

Linköping studies in science and technology.  
Dissertation, No. 1717

PHYSICALLY BASED RENDERING OF SYNTHETIC OBJECTS  
IN REAL ENVIRONMENTS

Joel Kronander



Division of Media and Information Technology  
Department of Science and Technology  
Linköping University, SE-601 74 Norrköping, Sweden  
Norrköping, December 2015

### **Description of the cover image**

Rendering of virtual furniture illuminated with spatially varying real world illumination captured in a physical environment at IKEA Communications AB.

The virtual scene was modeled by Sören and Per Larsson.

### **Physically based rendering of synthetic objects in real environments**

Copyright © 2015 Joel Kronander (unless otherwise noted)

*Division of Media and Information Technology  
Department of Science and Technology  
Campus Norrköping, Linköping University  
SE-601 74 Norrköping, Sweden*

ISBN: 978-91-7685-912-4

ISSN: 0345-7524

Printed in Sweden by LiU-Tryck, Linköping, 2015





# Abstract

This thesis presents methods for photorealistic rendering of virtual objects so that they can be seamlessly composited into images of the real world. To generate predictable and consistent results, we study physically based methods, which simulate how light propagates in a mathematical model of the augmented scene. This computationally challenging problem demands both efficient and accurate simulation of the light transport in the scene, as well as detailed modeling of the geometries, illumination conditions, and material properties. In this thesis, we discuss and formulate the challenges inherent in these steps and present several methods to make the process more efficient.

In particular, the material contained in this thesis addresses four closely related areas: HDR imaging, IBL, reflectance modeling, and efficient rendering. The thesis presents a new, statistically motivated algorithm for HDR reconstruction from raw camera data combining demosaicing, denoising, and HDR fusion in a single processing operation. The thesis also presents practical and robust methods for rendering with spatially and temporally varying illumination conditions captured using omnidirectional HDR video. Furthermore, two new parametric BRDF models are proposed for surfaces exhibiting wide angle gloss. Finally, the thesis also presents a physically based light transport algorithm based on Markov Chain Monte Carlo methods that allows approximations to be used in place of exact quantities, while still converging to the exact result. As illustrated in the thesis, the proposed algorithm enables efficient rendering of scenes with glossy transfer and heterogenous participating media.



# Populärvetenskaplig Sammanfattning

En av de största utmaningarna inom datorgrafik är att syntetisera, eller *rendera*, fotorealistiska bilder. Fotorealistic rendering används idag inom många tillämpningsområden såsom specialeffekter i film, datorspel, produktvisualisering och virtuell verklighet. I många praktiska tillämpningar av fotorealistic rendering är det viktigt att kunna placera in virtuella objekt i fotografier, så att de virtuella objekten ser verkliga ut. IKEA-katalogen, till exempel, produceras i många olika versioner för att passa olika länder och regioner. Grunden till de flesta bilderna i katalogen är oftast densamma, men symboler och standardmått på möbler varierar ofta för olika versioner av katalogen. Istället för att fotografera varje version separat kan man använda ett grundfotografi och lägga in olika virtuella objekt såsom möbler i fotot. Genom att på det här sättet möblera ett rum virtuellt, istället för på riktigt, kan man också snabbt testa olika möbleringar och därmed göra ekonomiska besparingar.

Den här avhandlingen bidrar med metoder och algoritmer för att rendera fotorealistic bilder av virtuella objekt som kan blandas med verkliga fotografier. För att rendera sådana bilder används fysikaliskt baserade simuleringar av hur ljus interagerar med virtuella och verkliga objekt i motivet. För fotorealistic resultat kräver simuleringarna noggrann modellering av objektens geometri, belysning och materialegenskaper, såsom färg, textur och reflektans.

