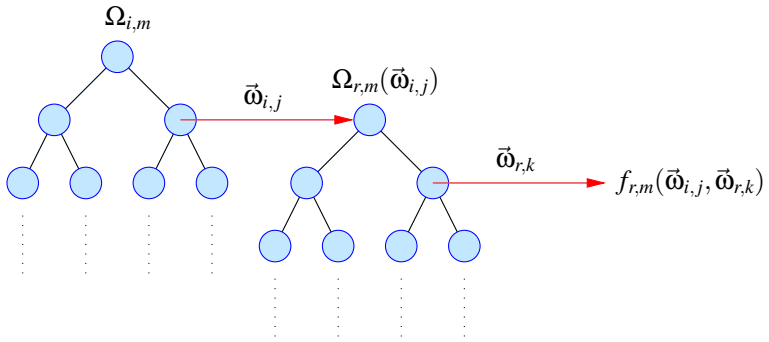


Figure 7.12: Data structure for BRDF resampling. The top level kd-tree contains measured incident directions $\vec{\omega}_{i,j} \in \Omega_{i,m}$, while each incident node points to the kd-subtree containing the measured exitant directions $\vec{\omega}_{r,k} \in \Omega_{r,m}(\vec{\omega}_{i,j})$. Each subtree node then contains measured BRDF samples $f_{r,m}(\vec{\omega}_{i,j},\vec{\omega}_{r,k})$.

by a function w and summed over the found incident directions to obtain $\hat{f}_{r,m}(\vec{\omega}_i,\vec{\omega}_r)$:

$$\begin{aligned}\hat{f}_r(\vec{\omega}_i,\vec{\omega}_r) \quad = \quad & \sum_{j=1}^{N}\sum_{k=1}^{N} w\left(\|\vec{\omega}_{i,j},\vec{\omega}_i\|\right) w\left(\|\vec{\omega}_{r,k},\vec{\omega}_r\|\right) f_{r,m}\left(\vec{\omega}_{i,j},\vec{\omega}_{r,k}\right), \\ & \vec{\omega}_{i,j} \in \Omega_{i,m},\ \vec{\omega}_{r,k} \in \Omega_{r,m}\left(\vec{\omega}_{i,j}\right). \end{aligned} \tag{7.3}$$

A simple cone filter was chosen as weighting function to resample the measured BRDF data in the validation simulations:

$$w(s_j) = \frac{1-\frac{s_j}{\max(s_1,\ldots,s_N)}}{\sum_{k=1}^{N} w(s_k)}, \qquad j \in [1,N]. \tag{7.4}$$

The resampling method is summarised as pseudocode in algorithm 7.1.

The distance metric for the nearest neighbour lookup must be adapted to the spherical topology. The concept of solid angle is amenable here (figure 7.13), with the distance between two vectors $\vec{\omega}_1$ and $\vec{\omega}_2$ equivalent to the solid angle ω subtended by $\vec{\omega}_1$ around $\vec{\omega}_2$, or vice versa. Since ω is proportional to the angle between the two vectors, we can use

$$\|\vec{\omega}_1,\vec{\omega}_2\| \propto \cos^{-1}\left(\vec{\omega}_1 \cdot \vec{\omega}_2\right), \tag{7.5}$$

which suffices for comparison during the search.

The bilevel paradigm is necessary in order to guarantee a balanced incident / exitant nearest neighbour lookup, i.e. with N neighbours in both dimensions. Without this constraint, BRDF lookups may be distorted by disproportionate incident and exitant solid angles, leading to excessive bias in one dimension, and noise in the other.

The BRDF resampling is however subject to a major caveat for points outside the measured range arising from the constraints of the goniophotometer; since

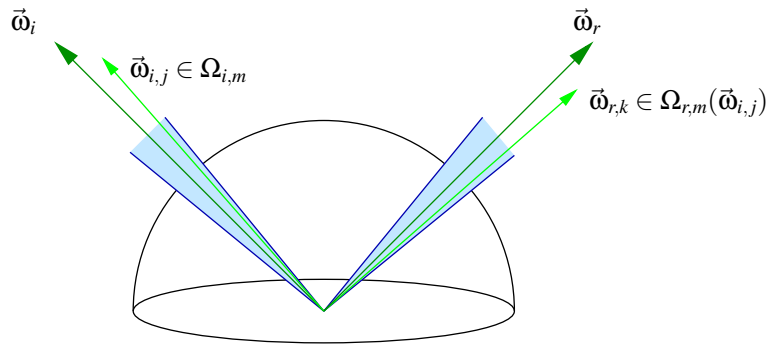


Figure 7.13: Nearest neighbour lookup over incident and exitant directions for BRDF resampling. A lookup consists of finding the N closest incident directions $\vec{\omega}_{i,j}$ around $\vec{\omega}_i$ in the set of incident samples $\Omega_{i,m}$, then finding the N closest exitant directions $\vec{\omega}_{r,k}$ around $\vec{\omega}_r$ in the set of exitant samples $\Omega_{r,m}(\vec{\omega}_{i,j})$ for each incident direction found. On a spherical topology, the search space is a solid angle.

there is no BRDF data available beyond the measurement boundary, the BRDF must be extrapolated, resulting in unknown deviations in these regions. In this case, the resampling method draws on nearest neighbours at the boundary, effectively implementing the constant extrapolation strategy used by the Voronoi integration to obtain the reflectance from the goniophotometer data[2] (section 7.3.2). Of course, this strategy still leads to underestimation of the molleton's off-specular peak, which continues to rise beyond the boundary. These deviations are not expected to be significant, however, due to the cosine term inherent in scattering during the light transport simulation.

The reflectance of the resampled molleton BRDF using 5 nearest neighbours is compared to the measurement from the integrating sphere reflectometer in figure 7.14. There is good agreement after artifact removal, with deviations within 2.5% up to an incident angle of ca. 78°, where the measurement boundary is exceeded and the resampling extrapolates over incident directions, causing the reflectance to level off. Comparing this to figure 7.10, we see the resampled reflectance agrees more favourably with the integrating sphere measurement than the reflectance upper bound obtained via Voronoi integration of the goniophotometer data. Although both methods use constant extrapolation at the boundary, the Voronoi integration computes the reflectance over the Voronoi sites corresponding to actual measurement points, whereas the resampled reflectance is computed over the vertices of a geodesic subdivision of the hemisphere [Dut90], which extend beyond the boundary. Because the Voronoi sites are all within the measurement bounds, the cosine weights assigned to boundary cells are overestimated, leading to a higher

[2] Unlike the photon map, BRDF resampling is not a density estimation problem and as such does not suffer from boundary bias resulting from a spurious decrease in density as discussed in chapter 5.

procedure BRDFresample($\vec{\omega}_i, \vec{\omega}_r, \Omega_{i,m}, N$) {Get resampled BRDF $\hat{f}_{r,m}(\vec{\omega}_i, \vec{\omega}_r)$ from incident kd-tree $\Omega_{i,m}$ containing BRDF samples via search for N nearest neighbours}
  $\hat{f}_{r,m}(\vec{\omega}_i, \vec{\omega}_r) = 0$
  S_i = BRDFnearest($\vec{\omega}_i, \Omega_{i,m}, N$) {Find nearest incident neighbours}
  for $j = 1$ to N **do** $\{\vec{\omega}_{i,j} \in S_i\}$
    $\hat{f}_{r,m}(\vec{\omega}_{i,j}, \vec{\omega}_r) = 0$
    S_r = BRDFnearest($\vec{\omega}_r, \Omega_{r,m}(\vec{\omega}_{i,j}), N$) {Find nearest exitant neighbours}
    for $k = 1$ to N **do** $\{\vec{\omega}_{r,k} \in S_r\}$
      $\hat{f}_{r,m}(\vec{\omega}_{i,j}, \vec{\omega}_r) = \hat{f}_{r,m}(\vec{\omega}_{i,j}, \vec{\omega}_r) + w(\|\vec{\omega}_{r,k}, \vec{\omega}_r\|) f_{r,m}(\vec{\omega}_{i,j}, \vec{\omega}_{r,k})$ {Sum exitant samples}
    end for
    $\hat{f}_{r,m}(\vec{\omega}_i, \vec{\omega}_r) = \hat{f}_{r,m}(\vec{\omega}_i, \vec{\omega}_r) + w(\|\vec{\omega}_{i,j}, \vec{\omega}_i\|) \hat{f}_{r,m}(\vec{\omega}_{i,j}, \vec{\omega}_r)$ {Sum incident samples}
  end for
return $\hat{f}_{r,m}(\vec{\omega}_i, \vec{\omega}_r)$

Algorithm 7.1: BRDF resampling algorithm

reflectance. The resampled reflectance, on the other hand, is obtained by also explicitly integrating over points beyond the boundary, extrapolating the BRDF in the process but using the correct cosine weights. The resampled refectance can therefore be regarded as the more accurate solution.

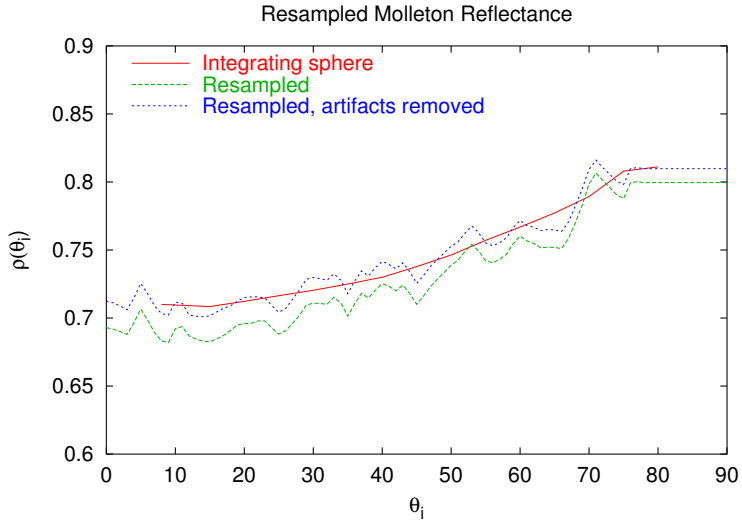


Figure 7.14: Resampled molleton reflectance using 5 nearest neighbours before and after BRDF artifact removal compared to reference obtained from integrating sphere.

7.3.4.1 Resampling Bias Test

The nearest neighbour search used in the BRDF resampling is prone to the same proximity bias investigated in chapter 5 for the photon map, and can therefore introduce an error into the resampling. The BRDFs under investigation here did not exhibit highly localised peaks, and this error was therefore not expected to be significant.

To quantify this bias, tests were conducted using the Lafortune model fitted to the molleton and aluminium BRDFs. The analytical model is particularly useful for the test because, unlike the measured BRDF data, it can be evaluated at arbitrary points. The BRDF was evaluated and stored in the kd-tree for incident and exitant direction pairs present in the measured BRDF data after artifact removal. This treats the Lafortune BRDF as measured data and ensures the sample distribution is identical to that of the actual measurement. The bias was estimated by comparing the analytical and resampled Lafortune BRDF for a number of sample points obtained from a geodesic subdivision of the hemisphere.

The resampled BRDF plots are shown for molleton and aluminium in figures E.9 and E.10, respectively. The resampled BRDF fitted to molleton exhibits deviations under the emerging specular peak beyond the measurement boundary. The constant extrapolation overestimates the falling flank of the peak at low incident angles, causing negative deviations. Conversely, it underestimates its rising flank at high incident angles, resulting in positive deviations. The ridge visible across the plot at 10° incidence results from samples in the goniophotometer arm's shadow which were deleted during artifact removal. Nearest neighbour search across the resulting gap creates a discontinuity in the rising flank of the specular component. The resampled BRDF fitted to aluminium exhibits similar deviations. As with molleton, resampling overestimates the falling flank of the peak beyond the measurement boundary, and underestimates its rise at grazing incident angles. At low incident angles there are deviations in the peak resulting from proximity bias, although these are small because the peak is relatively broad; an amenity brought about by sandblasting the aluminium. While the incident directions for the first three plots for both materials are within the measured range ($< 78^\circ$ incidence), requiring only extrapolation over the exitant directions, the last plot at 85° incidence requires extrapolation over the incident *and* exitant directions, which inevitably results in larger deviations.

Figure 7.15 shows plots of the relative mean bias in the resampled BRDFs as a function of incident angle. For molleton, the bias is within 0.5% for incident angles within the measured range ($\theta_i < 78^\circ$), where boundary values are only extrapolated over exitant directions. Beyond this limit, the resampled BRDF is extrapolated over incident directions as well, raising the bias to ca. 3% for 10 nearest neighbours. As indicated by the corresponding difference plots (figure E.10), the aluminium exhibits obvious deviations due to the falling and rising flank of the peak subject to boundary extrapolation. Here too, these deviations result in positive and negative bias which is more prominent than with molleton because the specular component dominates here, raising deviations to over 3% for 10 nearest neighbours once the peak crosses

the measurement boundary.

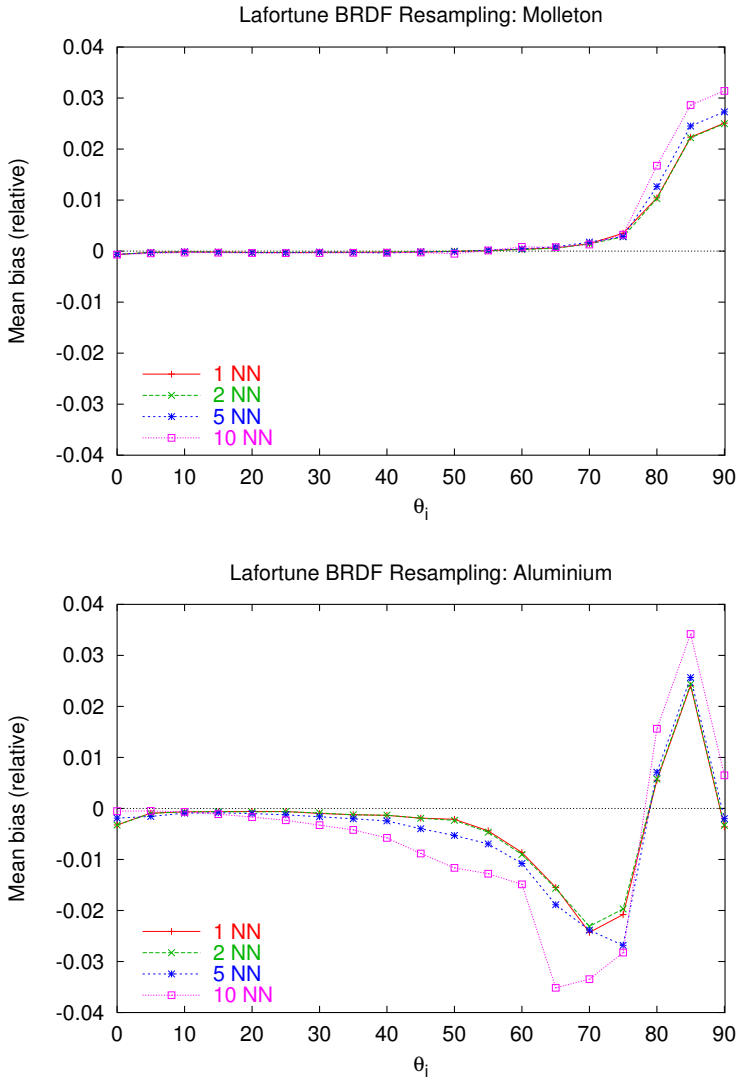


Figure 7.15: BRDF resampling bias test: mean bias of resampled Lafortune BRDF fitted to molleton (top) and aluminium (bottom) using 1, 2, 5, and 10 nearest neighbours (NN) compared to analytical evaluation of model.

A further test was conducted by comparing the analytically computed reflectance of the Lafortune BRDF to a numerical integration of resampled BRDF values. This was intended as an indication of the bias present in the scattering due to the cosine factor, which is more relevant to the validation simulations. As in the previous test, the Lafortune BRDF was resampled from sample points identically distributed to those in the measured BRDFs after artifact removal.

The results of the resampled reflectance test are shown in figure 7.16. As ex-

pected, the resampled reflectance levels off for both materials as a result of constant extrapolation, and deviations increase overall with the number of nearest neighbours. As a consequence, 5 nearest neighbours were deemed adequate for the validation simulations, for which the predicted bias in figure 7.15 is within 3% for both materials. Given this small error, resampling bias was not considered an issue and disregarded in the validation simulations.

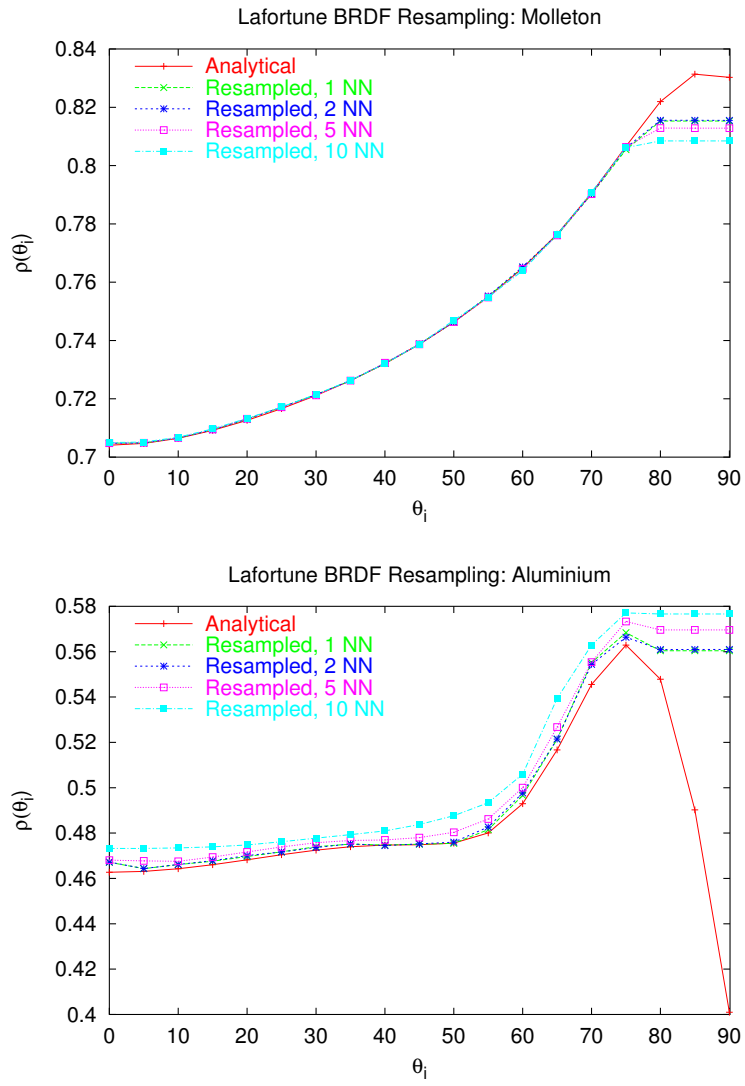


Figure 7.16: BRDF resampling bias test: analytically computed reflectance of Lafortune BRDF fitted to molleton (top) and aluminium (bottom) compared to a numerical integration of the resampled BRDF model using 1, 2, 5, and 10 nearest neighbours (NN).

7.3.5 BRDF Verification: Helmholtz Reciprocity Test

With the BRDF resampling code in place, we can further verify the goniophotometer BRDF data by subjecting it to a Helmholtz reciprocity test (equation 2.10). This test determines the relative deviations resulting from exchanging the incident and exitant directions while resampling the measured BRDFs. The deviations give an indication of the accuracy of the goniophotometer data. Due to the goniophotometer constraints, the test is limited to the measured range of directions (up to incident and exitant angles of ca. 78°). The resampling points within this range are once again obtained from a geodesic subdivision of the hemisphere.

Using 5 nearest neighbours for BRDF resampling, the average deviations are within 0.1% and 2% for the molleton and aluminium BRDFs, respectively. The substantially larger deviations for aluminium are expected due to the more localised nature of the BRDF compared to molleton, and the fact that the relative deviations are bloated by the low diffuse component outside the specular lobe. These deviations are a combined product of positioning errors on behalf of the goniophotometer, as well as resampling errors in regions with sparse samples, particularly at the boundary ($\theta_i, \theta_r \geq 78^\circ$).

Figure 7.17 is a plot of the molleton and aluminium BRDFs and their reciprocals in the plane of incidence for the incident angle typically used in the validation case studies ($\theta_i = 62^\circ$). The aluminium BRDF deviates noticeably for incident angles around 40° and in the boundary region beyond 70°. Reducing the number of nearest neighbours for the resampling introduces aliasing and only yields a marginal reduction in deviations. The deviations in the boundary region are expected, since this is where exitant samples (or incident samples for the reciprocal) are extrapolated. However, they do not come to bear in the simulations due to the cosine term applied in photon scattering.

Reciprocity deviations have consequences for the validation simulations: the resampled BRDF $\hat{f}_{r,m}(\vec{\omega}_i, \vec{\omega}_r)$ used by the forward raytracer will deviate from the reciprocal BRDF $\hat{f}_{r,m}(\vec{\omega}_r, \vec{\omega}_i)$ used by the backward raytracer[3], yielding deviations in illuminance. This was particularly noticeable in the caustic from the aluminium light shelf.

A simple countermeasure was attempted by averaging the resampled BRDF and its reciprocal in the simulations. However, this tended to yield complementary results: while deviations with forward raytracing were reduced, those with backward raytracing were often increased, or vice versa. Instead, a selective approach was adopted whereby an additional simulation using the reciprocal BRDF was performed. This reciprocal solution was compared to the nonreciprocal, and the more accurate of the two chosen for evaluation. This was only necessary for the component case study involving the light shelf.

[3] Here we assume $\vec{\omega}_i$ and $\vec{\omega}_r$ are in local coordinates and both point away from the surface.

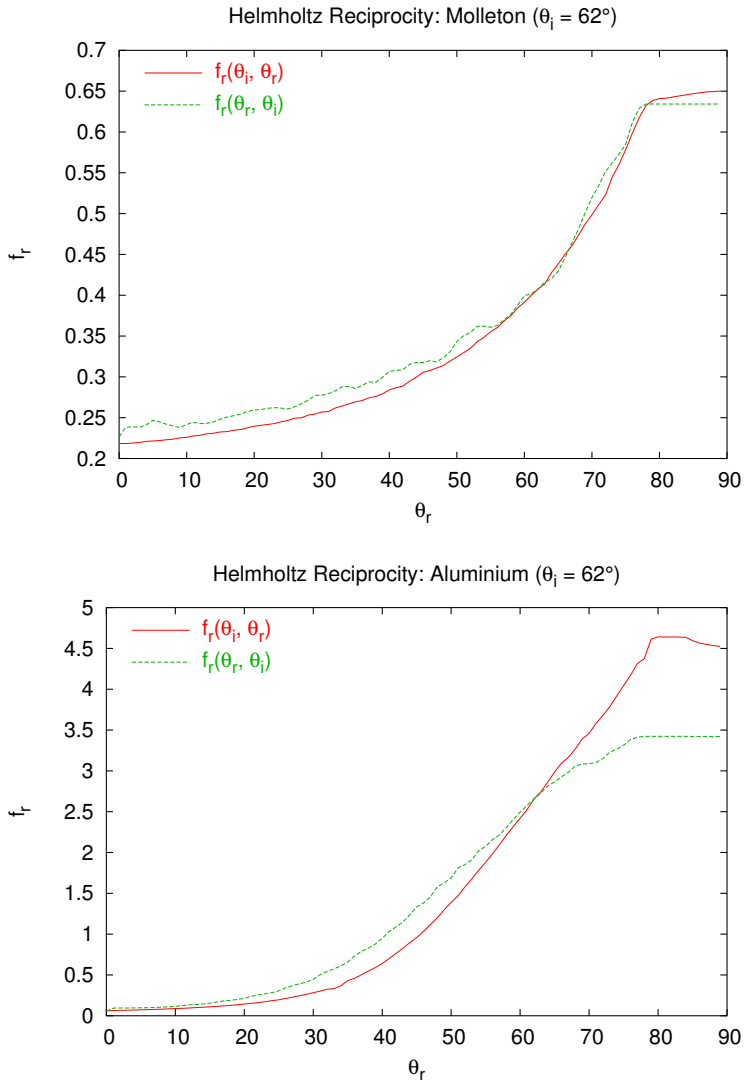


Figure 7.17: Helmholtz reciprocity within plane of incidence for molleton (top) and aluminium (bottom) BRDFs.

7.3.6 Simulation

Simulation runs using RADIANCE Classic and photon map were conducted for each case study to obtain the illuminance at the measured sensor positions.

The photon map simulations used global photons only (recall that this includes caustics), which were visualised directly rather than via final gathering, as in the analytical validation. Circa 1000000 photons were used, with a relatively large bandwidth of 5000 photons in order to suppress noise. Boundary bias in the density estimates was not expected to be significant, particularly for the middle sensors, neither was proximity bias, since the sandblasted aluminium light shelf does not

produce a highly localised caustic.

While the photon map parametrisation affected primarily the noise in the simulation, the parameters for RADIANCE Classic were critical as they can give rise to genuine bias if choosen too conservatively. This tendency was already evident in the analytical validation in the form of systematic underprediction resulting from premature termination of rays sampling the indirect illumination. Consequently, the number of ambient and specular reflections (-ab, -lr) for case studies involving interreflection was set to 40, while 2 reflections were sufficient for case studies involving a single reflection (the additional reflection accounts for background illuminance). As in the analytical validation, ray weights were not thresholded (-lw 0). Together, these settings ensure that RADIANCE Classic produces lighting levels comparable to those obtained with the photon map. Noise was expected in case studies involving caustics, which was adequately suppressed with ca. 8000 ambient samples for interreflection case studies, and 16000 for single reflection case studies. The irradiance cache was also parametrised for high accuracy (-aa 0.05, -ar 512).

The validation simulations require two principal components obtained by measurement: the BRDFs of the molleton lining the box interior and the aluminium light shelf, and the lamp's EDF. These components are discussed in the following sections.

7.3.6.1 BRDF Simulation

Incorporating the measured BRDF into the simulation requires sampling an arbitrary signal lacking an analytical representation. Monte Carlo inversion of arbitrary BRDFs with a numeric technique similar to that described in section 4.1.1 for photon emission would be prohibitively expensive in this case, since the construction for the CDF lookup table would require numerous nearest neighbour lookups in the kd-tree, and this process must be performed for every reflection. This is where the combined utility of the fitted Lafortune BRDF and the resampled BRDF data comes in; we can sample the former and compensate by scaling each sample with a factor proportional to the latter. We combine this with the scaling factor for floating point exponents in equation B.12 since this is required in order to sample the fitted model. This effectively gives us a bias-free estimator for the measured BRDF as if we had actually sampled it, except that we sample the Lafortune model analytically instead, which is far more efficient. We do this as in appendix B.2 by selecting a lobe j of the Lafortune BRDF based on its reflectance $\widetilde{\rho}_{l,j}$ and then sampling it with compensating factors to obtain the estimator:

$$
\begin{aligned}
\widehat{L}(\vec{\omega}_i) \quad = \quad & \overbrace{\frac{\hat{f}_{r,m}(\vec{\omega}_i,\vec{\omega}_r)}{f_{r,l}(\vec{\omega}_i,\vec{\omega}_r)}}^{\text{Measured BRDF compensation}} \cdot \overbrace{\frac{\widetilde{\rho}_l(\vec{\omega}_i)}{\widetilde{\rho}_{l,j}(\vec{\omega}_i)}}^{\text{Lobe compensation}} \cdot \overbrace{\frac{\widetilde{I}_{l,j}(\vec{\omega}_i)}{\widetilde{f}_{r,l,j}(\vec{\omega}_i,\vec{\omega}_r)}}^{\text{Direction compensation}} \cdot \\
& f_{r,l,j}(\vec{\omega}_i,\vec{\omega}_r)\, L(\vec{\omega}_r) \cos\theta_r, \\
\vec{\omega}_r \quad \in \quad & \Omega,
\end{aligned}
$$

$$j \in [1,N], \tag{7.6}$$

where $\hat{f}_{r,m}$ is the measured BRDF obtained by resampling (equation 7.3), and the remaining terms are defined as in appendix B.2. The respective compensating factors are identified in overbraces.

When using a photon map with constant flux (section 3.1.1.2), $\widetilde{\rho}_l$ is already accounted for by russian roulette, while the compensating factors modify the photon flux. This clearly violates the constant flux convention, but does not introduce bias because these are correction factors for nonuniform sampling. However, deviations in the fitted and measured BRDFs (where $\hat{f}_{r,m}/f_{r,l} \neq 1$) will introduce variance in the photon flux and consequently in the density estimates.

7.3.6.2 Lamp EDF Simulation

The ideal lamp for a validation would be a point source with uniform EDF. Reality complicates things; real lamps have finite area and can sometimes exhibit highly localised EDFs. Theoretically, this would require placing the lamp as far as possible from the box in order to approximate a point source. In practice, there are of course limitations, since this dramatically reduces the illuminance at the box window and compromises the accuracy of the sensor measurement made in the box interior. On the other hand, closer placement of the lamp reduces the uniformity of the distribution, revealing two lobes from the lamp reflector at distances under 3 m (figure 7.18). A compromise was found with a distance of 4.6 m, at which the two lobes have fully coalesced.

A significant challenge in the validation is obtaining the lamp's EDF required for the validation simulations. While this can be obtained from the lighting manufacturer as goniometric measurements based on the design data, the resolution is often too low (considering that in our setup the box only subtends a solid angle of ca. 0.01 sr at the light source). Furthermore, this data does not account for manufacturing tolerances, which we found to be quite significant with our lamp, particularly on behalf of the reflector. Custom high resolution goniometric measurements, on the other hand, are very costly.

A simple approach was devised to extract the lamp's EDF without actually performing any goniometric measurements. We assume our lamp is a point source as in far-field photometry, since its distance/size ratio easily satisfies the "five-times-rule" [Ash95]. Our EDF extraction setup is shown in figure 7.19. A diffuse 60 $\times$ 60 cm panel was placed in front of the box. Next, an image of the distribution pattern on the panel was taken with the HDR camera mounted on a low tripod, pointing upwards to avoid casting a shadow. The aluminium camera body and tripod were draped with black cloth to eliminate any specular reflection which may affect the illuminance on the panel. The relative positions and orientation of the lamp, camera, and panel were then measured and recorded, yielding the parameters shown in figure 7.19. These include the panel's inclination θ_p relative to the camera (defining the panel's normal $\vec{N}_p = [N_{p,x}, N_{p,y}, N_{p,z}]$ in terms of the camera's

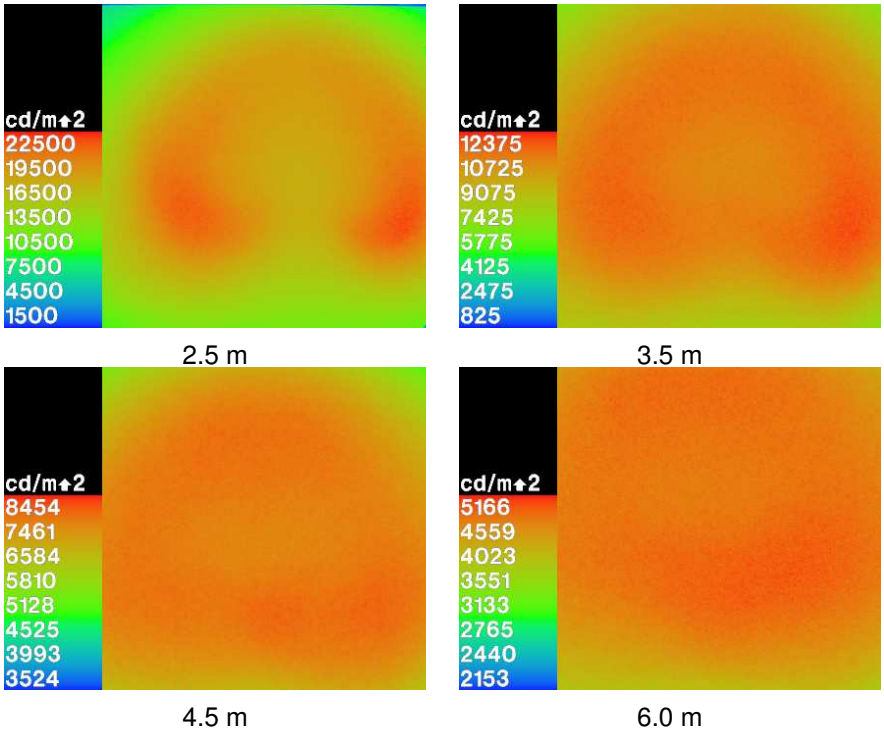


Figure 7.18: Lamp EDF at distances of 2.5–6.0 m. A distance of ca. 4.6 m was used for the validation, since the distribution is dominated by two lobes at closer placement, and the illuminance is too low at greater distances.

coordinate system), the focal length of the camera lens d_c (from the camera's specifications), and the distance d_p between the camera's centre of projection and the panel.

A ray $\vec{R}$ emitted from the light source will strike the panel at a point $\vec{p}_p$ which is projected onto a point $\vec{p}_c = [p_{c,x}, p_{c,y}, d_c]$ in the camera's projection plane (and thus in the HDR image), hence $\vec{p}_p = t\vec{p}_c$ for some factor t. By solving for t, we reverse the projection and obtain $\vec{p}_p$ in terms of the panel's coordinate system defined by $\vec{N}_p$[4]. Given two points that lie on the panel, $\vec{p}_p$ and the panel's midpoint $\vec{p}_m = [0, 0, d_p]$, we use the plane equation to obtain

$$(\vec{p}_p - \vec{p}_m)\vec{N}_p = 0 \tag{7.7}$$

$$t\vec{p}_c\vec{N}_p = d_p N_{p,z} \tag{7.8}$$

$$t = \frac{d_p N_{p,z}}{\vec{p}_c\vec{N}_p}. \tag{7.9}$$

Having solved for t, the HDR image pixels can be tabulated as EDF samples using $[p_{p,x}, p_{p,y}]$ as indices into a 2D light field slice [LH96] positioned at the panel.

[4]Note that we project in terms of $\vec{N}_p$ rather than θ_p to distinguish between upward and downward inclination of the panel.

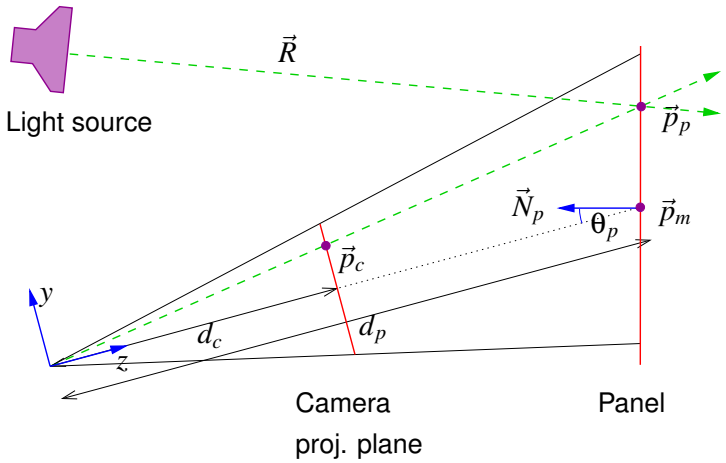


Figure 7.19: Extraction of lamp EDF. The HDR image of the panel taken with the camera is projected back onto the panel subject to the angle θ_p between the camera's view direction and the panel's normal.

The positioning of the panel on the box is crucial, with four markers on the panel serving as reference points for known coordinates on the box face. The extraction process automatically locates these markers in the HDR image and determines the offset and scaling required to align the extracted distribution with the box geometry.

To render the acquired EDF, the light field data is converted to a RADIANCE *brightdata* file, which modifies the light source's luminance [LS98]. The light field is then accessed by computing the intersection $\vec{p}_p$ of a source ray $\vec{R}$ with the light field plane (see figure 7.19), the latter essentially acting as an impostor for the panel used during acquisition. $\vec{p}_p$ is computed with an auxiliary user defined function, and the light field data is indexed at $[p_{p,x}, p_{p,y}]$ as in the acquisition. This results in a projection of the EDF onto the box and its interior. Note that the source ray's direction is irrelevant; it may point towards or away from the source, thus catering to backward resp. forward raytracing.

Simple as this EDF acquisition scheme is, it is severely limited. Only a small region of the EDF required for the validation is extracted, and it is constrained to the geometry used during acquisition; altering the position or orientation of the light source during rendering will invalidate the EDF data. Furthermore, the extracted EDF is only a 2D function of the coordinates on the panel, rather than the complete 4D function, since there is no information available on the rays' origins during acquisition (obtaining this data would require at least two planes bisecting the EDF [LH96]). This method is therefore unsuitable for near-field photometry.

Far more elaborate and comprehensive methods for EDF extraction exist, such as that proposed by Goesele et al [GGHS03]. The setup described here is similar to Goesele's using a diffuse reflector. However, no filtering or projection onto basis functions was done, but rather the raw luminance data obtained from the HDR camera was used.

Figure 7.20 shows an example of an extracted EDF. This distribution is rendered with RADIANCE Classic's direct component as well as with the photon map by visualising the direct photon hits on the panel, which provides a check for the accuracy of the photon emission. The renderings as well as the HDR image are normalised for comparison. This normalisation is not a conventional mapping of the maximum intensity to 1, but rather a mapping of the average intensity to 0.5[5]. The reason for this is that the maximum intensity is prone to noise in the photon map renderings and therefore falsifies the error metric, since we are interested in the mean bias, not the noise. Consequently, mapping the average to 0.5 results in a more robust normalisation which enables reliable quantitative comparison.

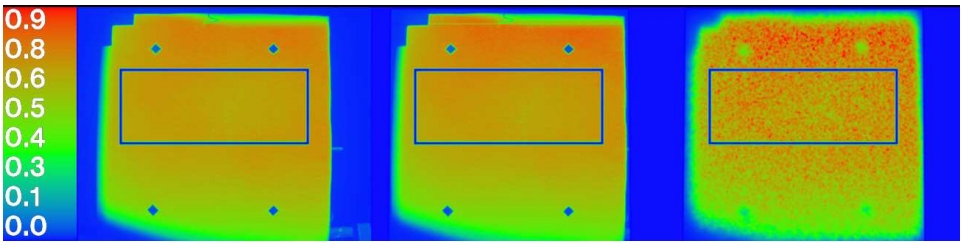


Figure 7.20: Falsecolour images of lamp EDF. Left: HDR camera image of panel in front of box. Centre: extracted distribution rendered with RADIANCE Classic's direct component. Right: extracted distribution rendered with direct photons. The superimposed boundary marks the location of the box window. The four markers are used to align the distribution with the box geometry.

RADIANCE Classic's direct calculation approximates the HDR image very well, with a mean deviation of ca. 1% within the window area. This accuracy is also achievable with the photon map; the photon map rendering used 1000000 photons with a bandwidth of 100 photons. This does require extreme parameters in order to resolve the distribution, however, with 100000 samples/sr for the emission PDF and over 2000 partitions on the light source surface, which impacts the time spent in the forward pass substantially. By contrast, using a uniform distribution (figure 7.21) produced deviations of up to 14%, which would particularly affect the interior illuminance in case studies involving caustics.

There are of course factors in the extraction scheme which inevitably introduce some error. The texture of the panel's surface produces minor artifacts in the HDR image. Furthermore, the periphery of the HDR image is slightly distorted by the camera's wide angle lens. Lastly, the relative position and orientation of the lamp, camera, and panel can only be measured with limited accuracy. Measured distances and angles are estimated to be accurate within $\pm$10mm and $\pm 0.5^\circ$, respectively. Given the parameters for our setup in equation 7.9, this results in a projection error of roughly $\pm$1%, which is deemed negligible.

[5]Obviously this does not necessarily map the maximum to 1, but this is inconsequential for the purpose of comparison.

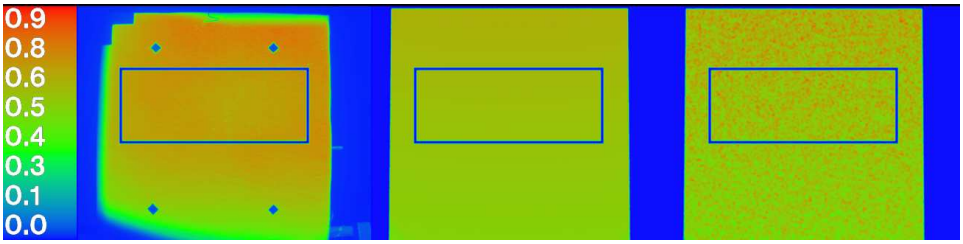


Figure 7.21: Falsecolour images of panel rendered with RADIANCE Classic (centre) and photon map (right) using a uniform EDF. The HDR camera image of the actual source distribution is shown at left.

7.3.7 Error Analysis

Figure 7.22 shows a breakdown of the errors conceivably involved in the validation, categorised according to their origin. The total error in the simulation is composed of a number of deviations arising from the measurement or simulation side. Ideally, we would like to subtract each error from the total in order to find the one error we are really interested in: the light transport simulation error. Practically, this is impossible, because the majority of the errors cannot be exactly determined. We can, however, roughly estimate each component error's contribution and include it in the data on either the physical or simulation side (depending on where the error arises) in order to obtain error margins.

The major contributors to deviations identified in figure7.22 are discussed in the following subsections.

7.3.7.1 Illuminance Sensors

These are deviations inherent in the illuminance sensors, which consist of an absolute error δ_{sens} of 5 lx and 50 lx for the low (0–10000 lx) and high (10000–100000 lx) ranges, respectively, as well as a relative error ε_{sens} of 4%. The sensor error becomes significant with low illuminance, particularly when measuring the background illuminance necessary for analytical estimates. Error estimates accounting for illuminance sensor tolerance are therefore included in the measurements and in the analytical solutions which include the background illuminance.

7.3.7.2 Spectral Distribution

The spectral distributions of the validation components are plotted in figure 7.23. These include the spectra of the light source emission, molleton reflectance (obtained from the integrating sphere reflectometer), and the illuminance sensor response, the latter closely approximating the photopic response function $V(\lambda)$. The validation simulations, on the other hand, are monochromatic and use only the photopic values obtained from the measurements. Potential deviations incurred by this simplification must be investigated and accounted for if necessary.

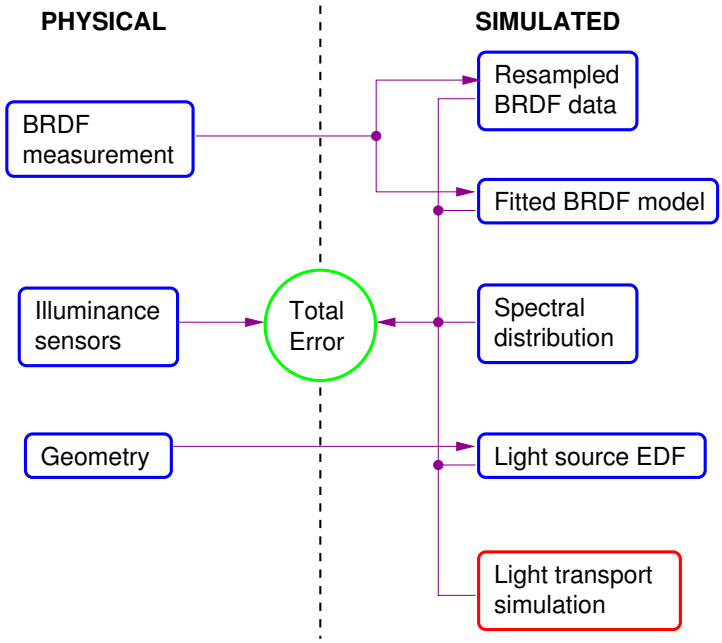


Figure 7.22: Principal components contributing to validation error. The light transport simulation error is sought, and can theoretically be obtained by subtracting the error contributions from the total. Since this requires the exact deviations introduced by each component, this is practically impossible. Instead, error estimates can be derived for the measurements and simulations.

The lamp spectrum exhibits peaks in the green and orange bands at ca. 540 and 590 nm, respectively. This is not critical, however, as long as the lamp's spectrum is not modified significantly by reflection off the molleton within the sensor's spectral response range, resulting in a spectral shift. While the molleton is not spectrally neutral, its reflectance exhibits a gentle gradient within the $V(\lambda)$ peak in figure 7.23. Consequently, it is not expected to have a significant effect.

To assess the extent of the spectral shift, a spectral error estimate can be derived by convolving the lamp emission, molleton reflectance, and sensor response spectra, and comparing this integral with its monochromatic counterpart used in the simulation. This estimate accounts for the indirect illuminance within a furnace type scene as used for the analytical validation in chapter 6. We therefore neglect geometric factors in the light transport. Since our validation box has an opening and only one side receives direct illumination, this estimate is considered an upper bound.

The upper bound for the measured indirect illuminance arriving at an interior sensor relative to the direct illuminance measured at the exterior sensors can be

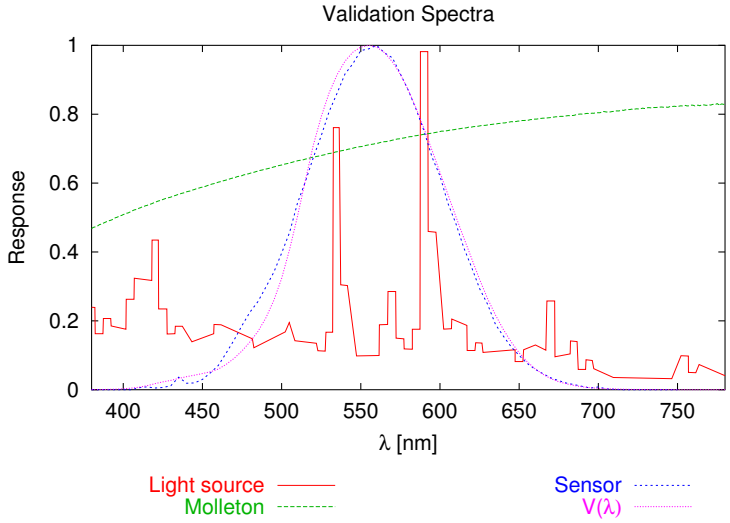


Figure 7.23: Spectral distribution of light source emission, molleton reflectance, and illuminance sensor response. $V(\lambda)$ is superimposed as reference.

approximated by

$$E_m \leq \frac{\int E_l(\lambda)\left[\sum_{i=1}^{n}\rho^i(\lambda)\right]V(\lambda)\,d\lambda}{\int E_l(\lambda)\,V(\lambda)\,d\lambda}, \tag{7.10}$$

where ρ is the spectral reflectance of the molleton as measured in the integrating sphere for normal incidence (where the material exhibits diffuse behaviour), E_l is the lamp's spectral illuminance, and n is the number of reflections. In the numerator, the lamp's emission spectrum is modified by one or more reflections off the molleton, and finally weighted by $V(\lambda)$ at the interior sensor. In the denominator, the lamp's emission spectrum is not modified, but rather weighted by $V(\lambda)$ directly at the exterior sensors.

For component case studies involving a single reflection $n = 1$, whereas for compound case studies involving interreflection $n = \infty$, giving rise to a geometric series. The reasoning behind this approximation is analogous to that of the spherical case study in chapter 6, namely the assumption that each point in the box receives the cumulative indirect illuminance from 1 reflection, 2 reflections, 3 reflections, etc.

On the simulation side, we have

$$E_s \leq \frac{\left[\int E_l(\lambda)\,V(\lambda)\,d\lambda\right]\sum_{i=1}^{n}\left[\int\rho(\lambda)\,V(\lambda)\,d\lambda\right]^i}{\int E_l(\lambda)\,V(\lambda)\,d\lambda} = \sum_{i=1}^{n}\rho_V^i, \tag{7.11}$$

where ρ_V is the photopic ($V(\lambda)$ weighted) molleton reflectance. Here, any spectral shift effected by the molleton on the lamp's emission spectrum is not taken into account, giving rise to a potential error.

Numeric evaluation and comparison of equations 7.10 and 7.11 for single reflection and interreflection predicted a relative error ε_{spec} of up to 0.5% and 4%

below the measured illuminance, respectively. This error estimate is included in the simulated illuminance data for validation case studies.

7.3.7.3 Light source EDF

These are deviations in the direct component of the simulation. They have been discussed in section 7.3.6.2 and were estimated at ca. 1–2% for both RADIANCE Classic and the photon map. As mentioned above, temporal fluctuations in the lamp's emission are eliminated by obtaining illuminance values from the interior box sensors relative to the direct illuminance (the spatial distribution of the EDF is not subject to fluctuation). The actual effect of the distribution error on the indirect illuminance is conceivably small, and this error is deemed negligible. The simulation parameters for figure 7.20 were used in the case study simulations to minimise this error.

Furthermore, the EDF is also subject to deviations in measurements of the acquisition geometry (relative position and orientation of the HDR camera, light source, and panel). The measured parameters have an estimated accuracy of $\pm$10mm and $\pm 0.5^\circ$, resulting in minor deviations during EDF extraction which may also be disregarded.

7.3.7.4 Fitted BRDF Model

These are deviations of the BRDF model fitted to the measured BRDF. They are subject to the model's ability to account for the measured BRDF's characteristics. The fit is also subject to deviations and limitations incurred by the goniophotometer during BRDF measurement, such as the inability to measure grazing angles. Because of these deviations, we sample the measured BRDF using the fitted BRDF as PDF as outlined in section 7.3.6.1. Deviations will then merely increase the noise level in the simulation, but not introduce bias.

7.3.7.5 Resampled BRDF Data

Errors incurred by the goniophotometer can still directly affect the measured BRDF data which is sampled in the simulation via the fitted BRDF. Primary contributors to this error are the goniophotometer shadowing artifacts and the lack of grazing angles in the data. The shadowing artifacts have been removed as outlined in section 7.3.1.2, while the missing grazing angles are "filled in" via constant extrapolation beyond the measured range by the resampling method used to evaluate the measured BRDF at arbitrary points. Resampling can also introduce some bias, but as discussed in section 7.3.4.1, this error is considered negligible and need not be accounted for.

Reciprocity deviations in the resampled BRDF data can also lead to noticeable deviations between forward and backward raytracing. As mentioned in section 7.3.5, these deviations depend on the accuracy of the goniophotometer and the

density of the samples used in BRDF resampling. The aluminium BRDF exhibits considerable reciprocity deviations and is treated accordingly by comparing simulations using the original and reciprocal resampled BRDFs, then choosing the more accurate of the two.

7.3.7.6 Light Transport Simulation

This is the error we seek. Systematic residual deviations after accounting for the other error components would indicate an error in the global illumination algorithm.

7.3.7.7 Error Bounds

To account for the potentially significant errors in our validation (i.e. spectral shift and sensor tolerance), we represent the measured and simulated illuminance as lower and upper bounds indicating possible error margins. These bounds are represented as bars in the illuminance plots for each case study (figures 7.24, 7.26, 7.29, 7.32, and 7.35). As discussed in section 7.2.3, they are expressed relative to the direct illuminance. Measured and simulated illuminance bounds are derived separately as follows:

Measurement: bounds account for sensor error. For a given measured illuminance E_m, the corresponding lower and upper bounds $E_{m,lo}$ and $E_{m,hi}$ relative to the averaged direct illuminance $E_{m,l}$ measured by the front sensors are defined as:

$$E_{m,lo} = \frac{E_m(1-\varepsilon_{sens})-\delta_{sens}}{E_{m,l}(1+\varepsilon_{sens})+\delta_{sens}}, \tag{7.12}$$

$$E_{m,hi} = \frac{E_m(1+\varepsilon_{sens})+\delta_{sens}}{E_{m,l}(1-\varepsilon_{sens})-\delta_{sens}}, \tag{7.13}$$

where ε_{sens} and δ_{sens} are the relative and absolute sensor errors for the range containing E_m as defined in section 7.3.7.1. Note that we also account for the sensor error in $E_{m,l}$.

Simulation: bounds are only derived for case studies involving interreflection using molleton (figures 7.32 and 7.35), where they account for potential underprediction due to spectral shift in the simulation. For a given simulated illuminance E_s, the lower and upper bounds $E_{s,lo}$ and $E_{s,hi}$ relative to the simulated direct illuminance $E_{s,l}$ are defined as:

$$E_{s,lo} = \frac{E_s}{E_{s,l}}, \quad E_{s,hi} = \frac{E_s(1+\varepsilon_{spec})}{E_{s,l}}, \tag{7.14}$$

where ε_{spec} is the estimated relative error due to spectral shift derived in section 7.3.7.2.

In order to ultimately assess the accuracy of the simulation, we evaluate the relative error between the measured and simulated illuminance, carrying the associated bounds over into the evaluation. This results in relative error margins ε_{lo} and ε_{hi} defined as:

$$\varepsilon_{lo} = \min\left(\varepsilon_{s,lo}, \varepsilon_{s,hi}\right), \quad \varepsilon_{hi} = \max\left(\varepsilon_{s,lo}, \varepsilon_{s,hi}\right), \tag{7.15}$$

where $\varepsilon_{s,lo}$ and $\varepsilon_{s,hi}$ are the error extremes defined by the closest and furthest pairs of measured and simulated illuminance bounds.

$$\varepsilon_{s,lo} = \frac{E_{s,lo}}{E_{m,hi}} - 1, \quad \varepsilon_{s,hi} = \frac{E_{s,hi}}{E_{m,lo}} - 1. \tag{7.16}$$

These bounds are indicated by bars in the relative error plots in figures 7.27, 7.30, 7.33, and 7.36.