För att de virtuella objekten ska se verkliga ut är det viktigt att belysa dem med samma ljus som de skulle ha haft om de var en del av den verkliga miljön. Därför är det viktigt att noggrant mäta och modellera ljusförhållanden på de platser i scenen där de virtuella objekten ska placeras. För detta använder vi *High Dynamic Range*-fotografi, eller HDR. Med hjälp av HDR-fotografi kan vi noggrant mäta hela omfånget av det infallande ljuset i en punkt, från mörka skuggor till direkta ljuskällor. Detta är inte möjligt med traditionella digitalkameror, då det dynamiska omfånget hos vanliga kameran sensorer är begränsat. Avhandlingen beskriver nya metoder för att rekonstruera HDR-bilder som ger mindre brus och artefakter än tidigare metoder. Vi presenterar också metoder för att rendera virtuella objekt som rör sig mellan regioner med olika belysning, eller där belysningen varierar i tiden. Metoder för att representera spatiellt varierande belysning på ett kompakt sätt presenteras också. För att noggrant beskriva hur glansiga ytor sprider eller reflekterar ljus, beskrivs också två nya parametriska modeller som är mer verklighetstroga än tidigare reflektionsmodeller. I avhandlingen presenteras också en ny metod för effektiv rendering av motiv som är mycket beräkningskrävande, till exempel scener med uppmätta

belysningsförhållanden, komplicerade material, och volimetriska modeller som rök, moln, textilier, biologisk vävnad och vätskor. Metoden bygger på en typ av så kallade Markov Chain Monte Carlo metoder för att simulera ljustransporten i scenen, och är inspirerad av nyligen presenterade resultat inom matematisk statistik.

Metoderna som beskrivs i avhandlingen presenteras i kontexten av fotorealistic rendering av virtuella objekt i riktiga miljöer, då majoriteten av forskningen utförts inom detta område. Flera av de metoder som presenteras i denna avhandling är dock tillämpbara inom andra domäner, såsom fysiksimulering, datorseende och vetenskaplig visualisering.



# Acknowledgments

During my years as a PhD student I have had the fortune to work with some amazing people. This has made my PhD studies a very enjoyable time! I would like to thank all the people that have made my PhD studies so much fun and have helped me in some way.

First of all, I would like to thank my thesis advisor, Jonas Unger, for his support and guidance over these years. Your ideas and enthusiasm have made working on this thesis a lot more fun than it would have been without you. It has been a great privilege to work with you, and I hope that our collaboration can continue in the future. I would also like to express my genuine gratitude towards my assistant supervisor Anders Ynnerman, who introduced me to the field of visualization and computer graphics. Thank you for all of our discussions and the guidance that you have given me.

Next, I would like to thank all of my colleagues in the computer graphics group. I have truly enjoyed working with all of you! In particular I would like to thank Per Larsson, for all the late nights working toward deadlines, all the nice renderings and making things work in practice, Stefan Gustavson for inspiring me to "think for myself", Reiner Lenz for many inspiring and interesting discussions, Andrew Gardner for being "the man", Gabriel Eilertsen for all of those hilarious puns, Ehsan Miandji; may your sub-identities never turn into Bernoulli matrices, Saghi Hajisharif for all the good collaborations and bringing a smile to the lab, and finally, Apostolia Tsirikoglou and Tanaboon Tongbuasirilai for all the fun and interesting collaborations! Thank you all for making it so motivating and fun to work with you!

During these years, I have also had the opportunity to work with some extraordinary researchers in external collaborations. Thomas Schön introduced me to machine learning, computational statistics, and Monte Carlo methods. Thank you for all of our inspiring discussions and motivating collaborations over the years! I would also like to thank Francesco Banterle for fun and interesting collaborations on Image Based Lighting, Johan Dahlin for many inspiring discussions and interesting research collaborations on Monte Carlo methods, and Torsten Möller for interesting collaborations on volume rendering and giving me the opportunity to visit his group in Vancouver. Thank you all for inspiring discussions and motivating collaborations!

I would also like to thank all other current and former colleagues and friends working at the Visualization center in Norrköping and at the Media and Information Technology division. Thank you for great research discussions, lots of fun, and all the coffee breaks! In particular, I would like to thank all of you whom I have had the pleasure to co-author papers with over the years,

Joakim Löw, Daniel Jönsson, Timo Ropinski, Stefan Lindholm and Patric Ljung, among others. I would also like to thank Eva Skärblom for all the help she has provided in practical and administrative matters.

This thesis has been proofread by several people, including Jonas Unger, Anders Ynnerman, Andrew Gardner, Thomas Schön, Reiner Lenz, Johan Dahlin, Ehsan Miandji, Gabriel Eilertsen, Saghi Hajisharif, Per Larsson, Martin Falk, Amanda Jonsson, and Tanaboon Tongbuasirilai. Your help has significantly improved this thesis. Thank you for your comments!