7.4 Validation Case Studies

The measurements are divided into component and compound case studies. Component case studies comprise preliminary measurements to validate individual light transport modes in order to maintain tractability, and, where possible, to compare against an analytical solution to check the validity of the measurements and the validation methodology as a whole. Compound case studies combine the component case studies, resulting in more complex light transport which cannot be validated with confidence if the individual components themselves have not undergone validation.

Illuminance data from the interior sensors was measured for window distances of ca. 20–100 cm in 2 cm increments. For all twelve sensors this resulted in a data set comprising some 490 sample points per case study. In all case studies, the box has an effective inclination of ca. 30° to the lamp, and only the floor (and lightshelf, if applicable) receives direct illuminance. Since we are primarily interested in the indirect illuminance, we restrict the results to the ceiling sensors in this chapter. Results for the complete compound case study data set can be found in appendix F.

7.4.1 Component Case Study 1: Absorbing Box

7.4.1.1 Description

Illuminance is measured in the “absorbing” box (interior clad in black molleton). The purpose of the case study is to determine the background illuminance incurred by the black molleton fabric, which must be included in analytic solutions derived in the subsequent component case studies.

7.4.1.2 Results

The low illuminance results in high error bounds for the measurements due to the sensor tolerance, indicated by the large error bars in figure 7.24. This translates to an average illuminance range of 14–25 lx. This error is taken into account in analytical solutions of subsequent component case studies.

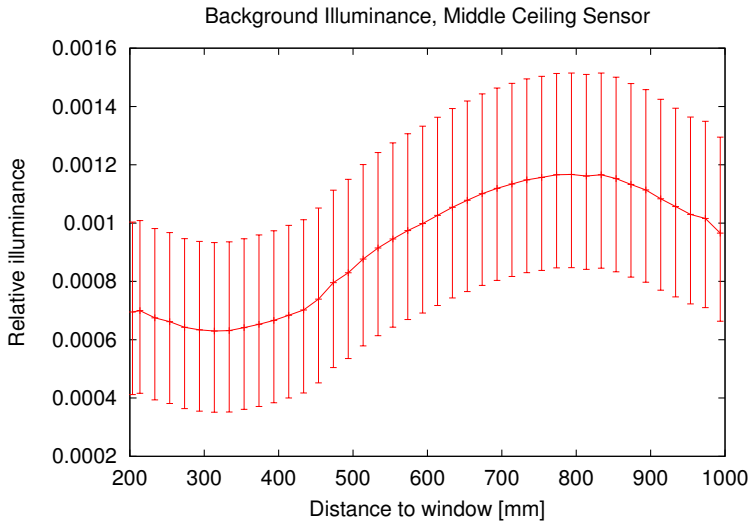


Figure 7.24: Results of component case study 1: measured background illuminance for centre ceiling sensor.

7.4.2 Component Case Study 2: Diffuse Patch Reflection

7.4.2.1 Description

Two identical light gray molleton patches are placed at equal window distance between the floor sensor tracks in the absorbing box, and the resulting illuminance measured at the ceiling (figure 7.25). The patches are placed in an area of direct illuminance, while the ceiling only receives indirect illuminance from the patches. The purpose is to derive an analytical solution based on diffuse reflection for the ceiling illuminance, which can then serve as a reference for both the measurements and the simulation. The legitimacy for considering the light transport as diffuse lies in the fact that the sensor to patch angle does not exceed 45°, such that the molleton's off-specularity does not come to bear, while the off-specular component reflected at grazing angles towards the back of the box is absorbed by the black molleton and does not affect the patches or sensors.

The analytical estimate for a ceiling sensor's illuminance E_s is based on simple form factor calculation [CW93]:

$$E_s = F_{1s}B_1 + F_{2s}B_2 + E_b \tag{7.17}$$

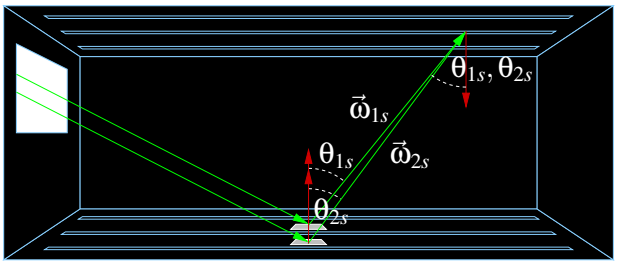


Figure 7.25: Component case study 2: diffuse patch reflection

$$= \rho\left(F_{1s}E_1 + F_{2s}E_2\right) + E_b \tag{7.18}$$

$$= \rho\frac{A}{\pi}\left(E_1\frac{\cos^2\theta_{1s}}{\left|\vec{\omega}_{1s}\right|^2} + E_2\frac{\cos^2\theta_{2s}}{\left|\vec{\omega}_{2s}\right|^2}\right) + E_b, \tag{7.19}$$

where:

- F_{1s} and F_{2s} are the form factors for sensor-patch light transport
- B_1 and B_2 are the patch luminosities
- ρ is the diffuse patch reflectance (measured at normal incidence)
- E_1 and E_2 are the patch illuminances from the light source
- E_b is the background illuminance at the sensor position
- A is the area of each patch
- $\vec{\omega}_{1s}$ and $\vec{\omega}_{2s}$ are the sensor-patch vectors
- θ_{1s} and θ_{2s} are the angles between the patch normals and $\vec{\omega}_{1s}$ and $\vec{\omega}_{2s}$, respectively.[6]

The background illuminance E_b is obtained for each ceiling sensor position from the previous case study. The direct patch illuminances E_1 and E_2 are measured with the three floor sensors at the patches' window distance, then interpolated for each patch's centre from the flanking sensor illuminances. The errors inherent in the background and floor sensor illuminances are taken into account by deriving lower and upper bounds for the analytical solution.

Obviously, the patch area A plays a crucial role and involves a tradeoff. The smaller the area, the more accurate the analytical solution, since the patch should ideally be of infinitesimal area. Conversely, a larger area increases the ceiling illuminance, and thus the accuracy of the measurements. A compromise was found with 9 $\times$ 9 cm patches, which yielded a maximum ceiling illuminance of ca. 120 lx.

[6] Because the sensor and patch normals are antiparallel, the same angles result between the sensor normals and $\vec{\omega}_{1s}$ and $\vec{\omega}_{2s}$, hence the squared cosine terms in equation 7.19.

A further component case study was planned involving patch interreflection. This called for an additional pair of patches attached to the ceiling in identical configuration to the floor patches. An analytical solution involving four patches would have been complex but feasible. Due to the small size of the patches, however, the interreflection was negligible. An analytical estimate predicted an illuminance increase below 1 lx at the ceiling sensors compared to simple patch reflection. This was also confirmed with measurements. Because the illuminance increase cannot be reliably measured with the sensors (1 lx increase versus $\pm$5 lx absolute error margin), the patch interreflection case study was aborted.

7.4.2.2 Results

Figure 7.26 compares the measured relative illuminance with the simulated values obtained with RADIANCE Classic and the photon map, as well as the analytical solution. The error bars in the measured data indicate the sensor tolerance, while the error bars in the analytical data indicate the errors inherent in the background and floor sensor illuminance. The estimated 0.5% underprediction inherent in the simulation due to spectral shift was negligible and omitted for clarity.

As expected, the patch reflection exhibits a characteristic gaussian curve centrered at the patch position. Both the measured and simulated gaussians are slightly asymmetric compared to the analytical solution, with a gentler slope towards the back of the box. This is attributed to the emerging off-specular component of the molleton patches.

The relative deviations (figure 7.27) lie within the error margin accounting primarily for the illuminance sensor tolerance, which becomes significant at such low levels. Average deviations for RADIANCE Classic and photon map are 3% and 2%, respectively.

7.4.3 Component Case Study 3: Light Shelf Caustic

7.4.3.1 Description

The aluminium light shelf is mounted on the outer window ledge of the absorbing box in order to create a caustic directed towards the ceiling (figure 7.28).

7.4.3.2 Results

The ceiling illuminance due to the caustic drops with increasing window distance, as expected (figure 7.29). The sandblasted aluminium gives rise to a broad caustic which can still be accurately simulated with RADIANCE Classic, though with a slightly higher degree of noise compared to the photon map. Both algorithms agree well with the measurement.

The photon map exhibits an illuminance drop up to 400mm. This is an obvious manifestation of the aluminium BRDF's reciprocity deviations revealed after subjecting it to the Helmholtz reciprocity test (section 7.3.5). Window distances of

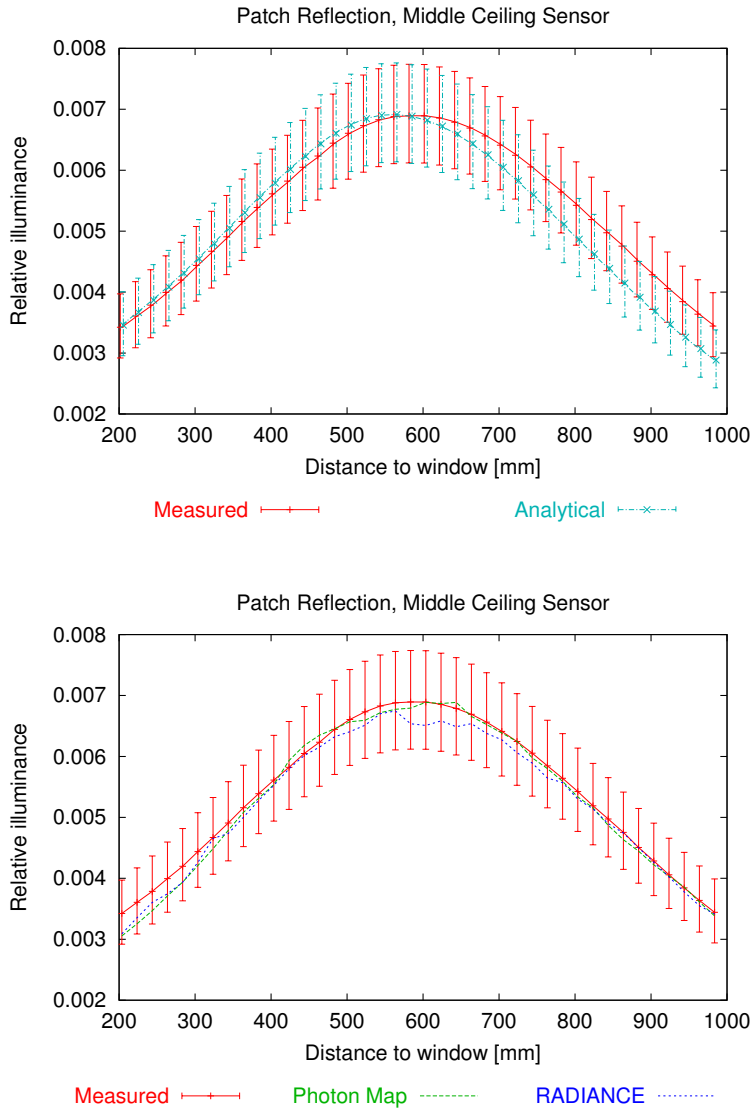


Figure 7.26: Results of component case study 2: analytical (top) and simulated (bottom) ceiling sensor illuminance from patch reflection.

200–400mm correspond to exitant angles on the light shelf of ca. 39°–58°, which lie in the region of deviation as shown in figure 7.17.

Results for a second set of simulations using the reciprocal aluminium BRDF confirm that the illuminance deviations up to 400mm are indeed caused by the BRDF: in this case, RADIANCE Classic exhibits the deviations, but not the photon map. On account of this, the original photon map simulation was dropped in favour of the reciprocal simulation, and used for subsequent error analysis.

Relative deviations for the ceiling sensor are graphed in figure 7.30. RADIANCE

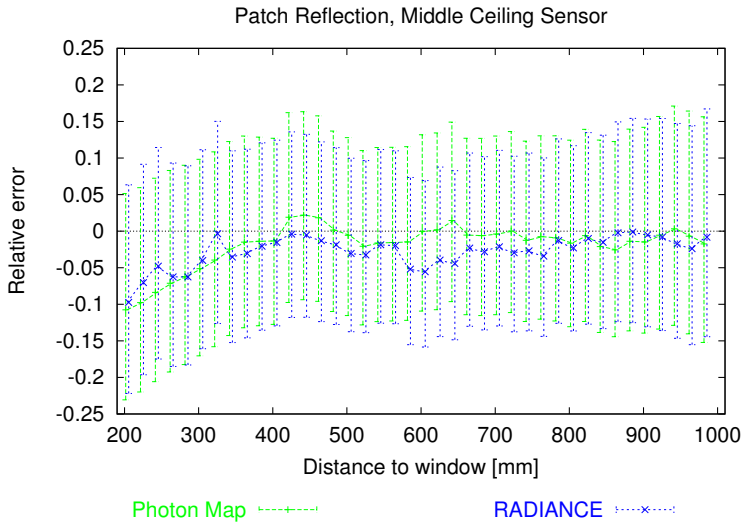


Figure 7.27: Results of component case study 2: relative error in simulated illuminance from patch reflection. Error bars are based on sensor tolerance in the measurement. Bars associated with identical points are separated for clarity.

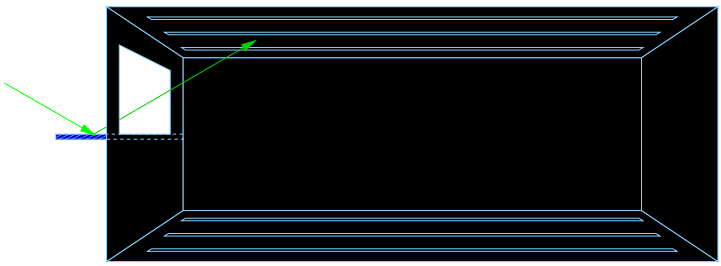


Figure 7.28: Component case study 3: light shelf caustic

Classic's deviations are mostly noise, which increases towards the back of the box since the density of ambient rays striking the light shelf drops. The absolute component of the illuminance sensor tolerance comes to bear in the error margin as the caustic fades with increasing window distance. Average deviations for RADIANCE Classic and photon map are 7% and 2%, respectively.

7.4.4 Compound Case Study 1: Diffuse Interreflection

7.4.4.1 Description

The box interior is clad with light gray molleton (figure 7.31). This is a generalisation of component case study 2, and no analytical solution is available. The error incurred by the light transport simulation is expected to weigh in far greater due to the interreflection. Furthermore, an increase in illuminance is expected towards the

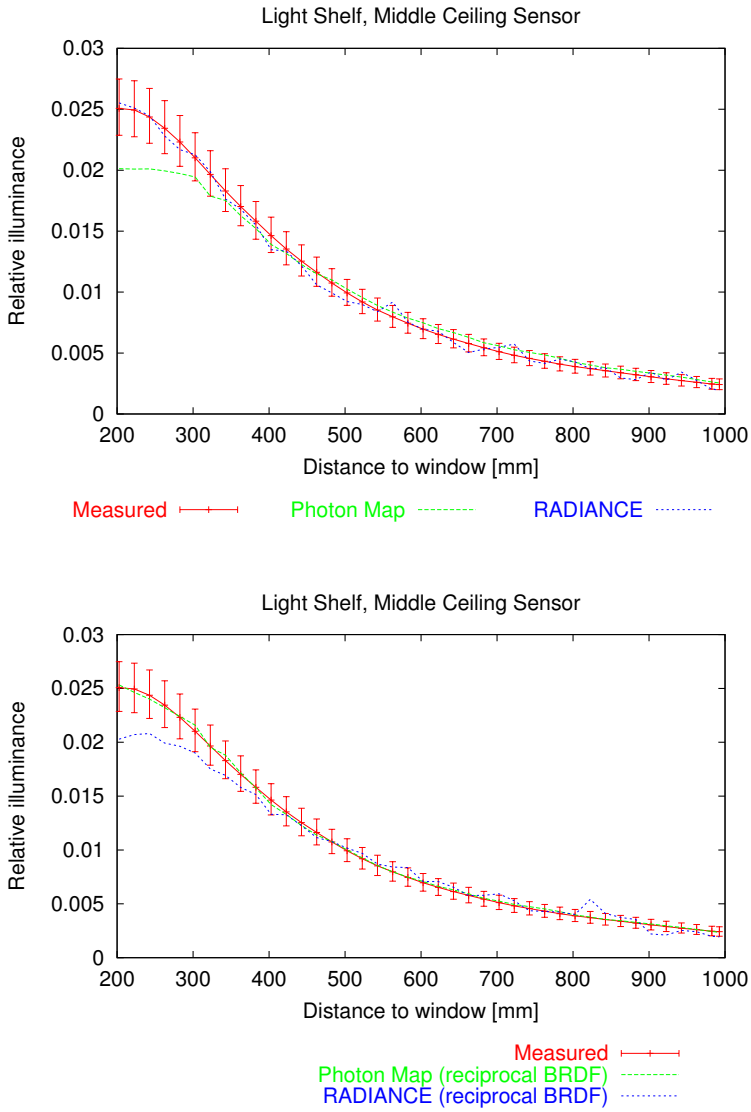


Figure 7.29: Results of component case study 3: simulated vs. measured ceiling sensor illuminance from light shelf caustic. Reciprocity deviations in the aluminium BRDF are evident when comparing simulations using the reciprocal BRDF (bottom) with those using the original (top).

back of the box due to the molleton's off-specular component.

7.4.4.2 Results

Initial simulations with RADIANCE Classic and photon map using the same molleton reflectance as in the component case studies ($\rho = 0.71$ at normal incidence) resulted in underprediction by some 10%. An investigation eventually revealed that

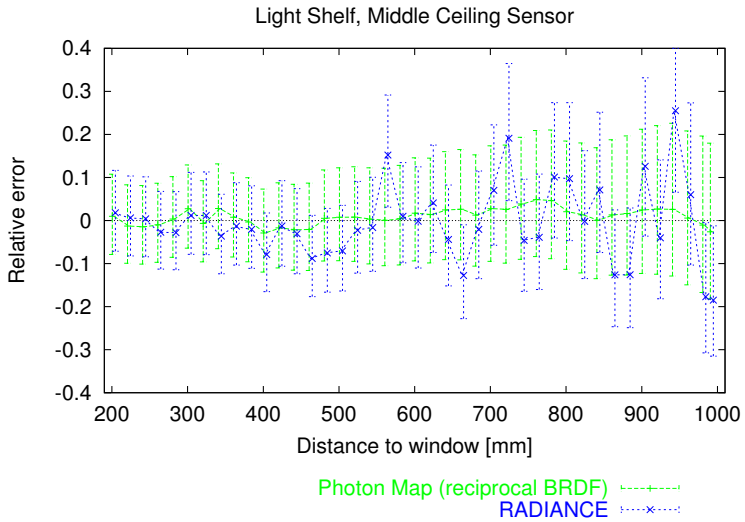


Figure 7.30: Results of component case study 3: relative error in simulated illuminance from light shelf caustic. Error bars are based on sensor tolerance in the measurement.

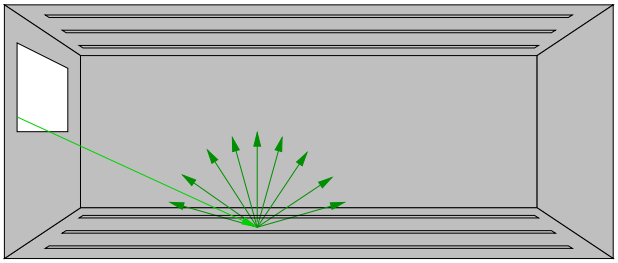


Figure 7.31: Compound case study 1: diffuse interreflection

the material's reflectance had actually *increased* to 0.724 for the side and ceiling sections, and even 0.754 for the floor. The cause of the increase in the molleton's reflectance appears to be due manufacturing tolerances, since the floor section was obtained from a different consignment than the other sections.

To account for the increase in reflectance the measured BRDF data was uniformly scaled by a factor of 0.724/0.71 and 0.754/0.71 for the corresponding surfaces during the simulation. While this is merely an increase of 2% and 6%, respectively, the interreflection amplifies this to yield a considerable increase in illuminance, particularly because the floor is directly illuminated and therefore contributes most. It was, however, assumed that the distribution of the BRDF was unchanged. The reflectance correction resulted in a substantially more accurate prediction as shown in figure 7.32. Error bars in the simulation account for the spectral shift effected by the molleton, which is amplified by interreflection and thus more significant than in the patch reflection case study. Error bars in the measurement account for

the sensor deviations, which are dominated by the relative component in contrast to the component case studies, where the absolute component has a greater impact due to the low illuminance.

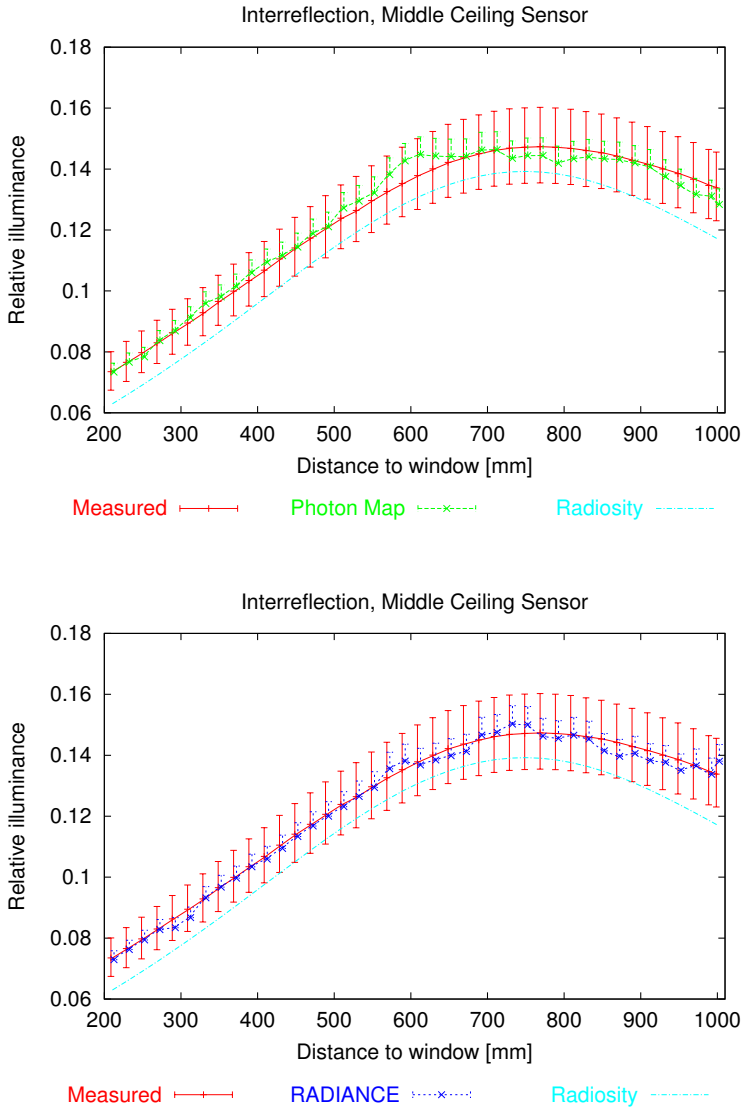


Figure 7.32: Results of compound case study 1: ceiling sensor illuminance from interreflection simulated with photon map (top) and RADIANCE Classic (bottom). Comparison with the radiosity solution obviates the impact of the molleton's off-specular component on the illuminance.

The off-specular component of the molleton BRDF clearly comes to bear here, and its effect is amplified by interreflection, as seen in the plot. This is evident by comparing the measurement with a purely diffuse radiosity solution using the

corrected reflectance. Not only does the off-specular component increase the illuminance by up to 14%, it also causes a slight shift of the illuminance peak towards the back of the box.

Figure 7.33 is a plot of the relative deviations for the ceiling sensor. The reflectance correction yields good agreement for both RADIANCE Classic and photon map, with remarkably low average deviations of 1% and 2%, respectively.

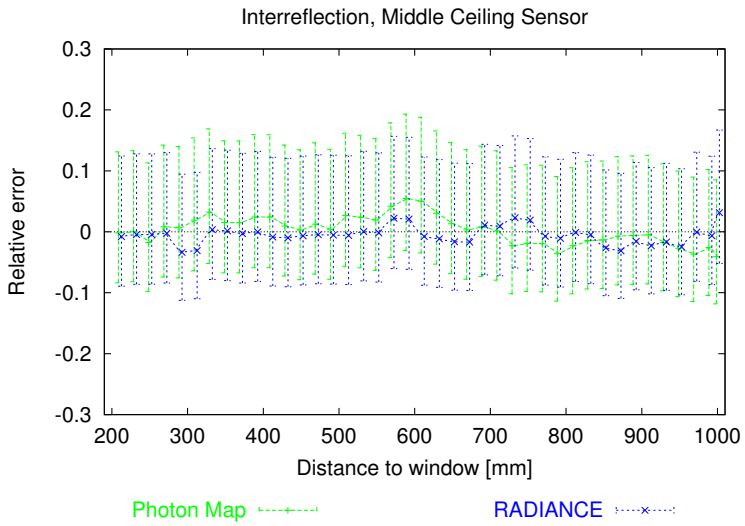


Figure 7.33: Results of compound case study 1: relative error in simulated illuminance from interreflection. Error bars are based on sensor tolerance in the measurement and spectral shift in the simulation.

7.4.5 Compound Case Study 2: Light Shelf Caustic and Diffuse Interreflection

7.4.5.1 Description

Combination of component case study 3 and compound case study 1. The aluminium light shelf is mounted on the outer window ledge of the reflecting box (figure 7.34). As in the previous case study, errors in the simulation are expected to be amplified by interreflection, and the illuminance is expected to rise with increasing window distance.

7.4.5.2 Results

As can be seen from the illuminance plot (figure 7.35), the region near the window is mostly under the influence of the caustic, producing a gentler rising slope in the illuminance compared to interreflection alone. As in the previous case study, the molleton BRDF data used in the simulations was scaled to match a reflectance at

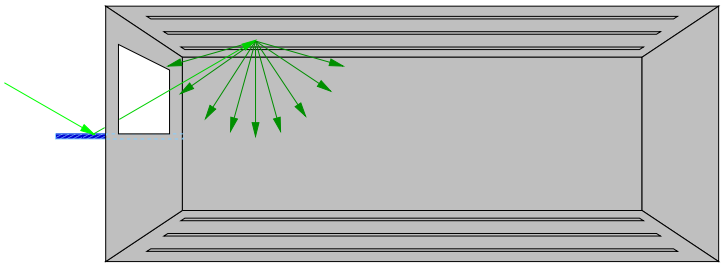


Figure 7.34: Compound case study 2: light shelf caustic and diffuse interreflection

normal incidence of 0.724 for the sides and ceiling, and 0.754 for the floor, thereby yielding similarly accurate predictions. As before, the error bars indicate estimated deviations due to spectral shift in the simulation and sensor tolerance in the measurement.

Relative deviations for the ceiling sensor are shown in figure 7.36. Here too, the reflectance correction yields low average deviations of 1% and 2% for RADIANCE Classic and photon map, respectively. The minor underprediction of the photon map near the window region is once again attributed to the aluminium BRDF's reciprocity deviations.

7.4.6 Performance Comparison

Table 7.1 lists relative computation times for the middle ceiling sensor. These times were obtained from simulation runs on a dual AMD Athlon 1.6 GHz system running Linux. Actual computation times ranged from 3 to 11 minutes for the photon map, and 47 to 434 minutes for RADIANCE Classic.

In all cases the photon map outperforms RADIANCE Classic by a factor of 16 or more. Calculating the illuminance for all sensors in a case study exacerbates this ratio even further; while there is no significant additional overhead for the photon map, RADIANCE Classic's cumulative computation time tends to increase linearly for every additional sensor position.

This discrepancy arises from the fact that the photon map is already a complete solution to the indirect illumination. Thus, the bulk of the photon map computation time is consumed during the distribution step; the actual gathering step is nearly instantaneous, since we have chosen to evaluate the photon illuminance directly for the validation, and therefore no ambient rays are traced. Once the photon map is built, we can reuse it with minimum penalty for an arbitrary number of sensor positions. RADIANCE Classic, on the other hand, must trace additional ambient rays for every sensor position, despite the fact that ambient files were used in order to reuse the irradiance cache accumulated from previous sensor positions. The high overhead of RADIANCE Classic for the validation case studies naturally stems from the parametrisation necessary to achieve accurate results. It is clear, however, that the photon map is capable of delivering comparable accuracy in a substantial

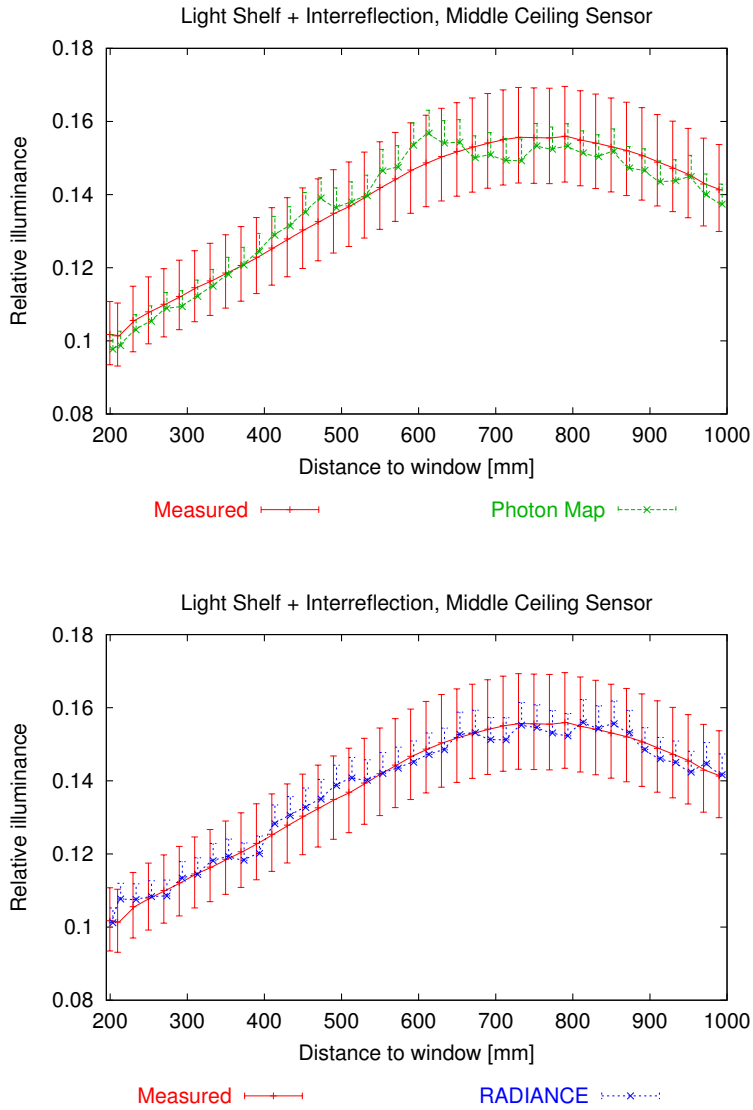


Figure 7.35: Results of compound case study 2: ceiling sensor illuminance from light shelf caustic and interreflection simulated with photon map (top) and RADIANCE Classic (bottom).

fraction of the time.

7.5 Conclusions

We have presented a physical validation based on measurements of a simple scale model with an artificial light source. The proposed methodology is characterised by:

- Emphasis on tractability and error minimisation.

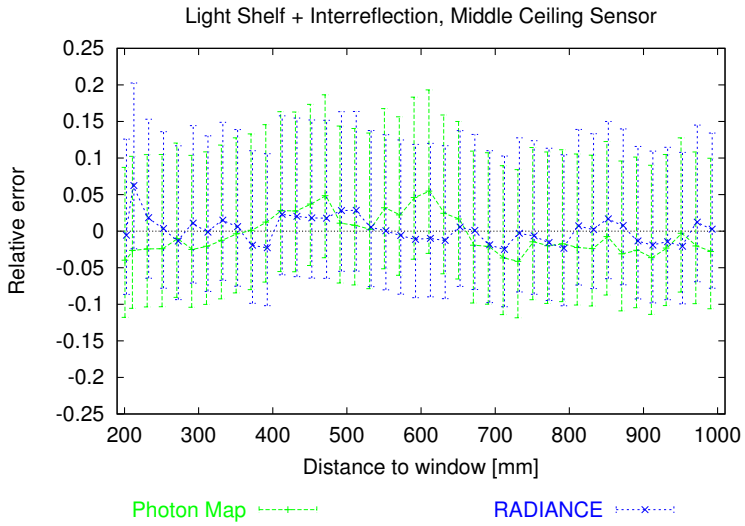


Figure 7.36: Results of compound case study 2: relative error in simulated illuminance from light shelf caustic and interreflection. Error bars are based on sensor tolerance in the measurement and spectral shift in the simulation.

Case study	Time [Photon map:RADIANCE Classic]
Patch reflection	1:39
Light shelf	1:22
Interreflection	1:16
Light shelf & interreflection	1:16

Table 7.1: Relative computation times for ceiling sensor illuminance.

- Simple component case studies testing individual light transport modes (single diffuse and specular reflection). These serve as foundation for the more complex compound case studies.
- Compound case studies testing interreflection (diffuse only and in conjunction with a single specular reflection).
- Analytical solutions where possible to assess the accuracy of the measurements themselves, not just the simulation.

Physical accuracy necessitated incorporating the EDF of the light source and the material BRDFs in the simulation. The EDF was obtained with a novel method based on extraction from HDR camera images. The BRDFs of the materials were obtained from goniophotometric measurements. Physical validation requires attention to detail and scrutiny in order to minimise errors on the physical side. Since the simulations used measured BRDF data, errors on the physical side will be carried over into the simulation. Consequently, both data sets were subjected to analysis and verification prior to simulation, as well as correction in the case of the BRDF

data due to limitations of the goniophotometer device. The BRDF data also necessitated the development of a resampling technique based on nearest neighbour lookups in a bilevel kd-tree in order to efficiently evaluate the BRDF at arbitrary incident and exitant directions during the simulation.

The uncertainties inherent in a physical validation were discussed, particularly pertaining to measurement inaccuracies. We have identified primary sources of error and accounted for them where necessary. These uncertainties are carried over into the resulting illuminance plots as error bounds. The most problematic factor of the experimental validation were the material BRDFs. An accurate simulation required using the measured BRDF data directly in place of a fitted analytical BRDF model. Instead, the fitted model served a subordinate purpose, providing the PDF sample ray distribution in the simulation. Furthermore, the molleton's reflectance was subject to manufacturing tolerances which contributed significantly to deviations in compound case studies due to interreflection, while the aluminium BRDF exhibited reciprocity deviations which were noticeable in component case studies.

The validation tested several aspects of both RADIANCE Classic and the photon map, including:

- Indirect illumination with respect to diffuse-diffuse as well as specular-diffuse (caustic) transfers.
- Non-trivial BRDFs (specular and off-specular).
- Light source EDF (particularly critical due to the small solid angle subtended by the scale model).

The results of the compound case studies show that both algorithms deliver very similar results. The average deviations are similarly low except for the floor sensors due to inaccuracies in the direct component. However, the stock RADIANCE code responsible for this component is common to both algorithms and thus only of marginal significance to the validation. Conversely, sensor positions dominated by indirect illumination are of primary interest, and for these both algorithms exhibit deviations under 10% compared to the measurements. This is well within the tolerances proposed by the CIE for lighting design software [Fis92] (10% for averaged and 20% for point illuminance), consequently the photon map is fundamentally suitable for lighting applications.

While both algorithms perform similarly as far as accuracy is concerned, the computation times differ substantially. In all case studies the photon map outperformed RADIANCE Classic by a minimum factor of 16. This disparity is attributed to the fact that the photon illuminance is evaluated directly rather than via final gathering, but even in the latter case performance is still generally superior to RADIANCE Classic. The photon map's primary advantage, namely the construction of a complete, reusable global illumination solution in the forward pass, comes to bear when sampling multiple sensor positions, while RADIANCE Classic's performance suffers under the burden of additional ambient rays for each sensor position.

It should be re-emphasised as on the outset that the primary intent of the validation was to test the fundamental algorithm behind the photon map, and that the validation should be considered exemplary. This implies that the accuracy and performance of the photon map obtained with these case studies is by no means guaranteed for all geometries and parametrisations. The forward pass is particularly efficient with mostly closed environments with high reflectance such as the validation test box, whereas RADIANCE Classic inevitably performs poorly in such situations due to the large number of ambient bounces required for accurate interreflection. Although complex in detail, the scope of the validation was infact too narrow to provide an assessment of the photon map's accuracy on a broad scale.

Physical validation methodologies are still in their infancy. This is borne out by our own experience and the scarcity of available references and measurement data. Research in physically based rendering is only recently addressing the need to develop a standardised framework for physical validation. The validation described here was an interdisciplinary undertaking, bringing together computer scientists and physicists. Its documentation on these pages is intended to contribute to the development of a standardised framework and aid in future validation prospects.

In concluding, we have shown that the adequately parametrised photon map algorithm delivers accuracy comparable to that of RADIANCE Classic, only generally faster. One a more fundamental level, this chapter and the previous have demonstrated the soundness of both the forward and backward raytracing paradigms in theory and practice. However, the validation does not single out the "better" algorithm. Both have their strengths and weaknesses, and are consequently suitable for complementary applications in daylighting. It is clear that forward raytracers excel at simulating caustics and as such are well suited to efficiently handling light redirecting systems typically used in daylighting configurations, but are burdened by distant light sources used by sky models during the forward pass. By contrast, backward raytracers have difficulty with caustics but can handle distant sources with ease by sampling them directly. The photon map extension to RADIANCE unifies these methods and makes the best of both worlds available to lighting consultants for the evaluation of configurations using both artifical and natural light.

As of this writing, the photon map extension has only seen experimental use within the RADIANCE community, but raised the interest particularly of those who have run into RADIANCE Classic's limitations when the issue of caustics arises. A gradual migration from RADIANCE Classic to photon map is therefore expected specifically for tasks involving the analysis of light redirecting systems, thereby establishing a possible future niche for photon map applications within the RADIANCE community.

Appendix A

The RADIANCE Rendering System

RADIANCE is a global illumination rendering system based on a Monte Carlo ray tracer developed by Greg Ward [War94, LS98] primarily for lighting and architecture. To the author's knowledge it is the only open-source physically based renderer that has seen commercial use and undergone validation [Gry89, MLH93]. It was chosen as a framework for the modifications that were implemented in the scope of this thesis since it has a sizeable user base which will (hopefully!) benefit from this work.

RADIANCE is not a single monolithic renderer but rather a suite of small, portable, command-line driven programs which perform specific tasks as part of a simulation. There are utilities for generating geometry, manipulating and converting renderings, a semi-interactive previewer, and the ray tracer itself. A fully interactive OpenGL based previewer is also available separately[1].

The set of geometric primitives supported by RADIANCE is indeed primitive, consisting of spheres, discs, cones, cylinders, and polygons. However, a set of utilities comes with RADIANCE which can generate approximations of more complex geometries based on these primitives. These include solids of revolution, parallelepipeds, and curved surfaces.

RADIANCE employs stratified sampling to determine the diffuse indirect (termed *ambient* in the RADIANCE literature) irradiance at a point. An additional supersampling pass with finer intervals is performed if this illumination exhibits a high gradient, from nearby specular objects for example. The specular component is sampled separately using importance sampling based on the BRDF (see appendix B.1).

A unique feature of RADIANCE is its *irradiance caching* scheme [WRC88, WH92], which reuses the ambient values from nearby points and interpolates between them. The spacing between the cached values is derived from a user-defined

[1] *Rshow* by Peter Apian-Bennewitz, available at `http://www.pab-opto.de/progs/rshow/rshow.html`

error metric. This can dramatically improve performance as the number of ambient computations (and hence the number of traced rays) is reduced. However, the irradiance cache can backfire and propagate inherent noise in the ambient values if the number of sample rays is too low, producing unsightly splotches on surfaces. The irradiance cache's parametrisation requires some experience to achieve acceptable results in difficult cases.

The direct illumination is sampled from area light sources associated with a scene object. In order to produce penumbrae their surfaces are partitioned and shadow rays cast towards each partition to prevent clustering, thereby determining the fraction of the visible area. The order in which light sources are sampled is optimized via an adaptive scheme which samples sources in order of decreasing potential contribution until some cutoff point, after which a statistical estimate is added for the remaining sources [War91a]. A special source primitive caters to daylight simulation in the form of a directional light source with a solid angle which is subtended at every point in the scene. This source is not associated with any scene geometry and at an infinite distance, thus simulating the sun. Arbitrary EDFs can be specified via functions or goniometric data files, the latter usually available from manufacturers of lighting fixtures.

As with all physically based renderers, the images produced by RADIANCE require adaptation of the dynamic range for proper viewing. A sophisticated tone mapping utility [LRP97] is included which models limitations of the HVS. In dark regions it accounts for low acuity, reduced colour saturation and contrast, whereas in bright regions it accounts for veiling due to internal scattering in the eye lens.

Appendix B

Analytical BRDF Models

B.1 The Ward Gaussian BRDF Model

The standard BRDF model employed in RADIANCE is Ward's Gaussian model [War92], which models specular peaks as Gaussian lobes. It is a general model capable of approximating most benign materials with simple BRDFs. The isotropic case is defined as follows (figure B.1):

$$f_{r,iso}\left(\vec{x},\vec{\omega}_i,\vec{\omega}_r\right)=\frac{\rho_d}{\pi}+\frac{\rho_s e^{-\left(\frac{\tan\delta}{\alpha}\right)^2}}{4\pi\alpha^2\sqrt{\left(\vec{N}_{\vec{x}}\cdot\vec{\omega}_i\right)\left(\vec{N}_{\vec{x}}\cdot\vec{\omega}_r\right)}}, \tag{B.1}$$

where

- ρ_d is the diffuse reflectance
- ρ_s is the specular reflectance
- $\vec{N}_{\vec{x}}$ is the surface normal
- δ is the angle between $\vec{N}_{\vec{x}}$ and the halfway vector $\vec{H}$ which is defined as

$$\vec{H}=\frac{\vec{\omega}_i+\vec{\omega}_r}{\|\vec{\omega}_i+\vec{\omega}_r\|} \tag{B.2}$$

- α is the RMS surface slope and gives a measure of the surface roughness.

The specular reflectance defines the amplitude of the Gaussian lobe while the roughness term defines its width.

The model is also capable of anisotropic effects by using two independent slope distributions α_u and α_v along two perpendicular axes $\vec{u}$ and $\vec{v}$ on the surface. The anisotropic case then becomes

$$f_{r,aniso}\left(\vec{x},\vec{\omega}_i,\vec{\omega}_r\right)=\frac{\rho_d}{\pi}+\frac{\rho_s e^{-\tan^2\delta\left(\left(\frac{\cos\phi}{\alpha_u}\right)^2+\left(\frac{\sin\phi}{\alpha_v}\right)^2\right)}}{4\pi\alpha_u\alpha_v\sqrt{\left(\vec{N}_{\vec{x}}\cdot\vec{\omega}_i\right)\left(\vec{N}_{\vec{x}}\cdot\vec{\omega}_r\right)}}, \tag{B.3}$$

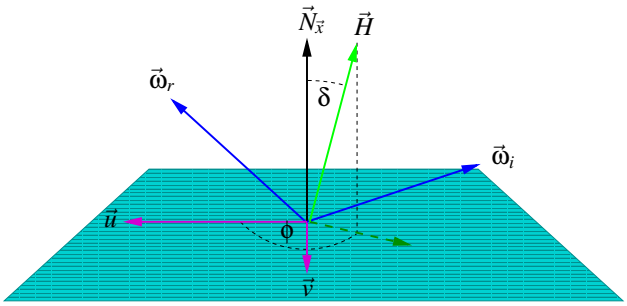


Figure B.1: Ward's Gaussian model. $\vec{\omega}_i$ and $\vec{\omega}_r$ are defined in terms of local coordinates.

where ϕ is the angle between $\vec{u}$ and the projection of $\vec{H}$ onto the surface.

Using Monte Carlo inversion techniques, a direction for a specular sampling ray is generated as follows:

$$\delta = \tan^{-1}\sqrt{\frac{-\log(\xi_1)}{\left(\frac{\cos\phi}{\alpha_u}\right)^2+\left(\frac{\sin\phi}{\alpha_v}\right)^2}} \tag{B.4}$$

$$\phi = \tan^{-1}\left(\frac{\alpha_u}{\alpha_v}\tan(2\pi\xi_2)\right), \tag{B.5}$$

$$\xi_1,\xi_2 \in [0,1],$$

where ξ_1 and ξ_2 are uniform random variables, and δ and ϕ are defined as above.

B.2 The Lafortune BRDF Model

This very general model due to Lafortune [LFTG97] sees widespread use. Its popularity stems from its simplicity and ability to cover a broad range of materials and reflection behaviour, while still being efficient and easy to implement. The parametrisation is however not as intuitive as with the Ward model, and best accomplished via a fit to measured data.

The model is a generalisation of the simple cosine lobe model [Lew93] based on a sum of independent exponentiated cosine lobes. Each lobe j is characterised by a coefficient vector $\vec{C}_j$, and an exponent n_j which specify the lobe's position, amplitude, and width for the given incident and outgoing directions. The BRDF is defined as

$$f_{r,l}(\vec{\omega}_i,\vec{\omega}_r) = \sum_{j=1}^{N}\left(C_{j,x}\omega_{i,x}\omega_{r,x}+C_{j,y}\omega_{i,y}\omega_{r,y}+C_{j,z}\omega_{i,z}\omega_{r,z}\right)^{n_j}, \tag{B.6}$$

where the incident and exitant directions $\vec{\omega}_i$ and $\vec{\omega}_r$ are defined in terms of the surface's local coordinate system. Negative BRDF values are clipped to 0. For a

more intuitive interpretation, the model can also be rewritten as

$$f_{r,l}(\vec{\omega}_i, \vec{\omega}_r) = \sum_{j=1}^{N} f_{r,l,j}(\vec{\omega}_i, \vec{\omega}_r) \tag{B.7}$$

$$= \sum_{j=1}^{N} (\vec{\omega}_j \cdot \vec{\omega}_r)^{n_j}, \tag{B.8}$$

$$\vec{\omega}_j = (C_{j,x}\omega_{i,x},\, C_{j,y}\omega_{i,y},\, C_{j,z}\omega_{i,z}). \tag{B.9}$$

This reduces the model to a traditional cosine lobe model with a scaled lobe axis $\vec{\omega}_j$ whose magnitude and position is based on the lobe coefficients $\vec{C}_j$ and the incident direction $\vec{\omega}_i$. In this way, the model can account for a reflectance which varies with incident angle.

The choice of lobe coefficients specifies the reflection characteristics the model accounts for:

lambertian reflection: $n_j = 0$

anisotropy: $C_{j,x} \neq C_{j,y}$

retro-reflection: $C_{j,x}, C_{j,y} > 0$

off-specularity: $|C_{j,x}|, |C_{j,y}| > C_{j,z}$, where the ratios $|C_{j,x}|/C_{j,z}$ and $|C_{j,y}|/C_{j,z}$ define the degree of off-specularity.

The individual lobe reflectances can be computed efficiently for integer exponents with a method by Arvo [Arv95, Dut03]. This is an approximation for lobes with floating point exponents, which we will denote $\widetilde{\rho}_{l,j}$, thereby also implying the approximated total reflectance $\widetilde{\rho}_l$. Similarly, $\widetilde{f}_{r,l,j}$ denotes the component BRDFs evaluated for each lobe using integer exponents $\lfloor n_j \rfloor$, implying the approximated total BRDF $\widetilde{f}_{r,l}$.

To apply Monte Carlo techniques, the BRDF can be importance sampled by selecting a lobe *j* according to its reflectance $\widetilde{\rho}_{l,j}$, and generating a direction on the unit hemisphere with a probability proportional to $\cos^{\lfloor n_j \rfloor}\theta$:

$$\theta = \cos^{-1}\left(\xi_1^{\frac{1}{\lfloor n_j \rfloor + 1}}\right) \tag{B.10}$$

$$\phi = 2\pi\xi_2, \tag{B.11}$$

$$\xi_1, \xi_2 \in [0,1].$$

The resulting direction is then transformed to the lobe axis $\vec{\omega}_j$ to obtain the sample direction $\vec{\omega}_r$. Sample directions which penetrate the surface are rejected. The sample's normalisation factor[1] $\widetilde{I}_{l,j}$ can be obtained for integer exponents with another method by Arvo similar to that for the lobe's reflectance $\widetilde{\rho}_{l,j}$ [Dut03]. Note that

[1] The normalisation factor is not the same as the reflectance in this case because the PDF does not account for $\cos\theta_r$ after transformation to the lobe axis.

because the reflectance and normalisation factor are based on integer exponents, the sampled directions must also be generated for integer exponents $\lfloor n_j \rfloor$.

Working with integer exponents imposes unacceptable constraints on the BRDF and introduces discontinuity with respect to the exponents, which complicates fits to measured BRDF data. A workaround is to compensate for the terms derived from integer exponents by scaling the samples as if they had effectively been based on a BRDF with floating point exponents. This gives us the bias-free estimator:

$$\begin{aligned} \widehat{L}(\vec{\omega}_i) &= \overbrace{\frac{\widetilde{\rho}_l(\vec{\omega}_i)}{\widetilde{\rho}_{l,j}(\vec{\omega}_i)}}^{\text{Lobe compensation}} \cdot \overbrace{\frac{\widetilde{I}_{l,j}(\vec{\omega}_i)}{\widetilde{f}_{r,l,j}(\vec{\omega}_i,\vec{\omega}_r)}}^{\text{Direction compensation}} f_{r,l,j}(\vec{\omega}_i,\vec{\omega}_r)\, L(\vec{\omega}_r)\cos\theta_r, \\ \vec{\omega}_r &\in \Omega, \\ j &\in [1,N], \end{aligned} \tag{B.12}$$

where the respective compensating factors for nonuniform sampling of the lobe and direction are indicated in overbraces.

Appendix C

Analytical Sky Models

The diffuse hemispherical sky source used in daylight simulation (see figure 1.4) accounts for sunlight scattered by the atmosphere. However, in reality the directional distribution of sky luminance is rarely uniform. Furthermore, cloud cover distributions are required in order to model more varied sky conditions. To account for such luminance distributions, *sky models* were developed, which may be considered analogous to goniometric data for artificial sources. Unlike the latter, however, sky models describe luminance distributions for a sensor at a fixed location. Like BRDFs, analytic sky models are generally empirical, i.e. based on comparisons with measured sky conditions, although it is also possible to integrate the data directly into the simulation as is the case with the RADIANCE system. The state of the art for measuring sky luminance distributions is the *skyscanner*, a luminance sensing device which scans the sky hemisphere and records both the diffuse (sky without sun) and directional (sun only) components [PBB+94]. However, skyscanner data is scarce and site specific, whereas analytical sky models are more flexible and easier to parametrise for a particular location and sky condition. Mardaljevic [Mar99] has demonstrated the validity of some of these models based on skyscanner data for a number of sky conditions. Note that all presently available sky models based on diffuse and direct luminance generate continuous luminance distributions, and do not explicitly model clouds, but merely account for the influence of cloud cover in the diffuse luminance.

C.1 CIE Overcast Sky

The simplest sky model is also the oldest, known as the Moon and Spencer sky before its standardisation by the CIE in 1955 [CIE94]. It accounts for the overcast sky's gradation from a dark horizon to a bright zenith, and is rotationally symmetric about the zenith. The luminance for an incident direction $\vec{\omega}_i$ at an angle θ_i to the zenith (see figure C.1) is given by:

$$L(\vec{\omega}_i) = L_z \frac{1 + 2\cos\theta_i}{3}, \tag{C.1}$$

where L_z is the zenith luminance. The horizon luminance is therefore 1/3 of the zenith luminance. Trivial as it may seem, the model agrees with measured data for dull sky conditions [KV93]. This model is primarily used in daylight factor evaluation.

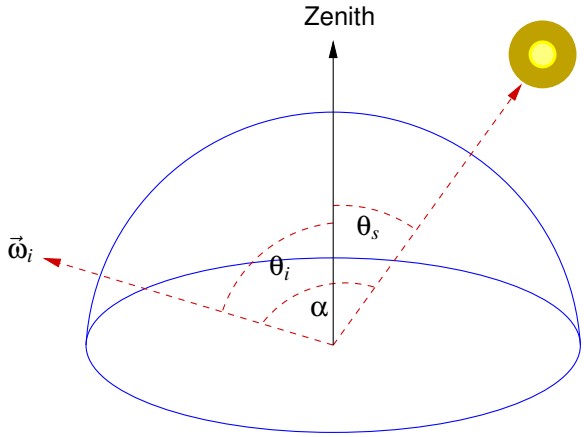


Figure C.1: Sky model geometry.

Figure C.2 is a 3D plot of this sky model. The plot represents the sky dome projected onto the ground plane, with luminance proportional to elevation and colour. The rotational symmetry and luminance maximum at the zenith are obvious in the plot.

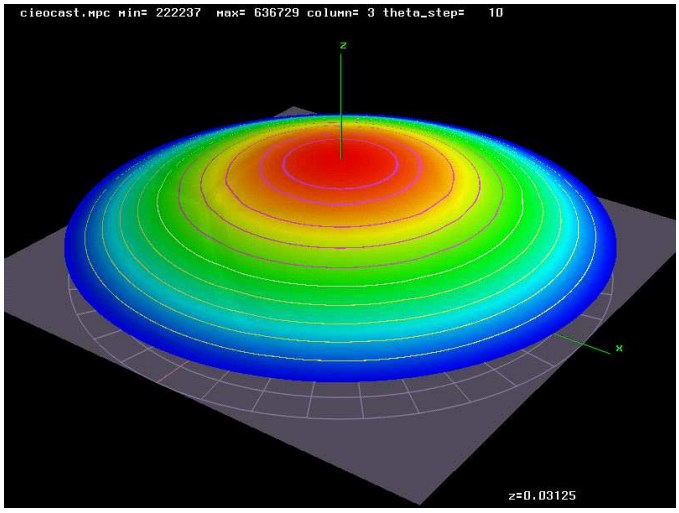


Figure C.2: CIE overcast sky model.

C.2 CIE Clear Sky

The more complex CIE clear sky was originally proposed by Pokrowski based on theory and measurements, which was modified by Kittler before being adopted as a standard by the CIE [CIE94]. The model accounts for a bright circumsolar region, a point with minimum sky luminance, and brightening towards the horizon.

$$L(\vec{\omega}_i) = L_z \frac{\left(0.91 + 10e^{-3\alpha} + 0.45\cos^2\alpha\right)\left(1 - e^{-0.32/\cos\theta_i}\right)}{\left(0.91 + 10e^{-3\theta_s} + 0.45\cos^2\theta_s\right)\left(1 - e^{-0.32}\right)}, \qquad \text{(C.2)}$$

where θ_s is the zenith angle to the sun's position, and α is the angle between the sun's position and the incident direction $\vec{\omega}_i$ under consideration (see figure C.1).

Figure C.3 is an example of a CIE clear sky with the sun located at an angle of $\theta_s = 30^\circ$ to the zenith. Note the peak representing the circumsolar region and the elevation (luminance increase) towards the horizon.

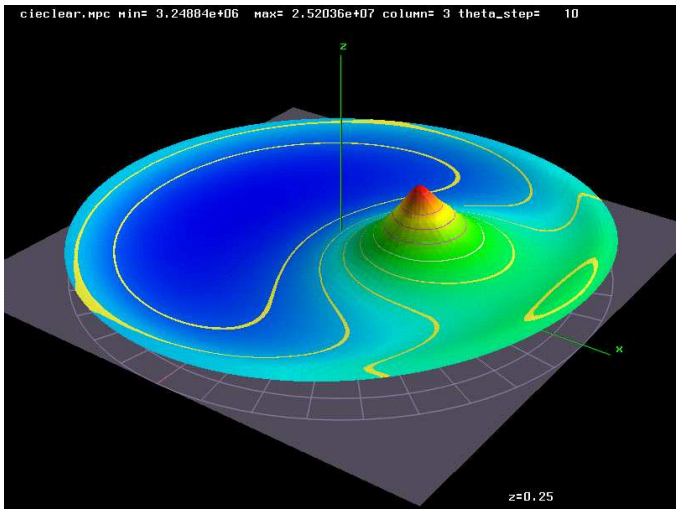


Figure C.3: CIE clear sky model.

The *gensky* utility in the RADIANCE suite can generate scene descriptions according to the CIE clear and overcast sky models.

C.3 Perez All-Weather Model

The CIE overcast and clear skies are extreme sky conditions and do not account for more common intermediate skies. The Perez all-weather model [PSM93] addresses this issue. It is parametrised with five distribution coeffcients $c_1, \ldots, c_5$ relating to

- darkening or brightening towards the horizon
- the luminance gradient towards the horizon

- the relative luminance of the circumsolar region
- the width of the circumsolar region
- the relative backscattered luminance.

These coefficients define a distribution function which scales the zenith luminance L_z:

$$d(\theta_i, \alpha) = \left(1 + c_1 e^{c_2/\cos\theta_i}\right)\left(1 + c_3 e^{c_4 \alpha} + c_5 \cos^2 \alpha\right) \quad \text{(C.3)}$$

$$L(\vec{\omega}_i) = L_z \frac{d(\theta_i, \alpha)}{d(0, \theta_s)}, \quad \text{(C.4)}$$

with θ_s and α defined as in the previous section.

Figure C.4 is an example of an intermediate sky obtained with the Perez model, with the sun located at an angle of $\theta_s = 30^\circ$ to the zenith. The moderate cloud cover causes the circumsolar region to widen.

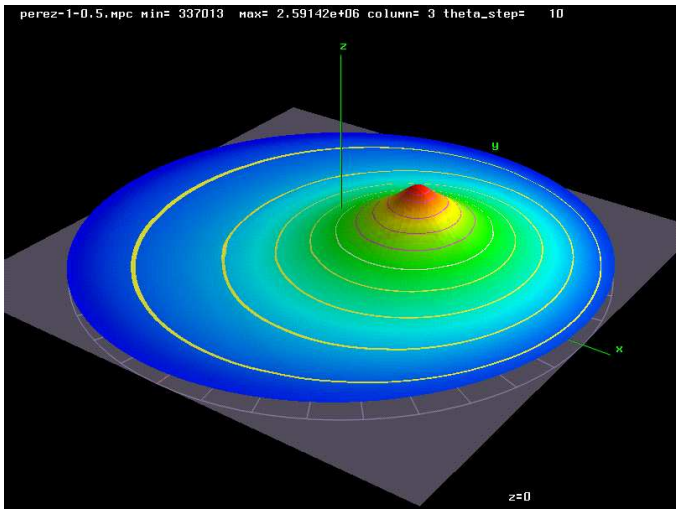


Figure C.4: Intermediate sky obtained with Perez model.

The Perez model is in widespread use due to its generality and ease of parametrisation. In a recent study comparing various sky models, the Perez model exhibited the lowest deviation from measured data [Var00]. This model has been improved specifically for computer graphics applications to account for the spectral properties of daylight [PSS99], based on sun position and atmospheric turbidity.

The *gendaylit* utility [Del95] is an extension to the RADIANCE system which generates scene descriptions according to the Perez sky model.

Appendix D

Analytical Validation with Lafortune BRDF

The analytical validation using the Lafortune BRDF model is based on the spherical furnace paradigm employed in chapter 6. For simplicity, the BRDF model consists of a constant diffuse term and a single cosine lobe aligned with the surface normal, i.e. $C_x = C_y = 0$.

The derivation of the reference value for the validation is based on a series expansion similar to equation 6.9, but complicated by the directional component of the BRDF. As a consequence, the irradiance is constant for any point on the inner sphere surface, but the radiance varies with the incident angle. The derivation can be simplified by treating directions as polar coordinates. The resultant expression appears to defy a closed form representation, and is instead distilled into a recursive schema.

The irradiance E at a point $\vec{x}_1$ on the inner sphere surface is defined as

$$E = E_l + \int_{\theta_{i,1}} \int_{\phi_{i,1}} L_i(\vec{x}_1, \theta_{i,1}, \phi_{i,1}) \cos\theta_{i,1} d\omega_{i,1}, \tag{D.1}$$

where E_l is the constant point light source irradiance and $\theta_{i,1}, \phi_{i,1}$ are incident angles relative to the surface normal at $\vec{x}_1$. The incident radiance L_i is recursively defined as

$$\begin{aligned} L_i(\vec{x}_1, \theta_{i,1}, \phi_{i,1}) &= L_r(\vec{x}_2, \theta_{r,2}, \phi_{r,2}) \\ &= E_l f_r(\theta_{i,l}, \phi_{i,l}, \theta_{r,2}, \phi_{r,2}) + \\ &\quad \int_{\theta_{i,2}} \int_{\phi_{i,2}} L_i(\vec{x}_2, \theta_{i,2}, \phi_{i,2}) f_r(\theta_{i,2}, \phi_{i,2}, \theta_{r,2}, \phi_{r,2}) \cos\theta_{i,2} d\omega_{i,2}, \end{aligned} \tag{D.2}$$

where $\vec{x}_2$ is a point contributing radiance incident at $\vec{x}_1$ along the direction $\theta_{i,1}, \phi_{i,1}$ (see figure D.1). The exitant direction $\theta_{r,2}, \phi_{r,2}$ is defined in terms of the surface normal at $\vec{x}_2$.

Due to the spherical topology, the expression can be simplified substantially:

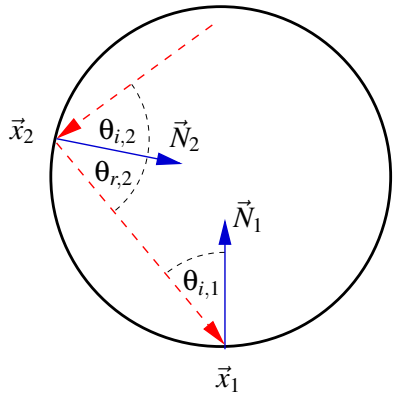


Figure D.1: Spherical validation geometry. The incident angle $\theta_{i,1}$ at $\vec{x}_1$ is equal to the exitant angle $\theta_{r,2}$ at $\vec{x}_2$.

- since the light source is at the sphere center, $\theta_{i,l} = \phi_{i,l} = 0$,
- the incident angle $\theta_{i,j}$ at $\vec{x}_j$ is equal to the exitant angle $\theta_{r,j+1}$ at $\vec{x}_{j+1}$,
- the BRDF and radiance are symmetric about the normal and position-independent, reducing them to $f_r(\theta_i, \theta_r)$ and $L(\theta)$.

Substituting D.2 into D.1 and simplifying, the expression now becomes

$$\begin{aligned} E &= E_l + \\ &\quad 2\pi \int_{\theta_1} \Big[E_l f_r(0, \theta_1) + \\ &\qquad 2\pi \int_{\theta_2} L(\theta_2) f_r(\theta_2, \theta_1) \cos\theta_2 \sin\theta_2 d\theta_2 \Big] \cos\theta_1 \sin\theta_1 d\theta_1 . \end{aligned} \tag{D.3}$$

The BRDF is defined as

$$f_r(\theta_i, \theta_r) = f_{r,d} + (C_z \cos\theta_i \cos\theta_r)^n , \tag{D.4}$$

where $f_{r,d}$ is the constant diffuse term, and C_z and n are the cosine lobe's Z-coefficient and exponent, respectively. Plugging this into D.3 yields

$$\begin{aligned} E &= E_l + \\ &\quad 2\pi \int_{\theta_1} \Big[\left[C_z^n \cos^n\theta_1 + f_{r,d} \right] E_l + \\ &\qquad 2\pi \Big[C_z^n \cos^n\theta_1 \int_{\theta_2} L(\theta_2) \cos^{n+1}\theta_2 \sin\theta_2 d\theta_2 + \\ &\qquad\quad f_{r,d} \int_{\theta_2} L(\theta_2) \cos\theta_2 \sin\theta_2 d\theta_2 \Big] \Big] \cos\theta_1 \sin\theta_1 d\theta_1 \\ &= E_l + 2\pi \Big[\frac{C_z^n}{n+2} \Big[E_l + 2\pi \int_{\theta_2} L(\theta_2) \cos^{n+1}\theta_2 \sin\theta_2 d\theta_2 \Big] + \end{aligned}$$

$$\frac{f_{r,d}}{2}\left[E_l + 2\pi\int_{\theta_2} L(\theta_2)\cos\theta_2\sin\theta_2 d\theta_2\right]\Bigg]. \tag{D.5}$$

We now expand the series:

$$\begin{aligned}
E &= E_l + \\
&\quad 2\pi\left[\frac{C_z^n}{n+2}\left[E_l + 2\pi\left[\frac{C_z^n}{2n+2}\left[E_l + 2\pi\int_{\theta_3} L(\theta_3)\cos^{n+1}\theta_3\sin\theta_3 d\theta_3\right] + \right.\right.\right. \\
&\qquad \left.\left.\left.\frac{f_{r,d}}{n+2}\left[E_l + 2\pi\int_{\theta_3} L(\theta_3)\cos\theta_3\sin\theta_3 d\theta_3\right]\right]\right] + \\
&\quad \frac{f_{r,d}}{2}\left[E_l + 2\pi\left[\frac{C_z^n}{n+2}\left[E_l + 2\pi\int_{\theta_3} L(\theta_3)\cos^{n+1}\theta_3\sin\theta_3 d\theta_3\right] + \right.\right. \\
&\qquad \left.\left.\left.\frac{f_{r,d}}{2}\left[E_l + 2\pi\int_{\theta_3} L(\theta_3)\cos\theta_3\sin\theta_3 d\theta_3\right]\right]\right]\right] \\
&= E_l\left[1 + 2\pi\left[\frac{C_z^n}{n+2} + \frac{f_{r,d}}{2}\right] + \right. \\
&\quad (2\pi)^2\left[\frac{C_z^n}{n+2}\left[\frac{C_z^n}{2n+2} + \frac{f_{r,d}}{n+2}\right] + \frac{f_{r,d}}{2}\left[\frac{C_z^n}{n+2} + \frac{f_{r,d}}{2}\right]\right] + \\
&\quad (2\pi)^3\left[\frac{C_z^n}{n+2}\left[\frac{C_z^n}{2n+2}\left[\frac{C_z^n}{2n+2} + \frac{f_{r,d}}{n+2}\right] + \right.\right. \\
&\qquad \left.\left.\frac{f_{r,d}}{n+2}\left[\frac{C_z^n}{n+2} + \frac{f_{r,d}}{2}\right]\right] + \right. \\
&\qquad \frac{f_{r,d}}{2}\left[\frac{C_z^n}{n+2}\left[\frac{C_z^n}{2n+2} + \frac{f_{r,d}}{n+2}\right] + \right. \\
&\qquad \left.\left.\left.\frac{f_{r,d}}{2}\left[\frac{C_z^n}{n+2} + \frac{f_{r,d}}{2}\right]\right]\right] + \quad \cdots \quad \Bigg].
\end{aligned} \tag{D.6}$$

This leads to the recursive expression

$$E = E_l\,[1+a] \tag{D.7}$$

$$a = 2\pi\left[\frac{f_{r,d}}{2}\,[1+a] + \frac{C_z^n}{n+2}\,[1+b]\right] \tag{D.8}$$

$$b = 2\pi\left[\frac{f_{r,d}}{n+2}\,[1+a] + \frac{C_z^n}{2n+2}\,[1+b]\right]. \tag{D.9}$$

Appendix E

Experimental Validation BRDFs

The BRDF plots of the molleton fabric used for the test box interior and the aluminium profile used as light shelf in caustic case studies were obtained with ISE's goniophotometer [AB95]. Each plot depicts the hemisphere of outgoing directions $\vec{\omega}_r \in \Omega_r$ for a fixed incident direction $\vec{\omega}_i$. $\vec{\omega}_r$ is projected onto the plane, with the BRDF $f_r(\vec{\omega}_i, \vec{\omega}_r)$ corresponding to elevation and colour. These 3D plots are more expressive than the traditional goniometric polar plots often encountered in the literature, which are usually restricted to the plane of incidence. Note that each plot is scaled separately and the colour bands denote relative elevation ranging from minimum (blue) to maximum (red). Consequently the plots are unsuitable for quantitative comparison.

Measurements were done for incident angles of 0°–78° in 1° intervals, with 78° being the maximum angle which can be measured reliably with the goniophotometer due to optical and mechanical limitations. This precludes obtaining BRDF data for grazing angles. A further problem incurred by the device are the self-shadowing artifacts from the goniophotometer's arm and sensor cable visible in the plots. The affected points were removed, and the resultant BRDF data was used for the BRDF model fit and the validation case study simulations.

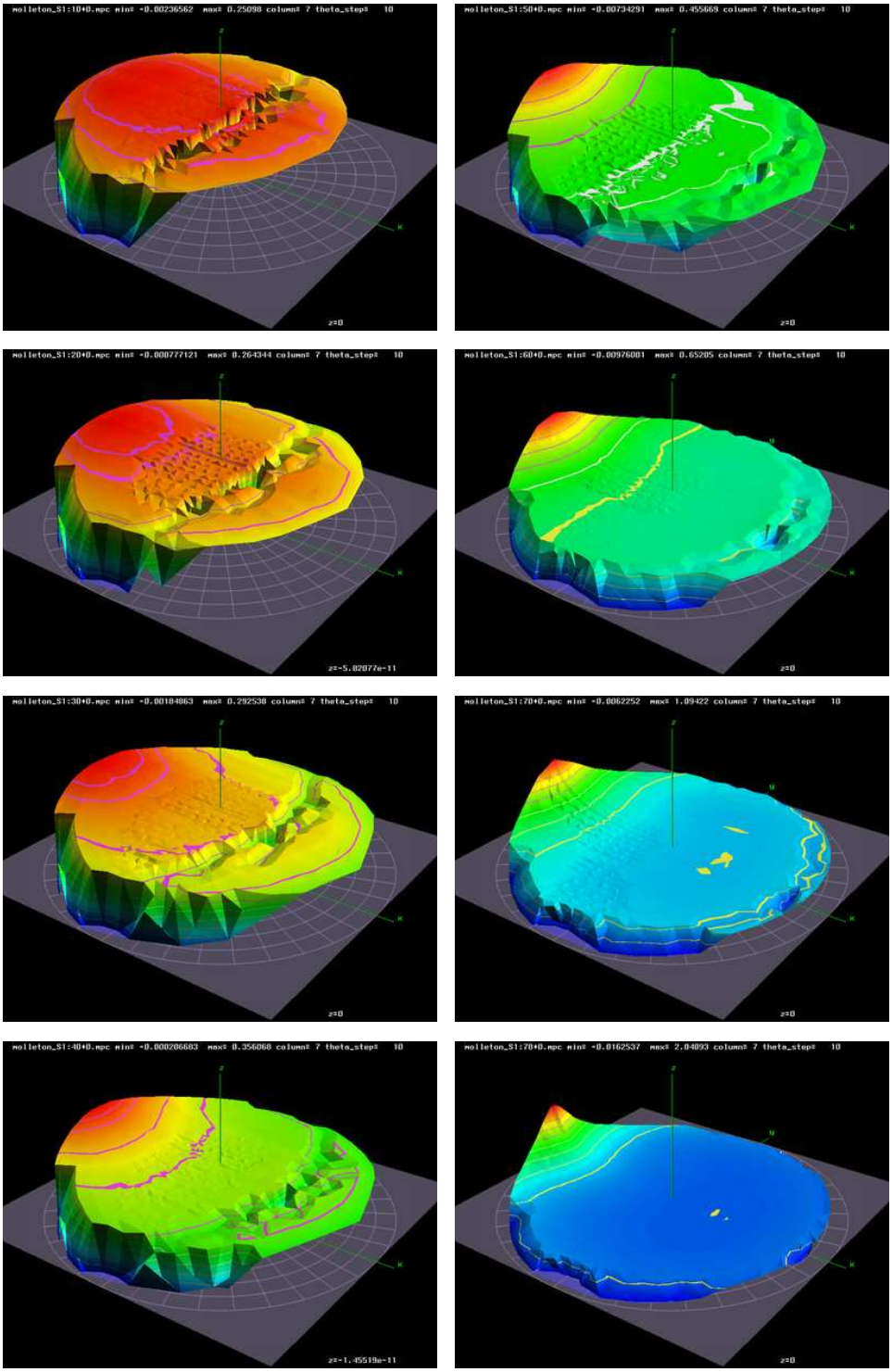


Figure E.1: Molleton BRDF for incident angles of 10°, 20°, 30°, 40° (left column), 50°, 60°, 70°, and 78° (right column). The material is lambertian at normal incidence, but exhibits an off-specular component towards grazing angles as evidenced by the emerging slope. The rifts and boundary falloff in the plots and are self-shadowing artifacts from the goniophotometer.

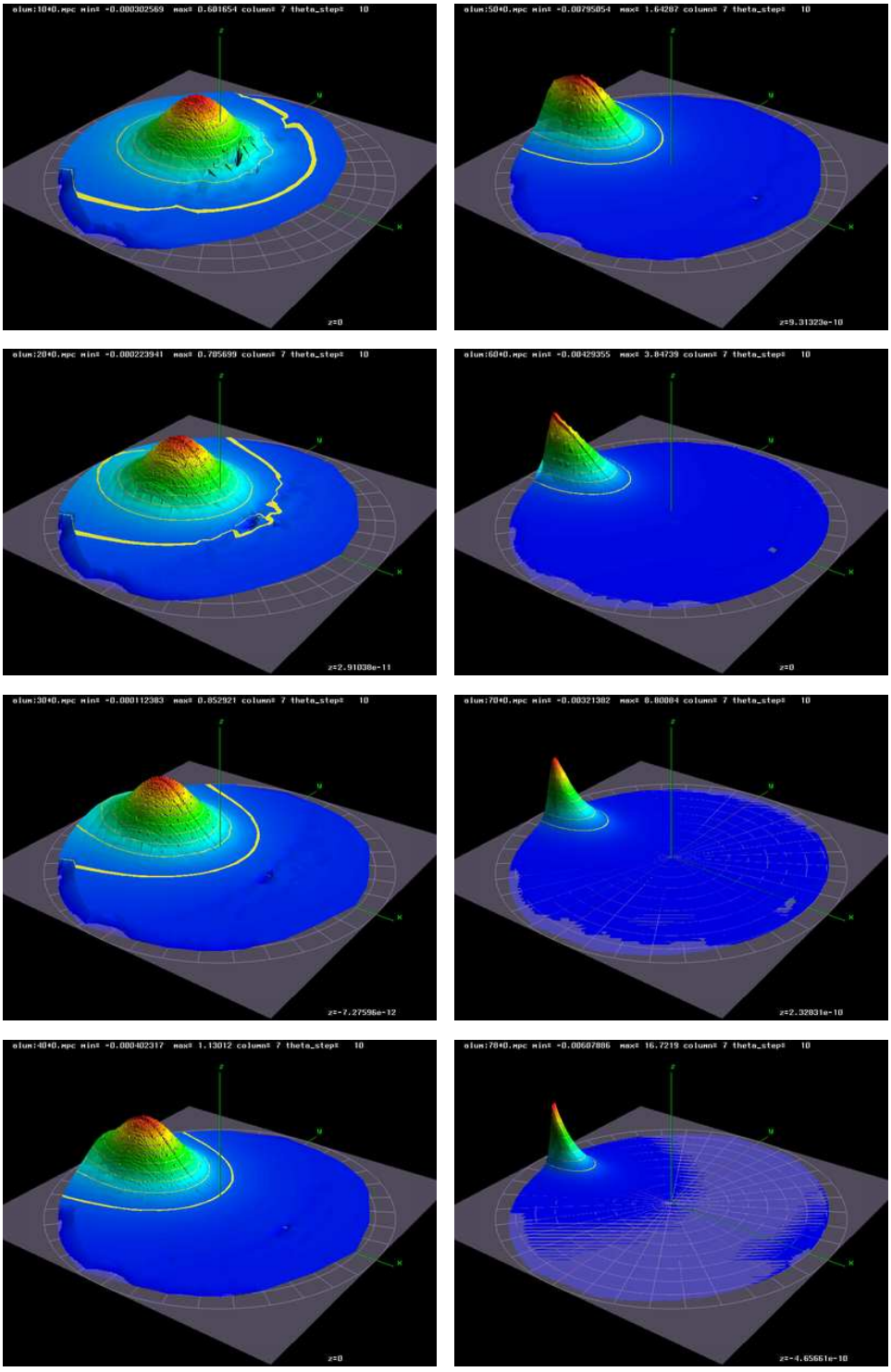


Figure E.2: Aluminium BRDF for incident angles of 10°, 20°, 30°, 40° (left column), 50°, 60°, 70°, and 78° (right column). The gaussian peak indicates specular reflection with elongation and off-specularity towards grazing angles.

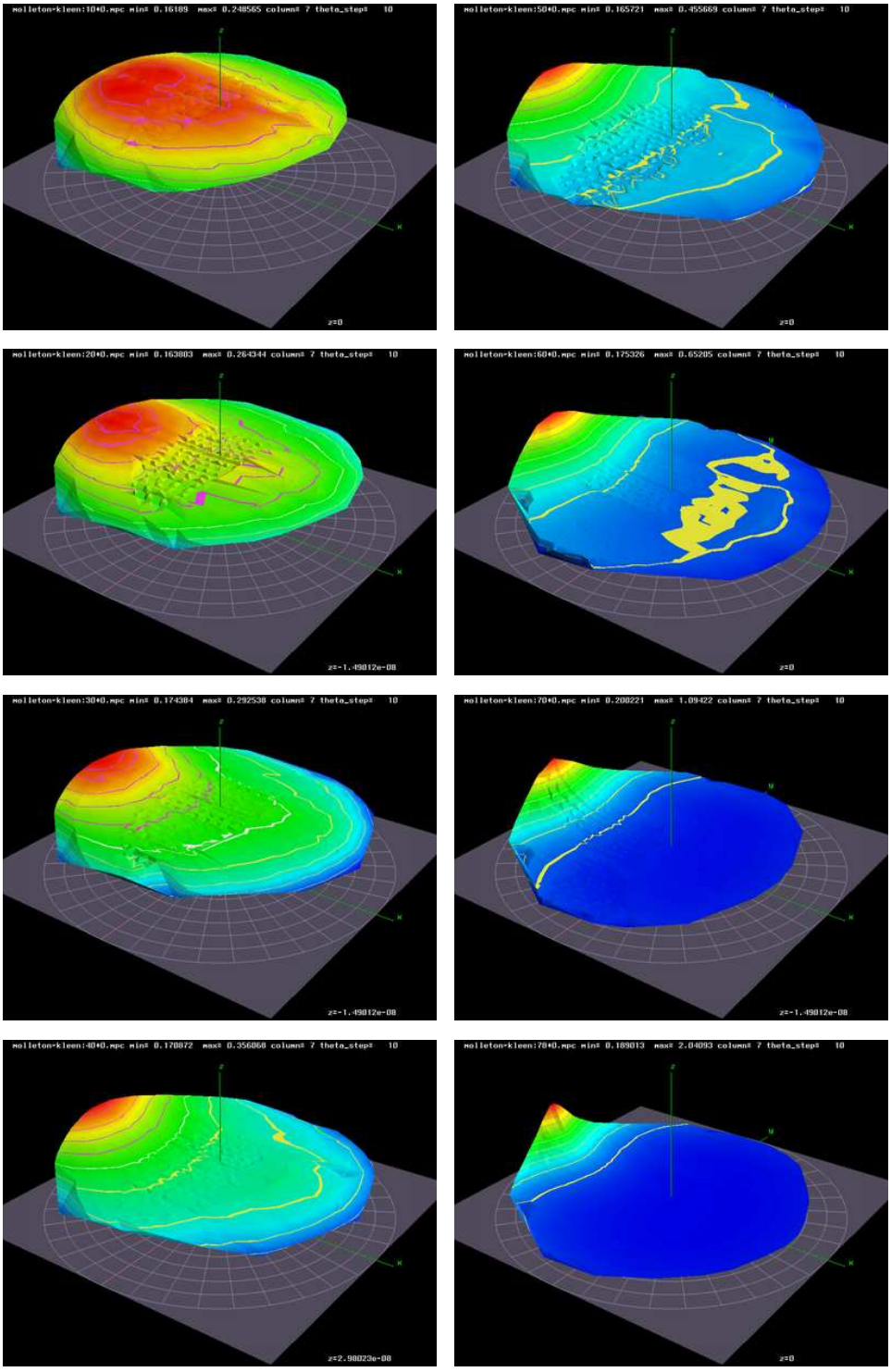


Figure E.3: Molleton BRDF after artifact removal for incident angles of 10°, 20°, 30°, 40° (left column), 50°, 60°, 70°, and 78° (right column). Samples subject to self-shadowing artifacts from the goniophotometer have been deleted.

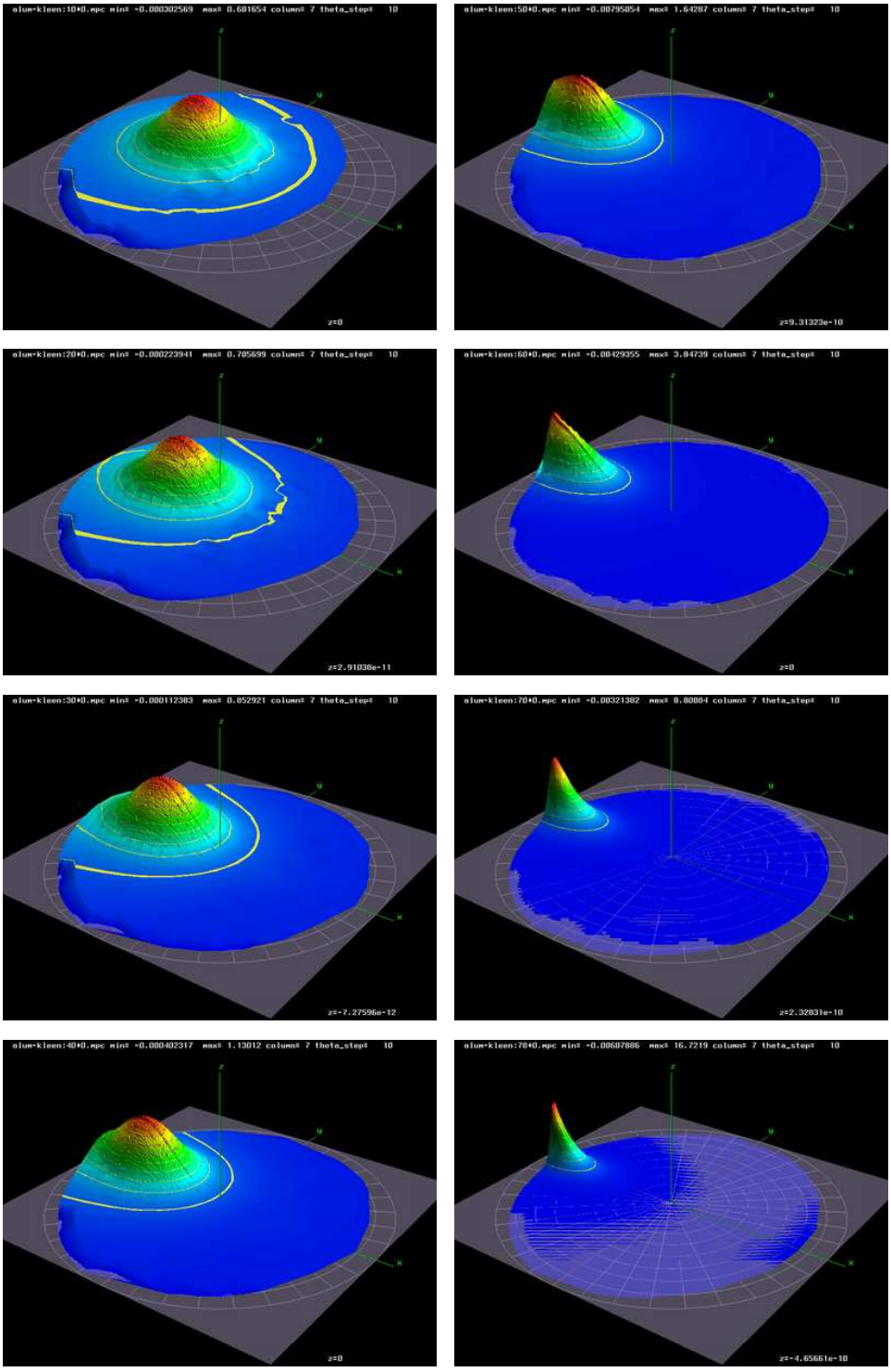


Figure E.4: Aluminium BRDF after artifact removal for incident angles of 10°, 20°, 30°, 40° (left column), 50°, 60°, 70°, and 78° (right column). Samples subject to self-shadowing artifacts from the goniophotometer have been deleted.

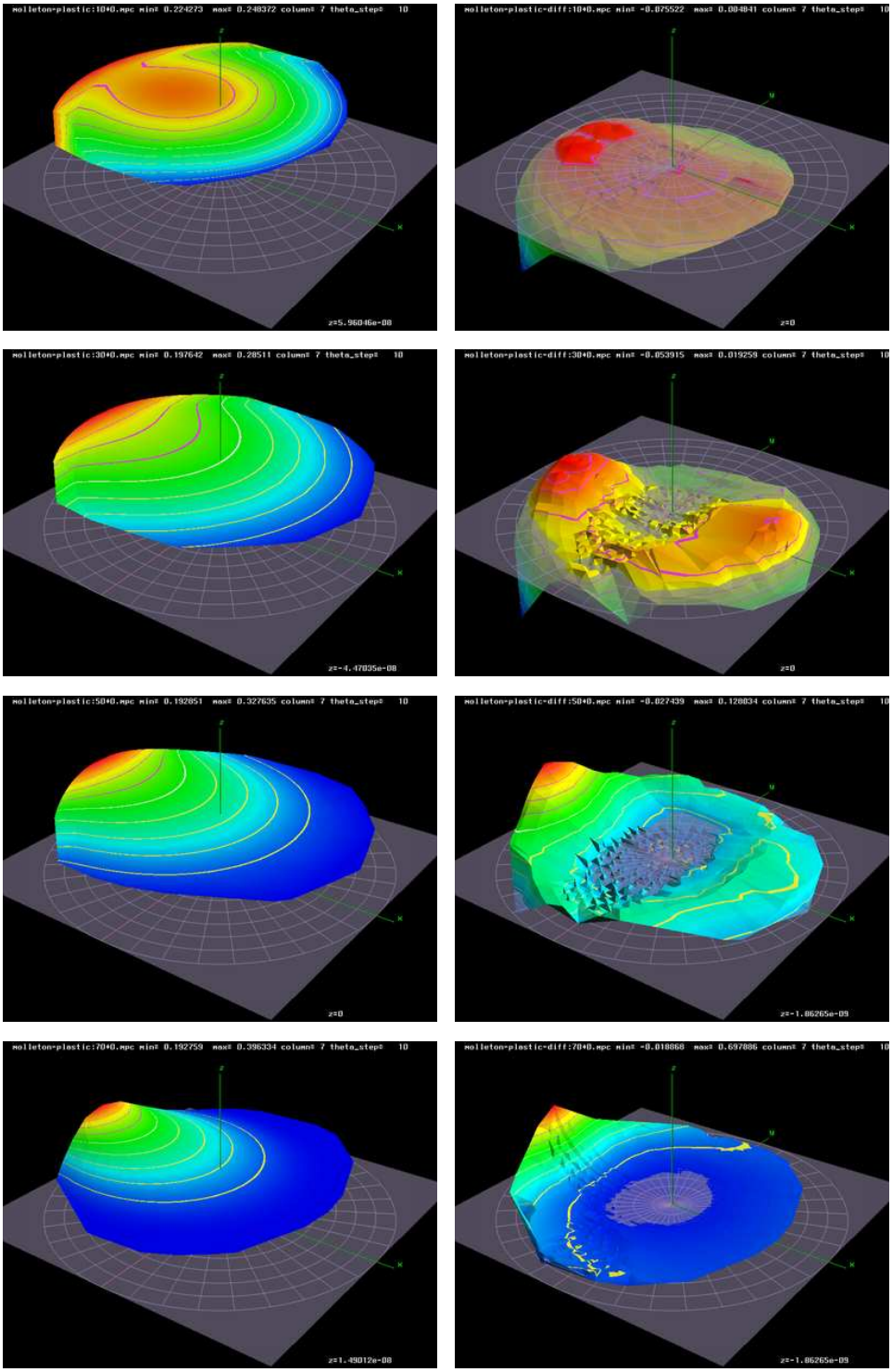


Figure E.5: Ward model fitted to molleton BRDF (left column) and difference to measured data (right column) for incident angles of 10°, 30°, 50°, and 70°. The model gives a poor approximation, mainly because it does not account for the off-specularity and consequently deviates considerably with increasing incident angle, giving rise to a slope in the difference plots.

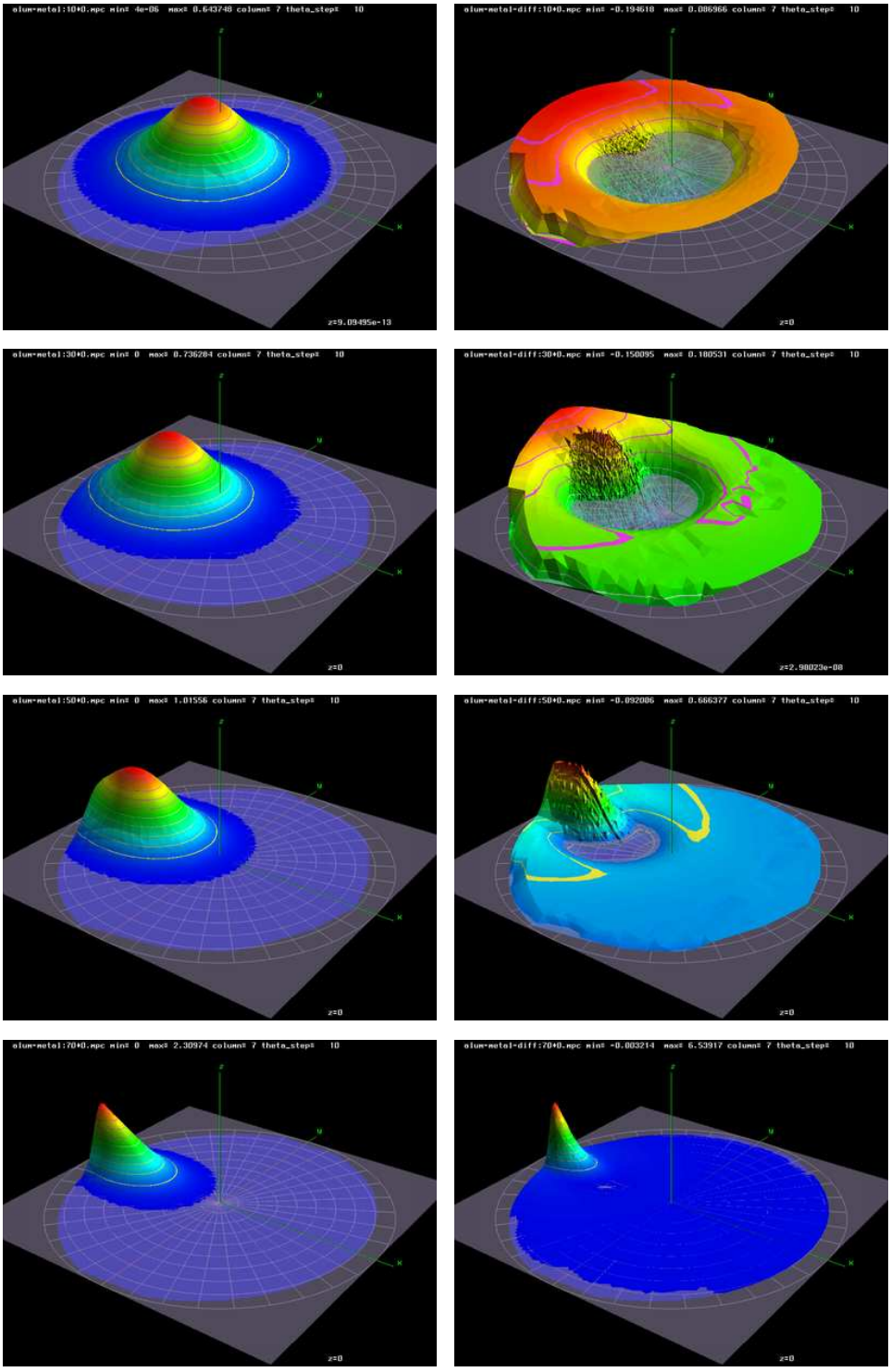


Figure E.6: Ward model fitted to aluminium BRDF (left column) and difference to measured data (right column) for incident angles of 10°, 30°, 50°, and 70°. The model gives a reasonable approximation, with deviations due mainly to the slight off-specularity of the measured peak.

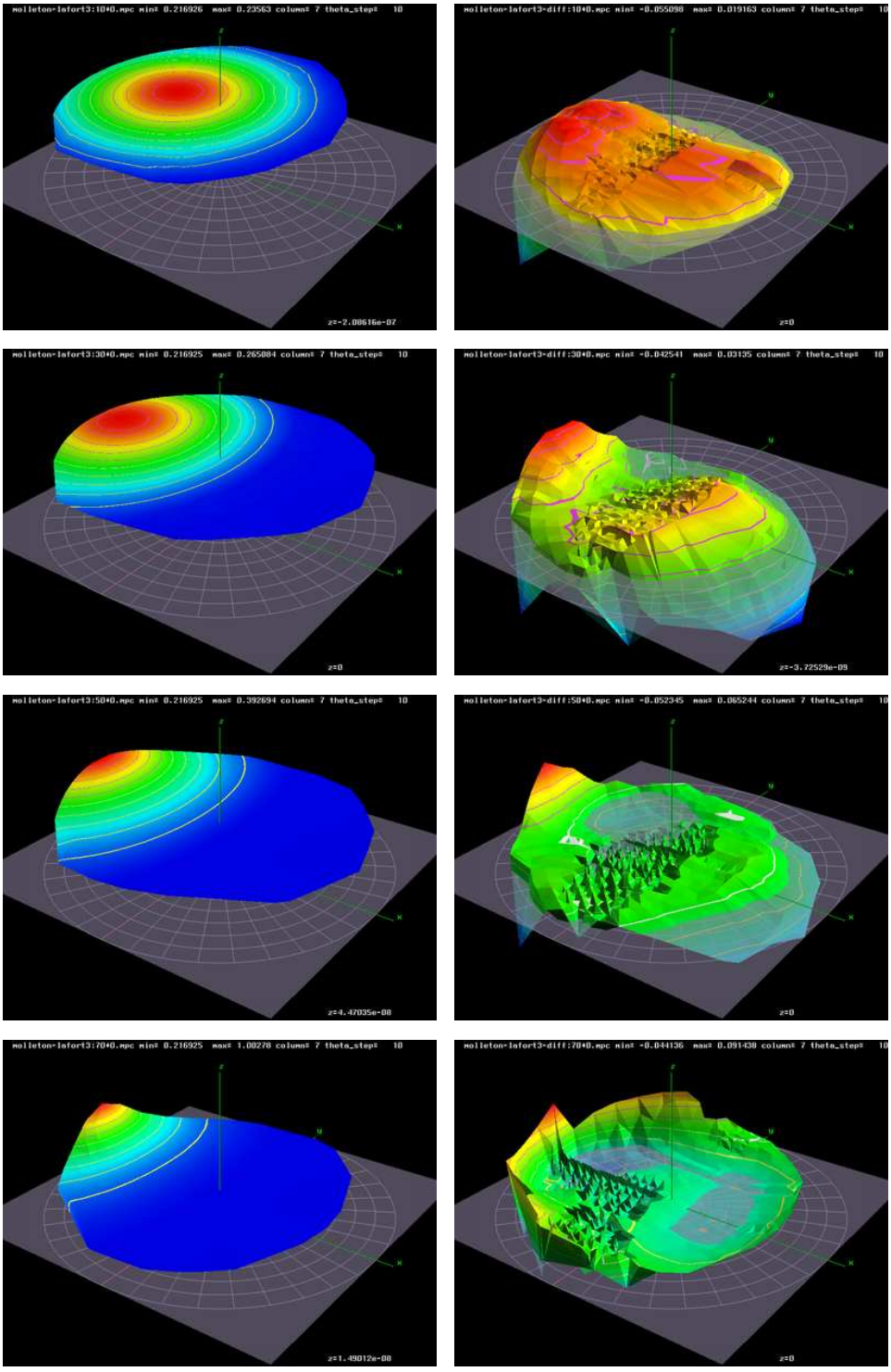


Figure E.7: Lafortune model fitted to molleton BRDF with 3 cosine lobes and Lambertian term (left column) and difference to measured data (right column) for incident angles of 10°, 30°, 50°, and 70°. The model gives a better approximation than the Ward model, accounting for the off-specularity.

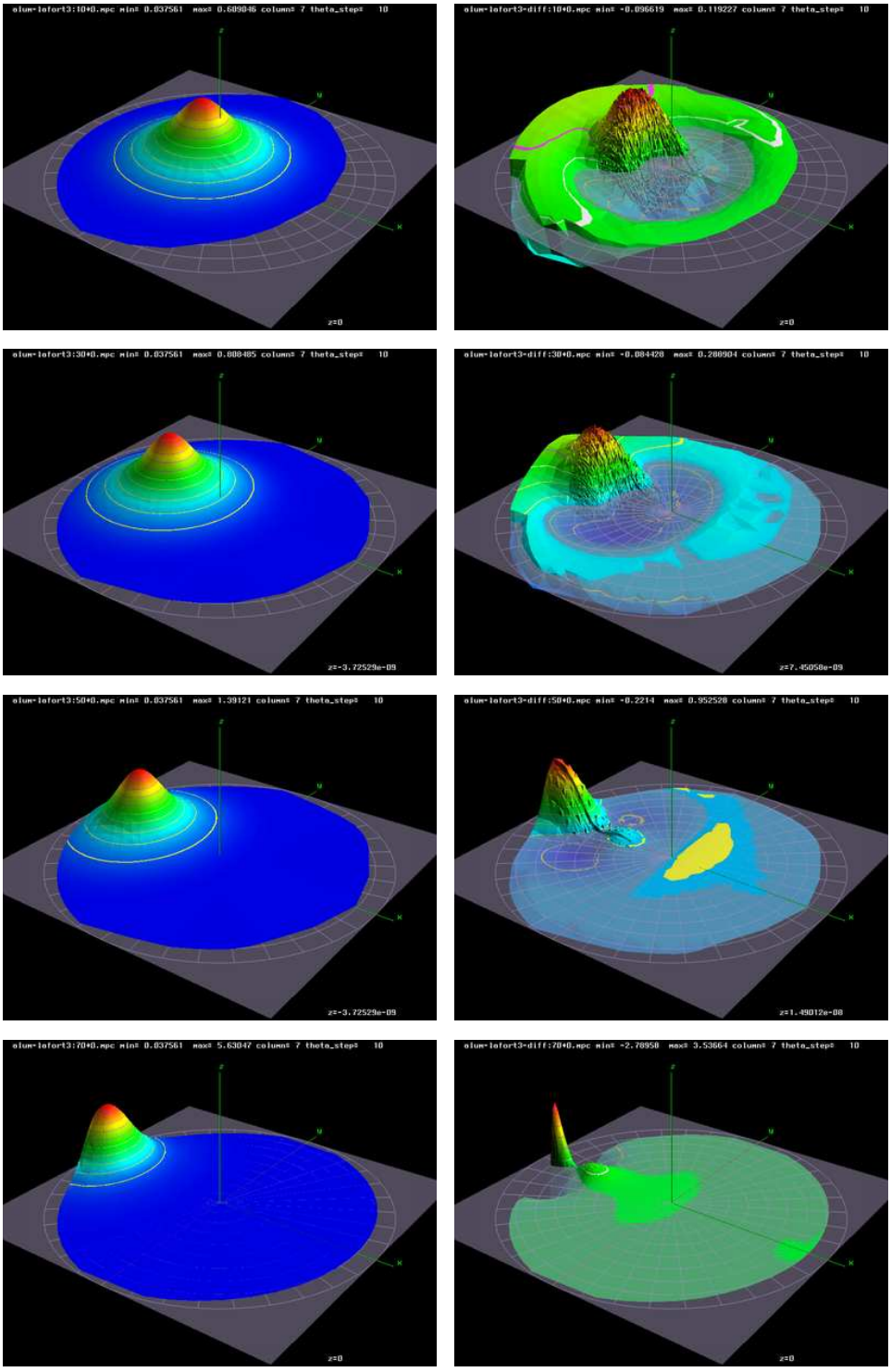


Figure E.8: Lafortune model fitted to aluminium BRDF with 3 cosine lobes and Lambertian term (left column) and difference to measured data (right column) for incident angles of 10°, 30°, 50°, and 70°. The model accounts for the slight off-specularity, but not for the elongation of the peak, producing two flanking lobes in the difference plots.

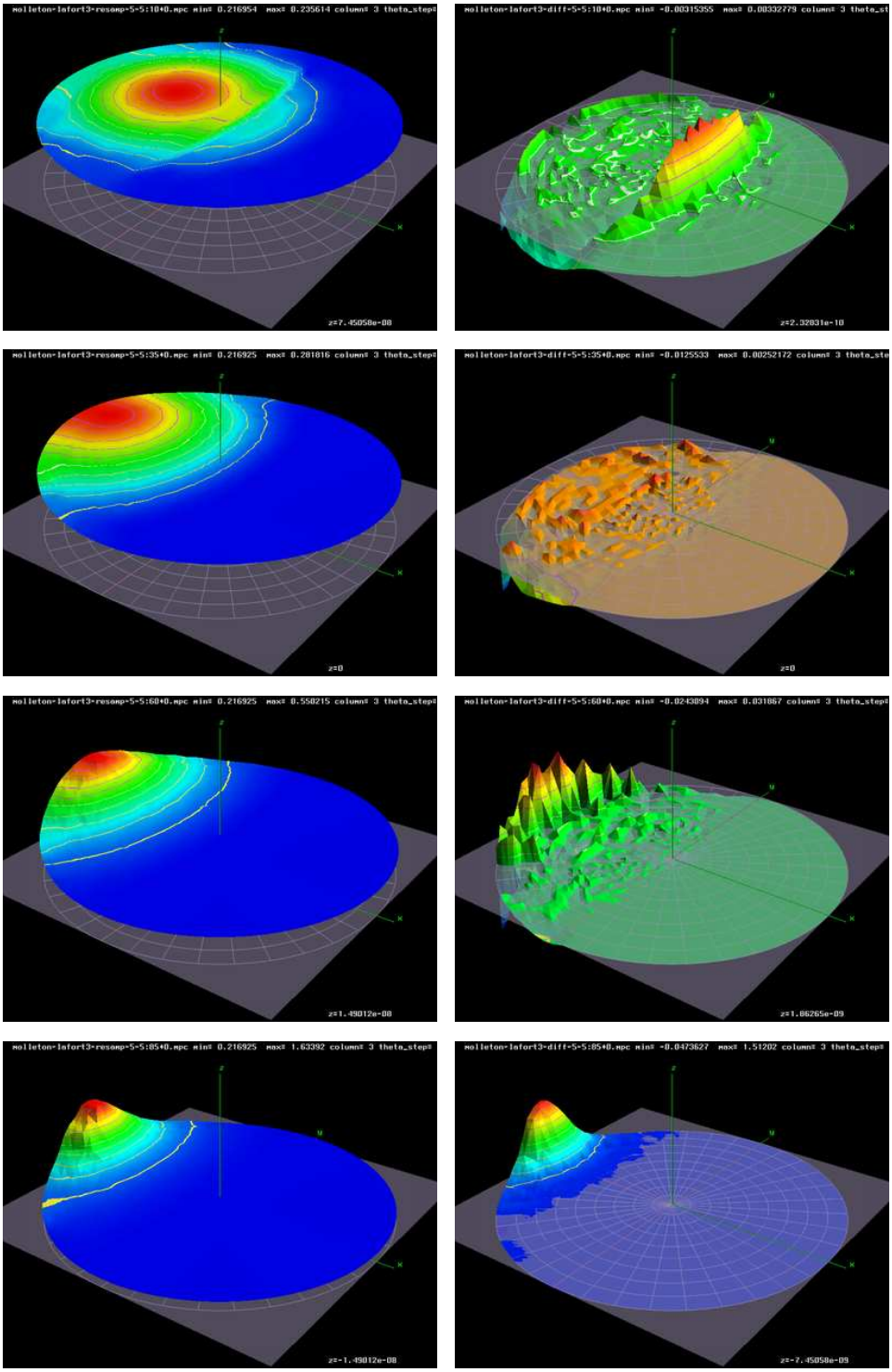


Figure E.9: Resampling bias test with 5 nearest neighbours: resampled Lafortune BRDF model fitted to molleton (left column) and difference to analytical evaluation of model (right column) for incident angles of 10°, 35°, 60°, and 85°.

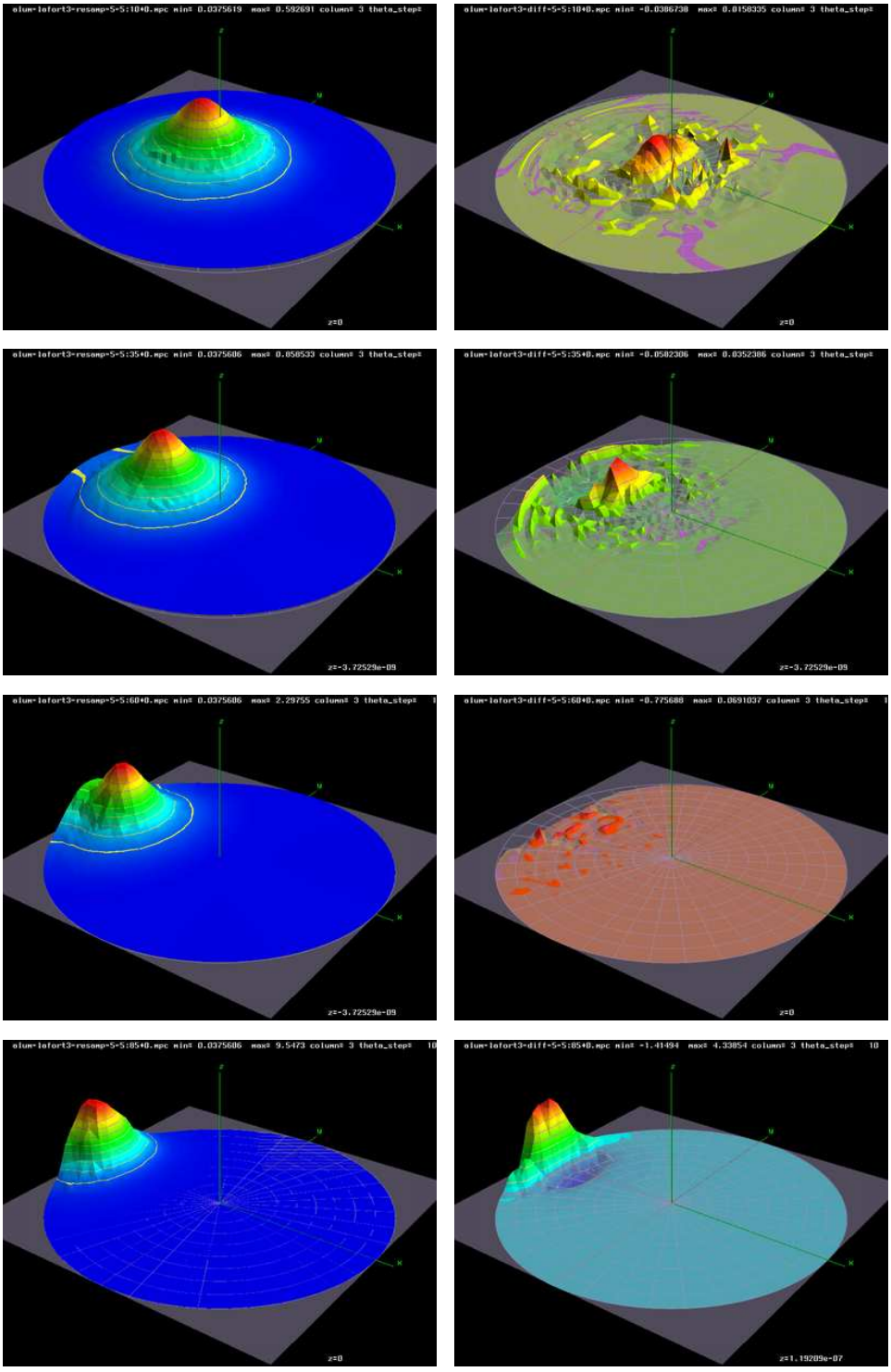


Figure E.10: Resampling bias test with 5 nearest neighbours: resampled Lafortune BRDF model fitted to aluminium (left column) and difference to analytical evaluation of model (right column) for incident angles of 10°, 35°, 60°, and 85°.

Appendix F

Experimental Validation Results

The positions of the illuminance sensors in the validation box interior and their designations are shown in figure F.1. Simulated and measured illuminance plots for each sensor are shown in figures F.2 – F.5 for compound case study 1 (interreflection), and figures F.6 – F.9 for compound case study 2 (light shelf and interreflection). Errors bars have been omitted.

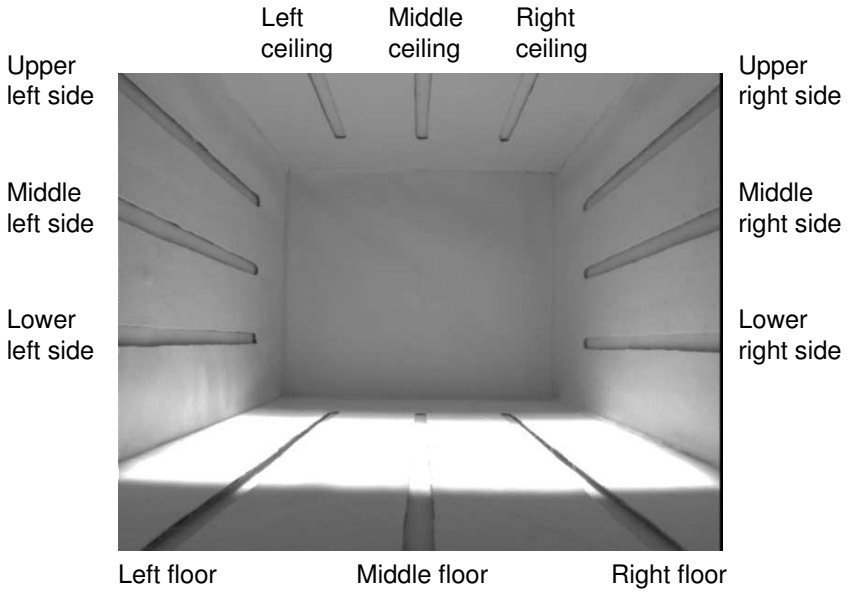


Figure F.1: Interior view of validation box with illuminance sensor positions and designations.

Table F.1 lists a breakdown of the relative errors for each sensor position. Substantial deviations only arise from the direct component for the lower side and floor sensors. These are caused by inaccurate penumbrae on behalf of RADIANCE's light source sampling which show up as stairstep aliasing artifacts in the simulated illuminance plots. Finer source subdivision (with the -ds parameter) would reduce

these deviations but introduces *aiming failures* resulting from shadow rays which fail to intersect the light source. This phenomenon arises with complex or very small and distant light source geometries and can result in considerable underprediction of the direct component, which must be avoided at all costs. Since we are primarily interested in validating the indirect illumination (the direct component is calculated identically for both RADIANCE Classic and photon map), the deviations in the affected plots should not be considered representative.

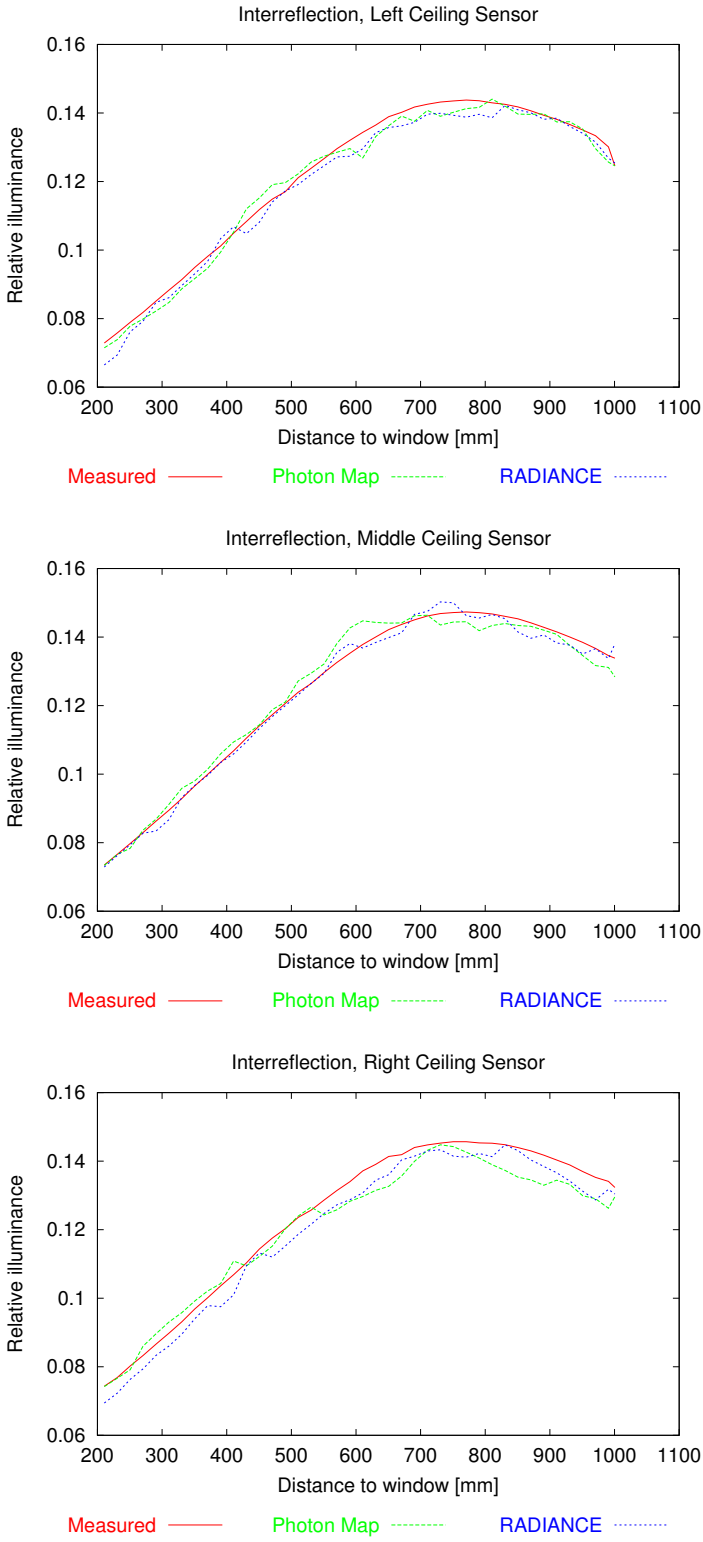


Figure F.2: Ceiling illuminance for compound case study 1 (interreflection).

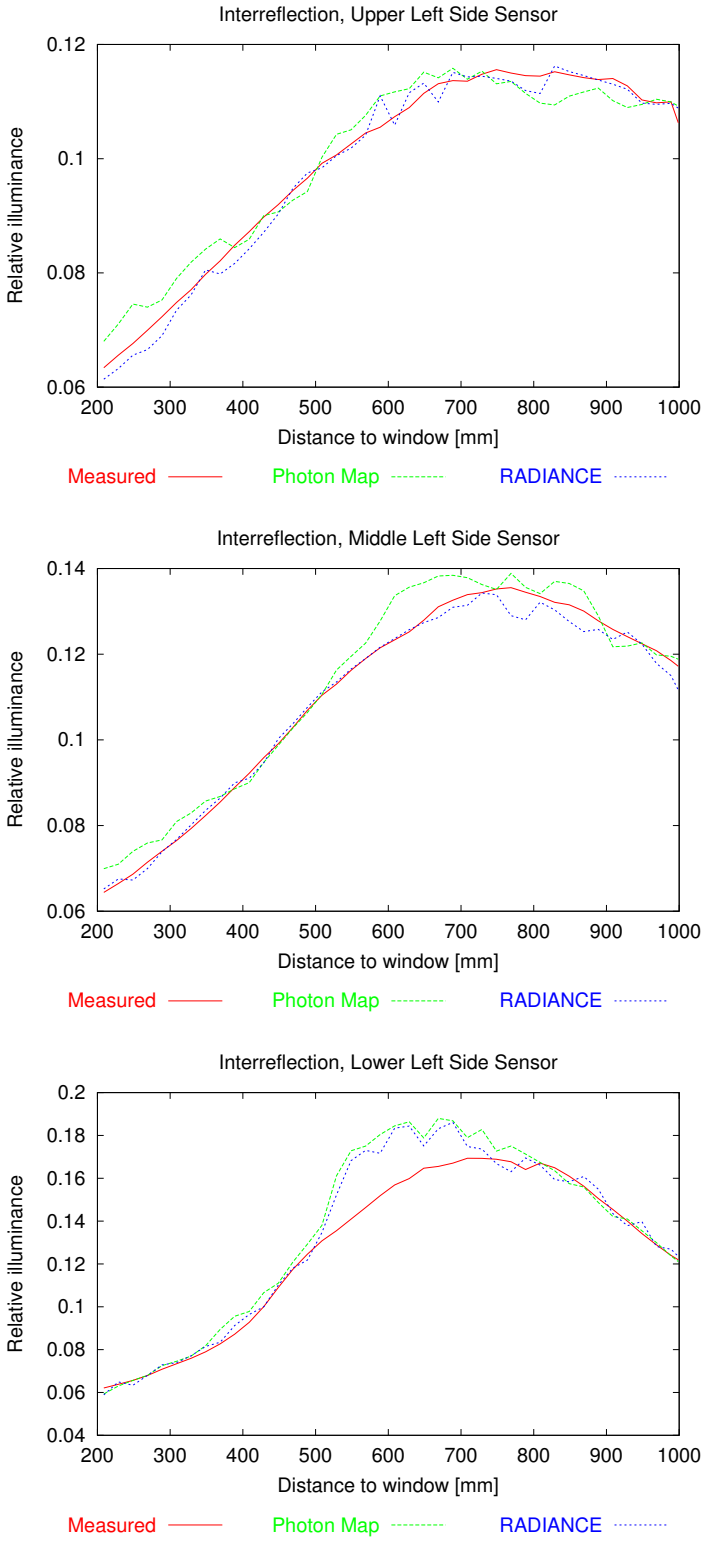


Figure F.3: Left side illuminance for compound case study 1 (interreflection).

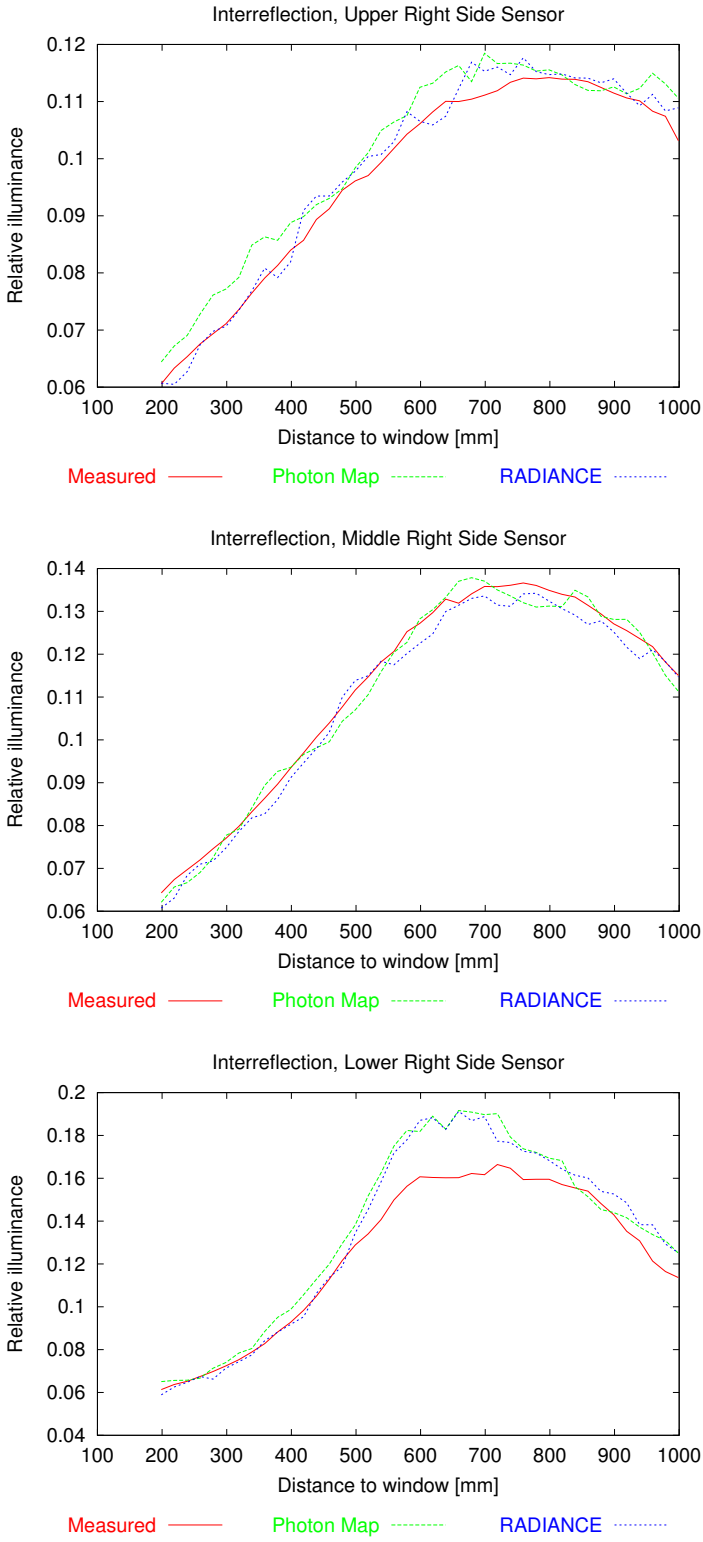


Figure F.4: Right side illuminance for compound case study 1 (interreflection).

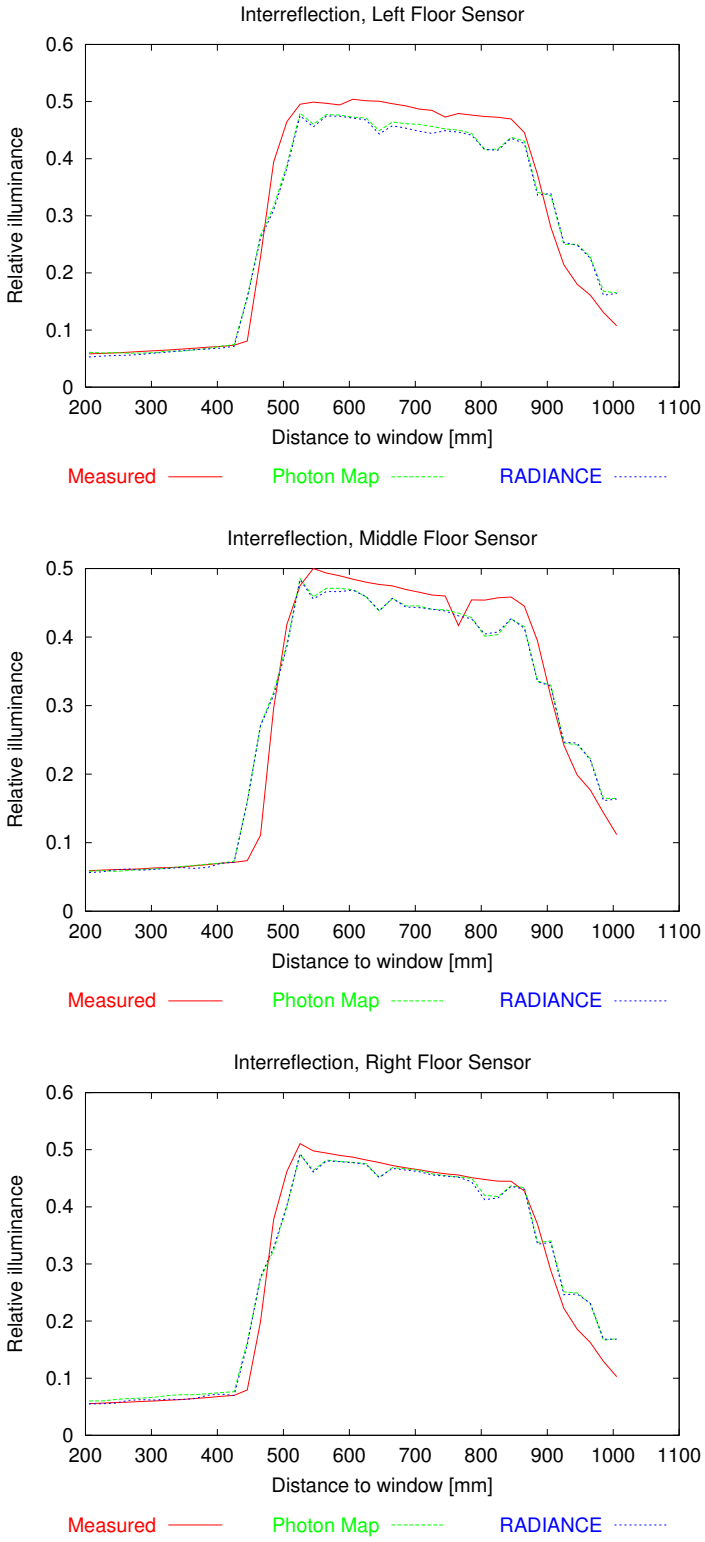


Figure F.5: Floor illuminance for compound case study 1 (interreflection).

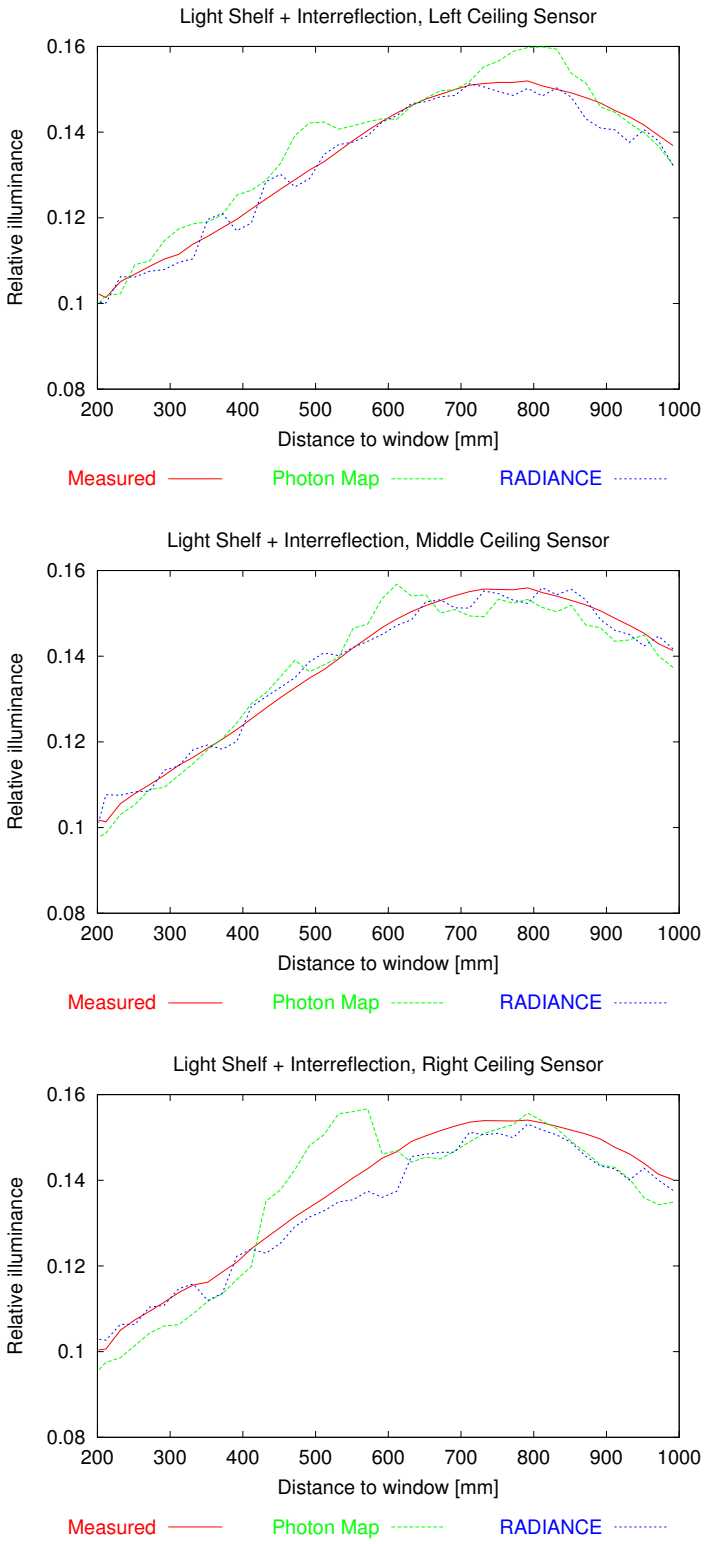


Figure F.6: Ceiling illuminance for compound case study 2 (light shelf and interreflection).

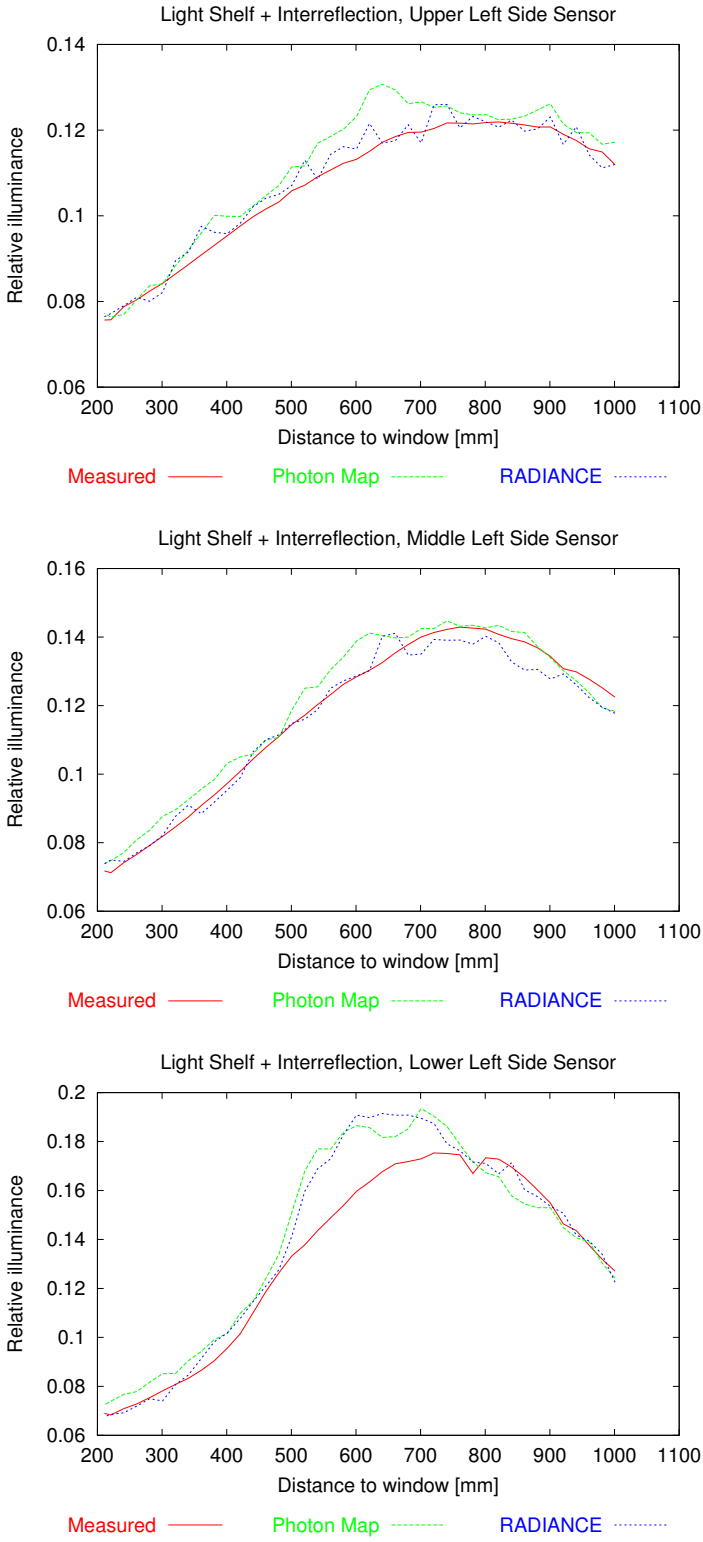


Figure F.7: Left side illuminance for compound case study 2 (light shelf and interreflection).

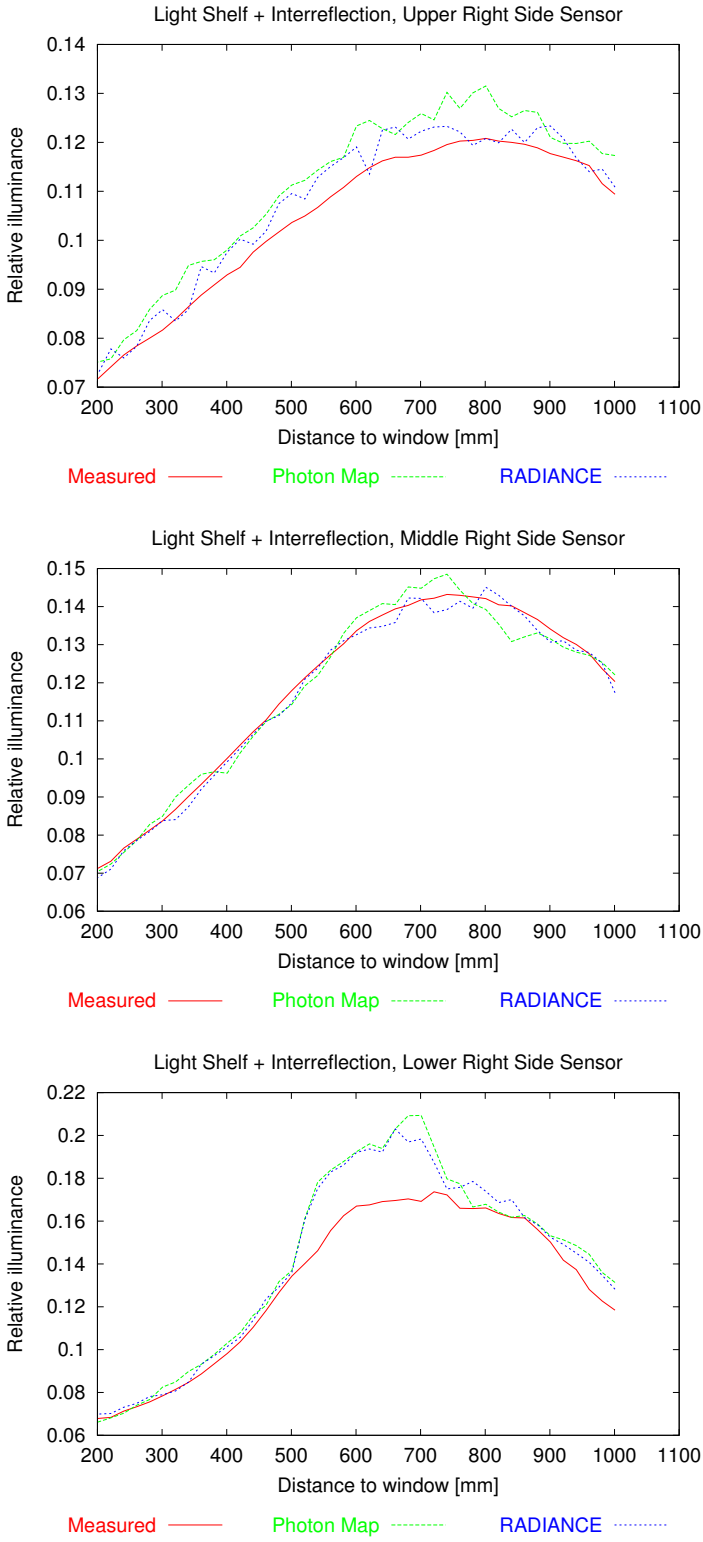


Figure F.8: Right side illuminance for compound case study 2 (light shelf and interreflection).

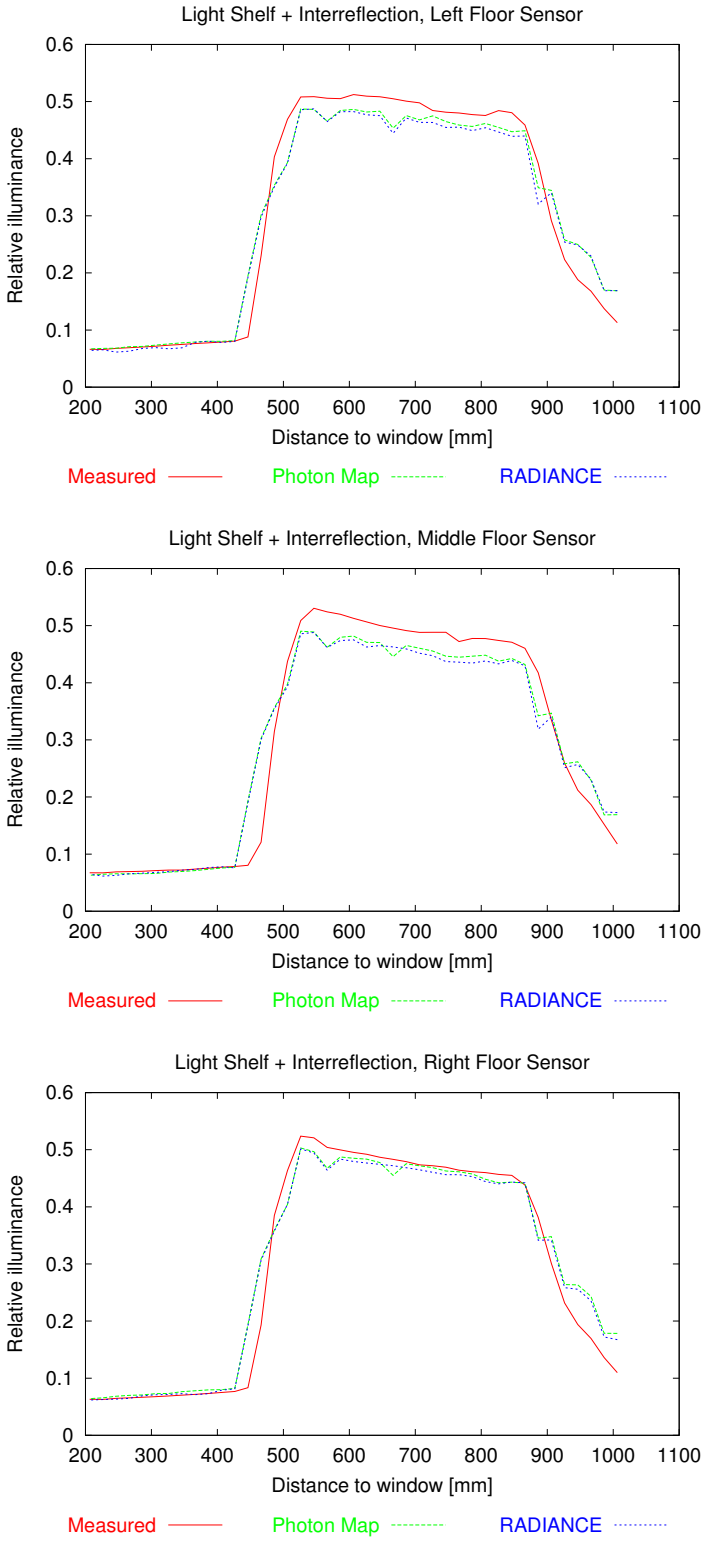


Figure F.9: Floor illuminance for compound case study 2 (light shelf and interreflection).

Sensor	Case study deviations [min/max/avg%]			
	Interreflection		Light shelf & Interreflection	
	Photon map	RADIANCE Classic	Photon map	RADIANCE Classic
Left ceiling	0.0/5.6/1.9	0.2/8.7/2.2	0.0/8.3/2.9	0.1/4.1/1.6
Middle ceiling	0.0/5.5/1.9	0.0/3.3/1.2	0.1/5.5/2.3	0.0/6.2/1.4
Right ceiling	0.0/6.2/3.1	0.0/6.6/3.2	0.1/12.6/4.6	0.0/6.3/2.4
Upper left side	0.0/10.2/3.1	0.0/5.2/1.7	0.0/12.5/4.0	0.0/7.3/2.1
Middle left side	0.1/8.6/3.2	0.0/4.8/1.4	0.2/8.3/3.5	0.1/6.0/2.5
Lower left side	0.0/22.6/5.9	0.0/19.4/4.7	1.0/23.3/7.9	0.0/19.5/5.8
Upper right side	0.3/11.0/4.4	0.2/6.1/2.1	2.2/9.8/6.1	0.0/6.5/3.1
Middle right side	0.2/4.3/2.1	0.0/6.4/2.4	0.1/6.7/2.1	0.0/3.4/1.4
Lower right side	0.4/19.5/8.3	0.0/19.3/7.0	0.1/23.7/7.7	0.1/19.8/6.9
Left floor	0.4/95.7/12.4	2.5/93.7/13.6	0.4/122.1/11.8	0.6/118.8/13.0
Middle floor	0.0/145.3/13.0	0.0/144.1/13.2	0.4/150.6/15.0	0.4/148.7/15.4
Right floor	0.2/106.1/13.3	0.0/99.6/11.1	0.4/132.9/13.3	0.2/130.4/11.8
Total	0.0/145.3/6.1	0.0/144.1/5.3	0.0/150.6/6.8	0.0/148.7/5.6

Table F.1: Breakdown of relative error for compound case studies. Deviations for each sensor are given as minimum/maximum/average percentage.

Bibliography

[AB94] Peter Apian-Bennewitz. Designing an apparatus for measuring bidirectional reflection/transmission. In *Optical Materials Technology for Energy Efficiency and Solar Energy Conversion XIII*, pages 697–706. SPIE, Vol 2255, April 1994.

[AB95] Peter Apian-Bennewitz. *Messung und Modellierung von lichtstreuenden Materialien zur Computer-Simulation von Tageslichtbeleuchtung*. PhD thesis, Albert-Ludwigs-Universität, Freiburg im Breisgau, December 1995.

[ABH98] Peter Apian-Bennewitz and Jochen von der Hardt. Enhancing and calibrating a goniophotometer. *Solar Energy Materials & Solar Cells*, 54(1-4):309–322, 1998.

[AK90] James Arvo and David Kirk. Particle transport and image synthesis. In *Proceedings Siggraph '90*, volume 24, pages 63–66, August 1990.

[Arv95] James Arvo. Applications of irradiance tensors to the simulation of non-lambertian phenomena. In *Proceedings Siggraph '95*, volume 29, pages 335–342, 1995.

[Ash95] Ian Ashdown. Near-field photometry: Measuring and modeling complex 3-D light sources. In *ACM SIGGRAPH '95 Course Notes*, pages 1–15, 1995.

[Ben75] Jon Louis Bentley. Multidimensional binary search trees for associative searching. *Communications of the ACM*, 18(9):509–517, September 1975.

[Ben79] Jon Louis Bentley. Multidimensional binary search trees in database applications. *IEEE Transactions on Software Engineering*, SE-5(4):333–340, July 1979.

[BF79] Jon Louis Bentley and Jerome H. Friedman. Data structures for range searching. *ACM Computing Surveys*, 11(4):397–409, December 1979.

[BSS94] Philippe Blasi, Bertrans Le Saëc, and Christophe Schlick. An importance driven monte carlo solution to the global illumination problem. In *Proceedings 5th Eurographics Workshop on Rendering*, 1994.

[CCWG88] Michael Cohen, Shenchang Chen, John Wallace, and Donald Greenberg. A progressive refinement approach to fast radiosity image generation. In *Proceedings Siggraph '88*, volume 22, pages 75–84, 1988.

[Chr00] Per H. Christensen. Faster photon map global illumination. *Journal of Graphics Tools*, 4(3):1–10, April 2000.

[Chr01] Per H. Christensen. Photon map tricks. In *Siggraph 2001 Course Notes 38*, August 2001.

[CIE94] CIE. Spatial distribution of daylight – luminance distributions of various reference skies. Technical Report CIE-110-1994, CIE, 1994.

[Cod93] P. D. Coddington. Analysis of random number generators using monte carlo simulation. Technical Report SCCS-526, Syracuse University, October 1993.

[CW93] Michael F. Cohen and John R. Wallace. *Radiosity and Realistic Image Synthesis*. Morgan Kaufmann, San Francisco, 1993.

[Del95] Jean-Jacques Delaunay. *Contribution á la modélisation de la lumière naturelle en vue de son application á la simulation de l'éclairage de locaux*. PhD thesis, Université Louis Pasteur, Strasbourg, 1995.

[DM01] Frederic Drago and Karol Myszkowski. Validation proposal for global illumination and rendering techniques. *Computers & Graphics*, 25(3):511–518, 2001.

[Dut90] Geoffrey Dutton. Locational properties of quaternary triangular meshes. In *Proceedings Fourth International Symposium on Spatial Data Handling*, pages 901–910, July 1990.

[Dut03] Philip Dutré. Global illumination compendium. Available at `http://www.cs.kuleuven.ac.be/~phil/GI`, September 2003.

[EGW98] Karlheinz Eder, Manuel Goller, and Jan Wienold. Tageslichtnutzung - transport von konzentriertem sonnenlicht in lichtrohren. In *Proceedings Licht '98*, September 1998.

[FBF77] J. H. Friedman, J. L. Bentley, and R. A. Finkel. An algorithm for finding best matches in logarithmic expected time. *ACM Transactions on Mathematical Software*, 3(3):209–226, September 1977.

[Fis92] A. Fisher. Tolerances in lighting design. In *Proceedings of the CIE seminar on computer programs for light and lighting*, 1992.

[GCS94] Steven Gortler, Michael Cohen, and Philipp Slusallek. Radiosity and relaxation methods. *IEEE Computer Graphics & Applications*, 14(6):48–58, November 1994.

[GGHS03] Michael Goesele, Xavier Granier, Wolfgang Heidrich, and Hans-Peter Seidel. Accurate light source acquisition and rendering. In *Proceedings SIGGRAPH 2003*, 2003.

[Gla95] Andrew S. Glassner. *Principles of Digital Image Synthesis*. Morgan Kaufmann, San Francisco, 1995.

[Gry89] Anat Grynberg. Validation of RADIANCE. Technical Report LBID 1575, Lawrence Berkeley Laboratory, July 1989.

[GTGB84] Cindy M. Goral, Kenneth E. Torrance, Donald P. Greenberg, and Bennet Battaile. Modelling the interaction of light between diffuse surfaces. In *Proceedings Siggraph '84*, volume 18, pages 213–222, 1984.

[Har95] Jochen von der Hardt. Optimierung eines messverfahrens zur untersuchung des streuverhaltens lichtlenkender bzw. -streuender elemente. Master's thesis, Albert-Ludwigs-Universität, Freiburg im Breisgau, October 1995.

[Hec90] Paul S. Heckbert. Adaptive radiosity textures for bidirectional ray tracing. In *Proceedings Siggraph '90*, volume 24, pages 145–154, August 1990.

[Hel98] Peter Hellekalek. Good random number generators are (not so) easy to find. *Mathematics and Computers in Simulation*, 46:485–505, 1998.

[HH64] J.M. Hammersley and D.C. Handscomb. *Monte Carlo Methods*. Methuen & Co., London, 1964.

[HP02] Heinrich Hey and Werner Purgathofer. Advanced radiance estimation for photon map global illumination. *Computer Graphics Forum*, 21(3):541–545, 2002.

[HSA91] Pat Hanrahan, David Salzman, and Larry Aupperle. A rapid hierarchical radiosity algorithm. In *Proceedings Siggraph '91*, volume 25, pages 197–206, 1991.

[HTSG91] X. He, Kenneth Torrance, François Sillion, and Donald Greenberg. A comprehensive physical model for light reflection. In *Proceedings Siggraph '91*, volume 25, pages 175–186, 1991.

[ICG86] David S. Immel, Michael F. Cohen, and Donald P. Greenberg. A radiosity method for non-diffuse environments. In *Proceedings Siggraph '86*, volume 20, pages 133–142, 1986.

[IES86] IES. Nomenclature and definitions for illuminating engineering. Technical Report RP-16-1986, ANSI/IES, New York, 1986.

[JC95] Henrik Wann Jensen and Niels Jørgen Christensen. Photon maps in bidirectional monte carlo ray tracing of complex objects. *Computers and Graphics*, 19(2):215–224, March 1995.

[JC98] Henrik Wann Jensen and Per H. Christensen. Efficient simulation of light transport in scenes with participating media using photon maps. In *Proceedings Siggraph '98*, volume 32, pages 311–320, July 1998.

[JC00] Henrik Wann Jensen and Niels Jørgen Christensen. A practical guide to global illumination using photon maps. In *Siggraph 2000 Course Notes 8*, July 2000.

[Jen95] Henrik Wann Jensen. Importance driven path tracing using the photon map. In P. M. Hanrahan and W. Purgathofer, editors, *Rendering Techniques '95*, pages 326–335. Springer-Verlag, Wien, 1995.

[Jen96] Henrik Wann Jensen. Global illumination using photon maps. In X. Pueyo and P. Schröder, editors, *Rendering Techniques '96*, pages 21–30. Springer-Verlag, Wien, 1996.

[Jen97] Henrik Wann Jensen. Rendering caustics on non-lambertian surfaces. *Computer Graphics Forum*, 16(1):57–64, March 1997.

[Jen01] Henrik Wann Jensen. *Realistic Image Synthesis using Photon Mapping*. A K Peters, Natick, MA, 2001.

[Kaj86] James Kajiya. The rendering equation. In *Proceedings Siggraph '86*, volume 20, pages 143–150, 1986.

[Kuc02] Thomas Kuckelkorn. *Nichtabbildende Konzentratoren für den Sonnenschutz in Gebäuden*. PhD thesis, Technische Universität München, Fakultät für Architektur, April 2002.

[KV93] R. Kittler and P. Valko. Radiance distribution on densely overcast skies: Comparison with cie luminance standard. *Solar Energy*, 51(5):349–355, 1993.

[KW86] Malvin H. Kalos and Paula A. Whitlock. *Monte Carlo Methods: Volume I: Basics*. John Wiley & Sons, New York, 1986.

[Lan91] Brigitta Lange. The simulation of radiant light transfer with stochastic ray tracing. In *Proceedings 2nd Eurographics Workshop on Rendering*, 1991.

[Lew93] R. R. Lewis. Making shaders more physically plausible. In *Proceedings 4th Eurographics Workshop on Rendering*, pages 47–62, 1993.

[LFTG97] Eric LaFortune, Sing-Choong Foo, Kenneth Torrance, and Donald Greenberg. Non-linear approximation of reflectance functions. In *Proceedings Siggraph '97*, volume 31, pages 117–126, 1997.

[LH96] Marc Levoy and Pat Hanrahan. Light field rendering. In *Proceedings SIGGRAPH '96*, pages 31–42, 1996.

[Lit90] J.P. Littlefair. Predicting annual lighting use in daylit buildings. *Building & Environment*, 25(1):43–54, 1990.

[LP88] Thomas A. Leonard and Michael Pantoliano. BRDF round robin. In *Stray Light and Contamination in Optical Systems*, volume 967 of *SPIE*. Society of Photo-Optical Instrumentation Engineers, 1988.

[LPR89] Thomas A. Leonard, Michael Pantoliano, and James Reilly. Results of a CO_2 BRDF round robin. In *Scatter from Optical Components*, volume 1165 of *SPIE*. Society of Photo-Optical Instrumentation Engineers, 1989.

[LRP97] Gregory Ward Larson, Holly Rushmeier, and Christine Piatko. A visibility tone reproduction operator for high dynamic range scenes. Technical Report LBNL 39882, Lawrence Berkeley National Laboratory, January 1997.

[LS98] Greg Ward Larson and Rob Shakespeare. *Rendering with RADIANCE*. Morgan Kaufmann, San Francisco, 1998.

[Mar99] John Mardaljevic. *Daylight Simulation: Validation, Sky Models, and Daylight Coefficients*. PhD thesis, De Montfort University, Leicester, December 1999.

[MF03] F. Maamari and M. Fontoynont. Analytical tests for investigating the accuracy of lighting programs. *Lighting Research and Technology*, 35(3):225–242, September 2003.

[MLH93] J. Mardaljevic, K. Lomas, and D. Henderson. Advanced daylighting design for complex spaces. In *Proceedings CLIMA 2000*, November 1993.

[Mys97] Karol Myszkowski. Lighting reconstruction using fast and adaptive density estimation techniques. In *Rendering Techniques '97*, pages 251–262. Springer-Verlag, Wien, 1997.

[NM65] J.A. Nelder and R. Mead. A simplex method for function minimization. *Computer Journal*, 7:308–313, 1965.

[NRH+77] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis. Geometrical considerations and nomenclature for reflectance. Technical Report 160, National Bureau of Standards, October 1977.

[PBB+94] R. Perez, H.G. Beyer, A. Bittar, G. Czeplak, L. Dahlgren, J.J. Delaunay, D. Dumortier, M. Fontoynont, P. Ineichen, J. Love, B. McArthur, J.A. Olseth, M. Rommel, and A. Skartveit. Broad-band visible radiation data acquisition and analysis. Technical Report IEA-SHCP-17E-1, International Energy Agency, August 1994.

[PP98] Ingmar Peter and Georg Pietrek. Importance driven construction of photon maps. In *Rendering Techniques '98*, pages 269–280. Springer-Verlag, Wien, 1998.

[PSM93] R. Perez, R. Seals, and R. J. Michalsky. All-weather model for sky luminance distribution – preliminary configuration and validation. *Solar Energy*, 50(3):235–245, 1993.

[PSS99] A. J. Preetham, Peter Shirley, and Brian Smits. A practical analytic model for daylight. In *Proceedings Siggraph '99*, volume 33, pages 91–100, 1999.

[PTVF92] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992.

[RT90] Holly E. Rushmeier and Kenneth E. Torrance. Extending the radiosity method to include specularly reflecting and translucent materials. *ACM Transactions on Graphics*, 9(1):1–27, January 1990.

[Rub81] Reuven Y. Rubinstein. *Simulation and the Monte Carlo Method*. John Wiley & Sons, 1981.

[Rus97] Szymon Rusinkiewicz. A survey of BRDF representation for computer graphics. Technical report, Standford University, 1997.

[RW01] Christoph F. Reinhart and Oliver Walkenhorst. Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds. *Energy and Buildings*, 33(7):683–697, 2001.

[SAG94] Brian Smits, James Arvo, and Donald Greenberg. A clustering algorithm for radiosity in complex environments. In *Proceedings Siggraph '94*, volume 28, pages 435–442, 1994.

[SAWG91] François X. Sillion, James R. Arvo, Stephen H. Westin, and Donald P. Greenberg. A global illumination solution for general reflectance distributions. In *Proceedings Siggraph '91*, volume 25, pages 187–196, 1991.

[Sed92] Robert Sedgewick. *Algorithms in C++*. Addison-Wesley, Reading, MA, 1992.

[Sil86] Bernard W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London / New York, 1986.

[SJ00] Brian Smits and Henrik Wann Jensen. Global illumination test scenes. Technical Report UUCS-00-013, University of Utah, 2000.

[SKKA01] L. Szirmay-Kalos, L. Kovacs., and A. M. Abbas. Testing monte-carlo global illumination methods with analytically computable scenes. In *Ninth International Conference in Central Europe on Computer Graphics, Visualization and Interactive Digital Media*, 2001.

[SP89] François Sillion and Claude Puech. A general two-pass method for integrating specular and diffuse reflection. In *Proceedings Siggraph '89*, volume 23, pages 335–344, 1989.

[SS95] Peter Schröder and Wim Sweldens. Spherical wavelets: Efficiently representing functions on the sphere. In *Proceedings Siggraph '95*, volume 29, pages 161–172, 1995.

[Suy02] Frank Suykens. *On Robust Monte Carlo Algorithms for Multi-Pass Global Illumination*. PhD thesis, Katholieke Universiteit Leuven, September 2002.

[SW00] Frank Suykens and Yves D. Willems. Density control for photon maps. In *Rendering Techniques 2000*, pages 23–34. Springer-Verlag, Wien, 2000.

[SWH+95] Peter Shirley, Bretton Wade, Philip Hubbard, David Zareski, Bruce Walter, and Donald Greenberg. Global illumination via density estimation. In *Proceedings 6th Eurographics Workshop on Rendering*, pages 187–199, June 1995.

[TS65] Kenneth. E. Torrance and E. M. Sparrow. Off-specular peaks in the directional distribution of reflected thermal radiation. *Transactions of the ASME*, pages 1–8, November 1965.

[TS67] Kenneth. E. Torrance and E. M. Sparrow. Theory of off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, pages 1105–1114, September 1967.

[Var00] E. Vartiainen. *Daylight Modeling and Optimization of Solar Facades*. PhD thesis, Helsinki University of Technology, 2000.

[VG97] Eric Veach and Leonidas J. Guibas. Metropolis light transport. In *Proceedings Siggraph '97*, volume 31, pages 65–76, 1997.

[Wal98] Bruce Walter. *Density Estimation Techniques for Global Illumination*. PhD thesis, Cornell University, August 1998.

[War91a] Greg Ward. Adaptive shadow testing for ray tracing. In *Proceedings 2nd Eurographics Workshop on Rendering*, 1991.

[War91b] Greg Ward. Real pixels. In James Arvo, editor, *Graphics Gems II*, pages 80–83. Academic Press, 1991.

[War92] Greg Ward. Measuring and modeling anisotropic reflection. In *Proceedings Siggraph '92*, volume 26, pages 265–272, 1992.

[War94] Greg Ward. The RADIANCE lighting simulation and rendering system. In *Proceedings Siggraph '94*, volume 28, pages 459–472, 1994.

[WAT92] Stephen Westin, James Arvo, and Kenneth Torrance. Predicting reflectance functions from complex surfaces. In *Proceedings Siggraph '92*, volume 26, pages 255–264, 1992.

[WCG87] John R. Wallace, Michael F. Cohen, and Donald P. Greenberg. A two-pass solution to the rendering equation: A synthesis of ray tracing and radiosity methods. In *Proceedings Siggraph '87*, volume 21, pages 311–320, 1987.

[WH92] Greg Ward and Paul Heckbert. Irradiance gradients. In *Proceedings 3rd Eurographics Workshop on Rendering*, 1992.

[Whi80] Turner Whitted. An improved illumination model for shaded display. *Communications of the ACM*, 23(6):343–349, June 1980.

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

[WJ95] M. P. Wand and M. C. Jones. *Kernel Smoothing*. Chapman and Hall, 1995.

[WRC88] Greg Ward, Francis Rubinstein, and Robert Clear. A ray tracing solution for diffuse interreflection. In *Proceedings Siggraph '88*, volume 22, pages 85–92, 1988.

[WW78] W.T. Welford and R. Winston. *The Optics of Nonimaging Concentrators*. Academic Press, 1978.