I would also like to thank all my friends and family, who have made my life outside of work so much more rich and fulfilling over these years. Thank you for understanding my physical (and mental!) absence during the many intense periods of work and deadlines leading up to this thesis. Amanda, you make my life complete! Thank you for all your love and support in times of doubt!

Finally I would also like to thank you, the reader, for taking your time to read this thesis! I hope that it in some way can inspire you to achieve great things, far surpassing the content of this thesis!

# Contents

<b>Abstract</b>	<b>v</b>
<b>Populärvetenskaplig Sammanfattning</b>	<b>vii</b>
<b>Acknowledgments</b>	<b>ix</b>
<b>I Background</b>	
<b>1 Introduction</b>	<b>1</b>
1.1 Towards virtual photo sets	1
1.2 Photorealistic rendering	3
1.3 Contributions	7
1.4 Publications	8
1.5 Thesis outline	10
1.5.1 Outline of part I	10
1.5.2 Outline of part II	11
<b>2 Fundamentals of Light transport</b>	<b>17</b>
2.1 Light transport model	17
2.2 Radiometry	18
2.2.1 Domains and measures	19
2.2.2 Radiometric quantities	20
2.3 Rendering equation	21
2.4 Radiative transfer equation	22
2.5 Path integral formulation	26
2.6 Simulating light transport	28
<b>3 Monte Carlo rendering</b>	<b>29</b>
3.1 Monte Carlo estimators	30
3.1.1 Constructing estimates from random samples	30
3.1.2 Importance Sampling	32
3.1.3 Independent Sampling methods	33
3.1.4 Markov chain Monte Carlo methods	36
3.2 Monte Carlo light transport simulation	38
3.2.1 Path tracing methods	38
3.2.2 Caching methods	41
3.2.3 Metropolis Light Transport	41
3.3 Contributions	48

3.3.1	Pseudo-marginal Metropolis Light Transport	48
3.3.2	Rendering heterogeneous media using MLT	50
3.4	Summary and future work	53
<b>4</b>	<b>Image Based Lighting</b>	<b>55</b>
4.1	Light transport in mixed reality scenes	56
4.1.1	Differential rendering	56
4.2	Traditional IBL	58
4.3	Video based lighting	60
4.3.1	Previous approaches	60
4.3.2	Contributions	61
4.4	Rendering with spatially varying illumination	62
4.4.1	Previous approaches	63
4.4.2	Contributions	64
4.5	Summary and future work	67
<b>5</b>	<b>HDR Imaging</b>	<b>69</b>
5.1	State-of-the-art HDR capture	70
5.1.1	Digital cameras and raw sensor data	70
5.1.2	Dynamic range	72
5.1.3	Exposure bracketing	73
5.1.4	Single Shot Techniques	73
5.1.5	Previous methods for HDR reconstruction	77
5.2	Contributions	78
5.2.1	Radiometric camera model	79
5.2.2	Unified reconstruction	80
5.2.3	Local polynomial model	82
5.2.4	Maximum localized likelihood fitting	84
5.2.5	Adapting the filtering window	85
5.2.6	Practical results and comparisons	88
5.3	Summary and Future work	91
<b>6</b>	<b>Surface Reflectance Models</b>	<b>93</b>
6.1	BRDF acquisition and representation	94
6.1.1	Theoretical foundations	94
6.1.2	Parameterizations and symmetry properties	96
6.1.3	Acquisition	97
6.1.4	BRDF Models	99
6.2	Contributions	102
6.2.1	Observations from measured data	103
6.2.2	New parametric BRDF models for glossy reflectance.	105
6.2.3	Anisotropic model	107

## Contents

---

6.3 Summary and Future work	108
<b>7 Concluding remarks</b>	<b>113</b>
7.1 Summary of contributions	113
7.2 Future work	114
<b>Bibliography</b>	<b>117</b>
<b>II Publications</b>	
<b>Paper A</b>	<b>139</b>
<b>Paper B</b>	<b>165</b>
<b>Paper C</b>	<b>175</b>
<b>Paper D</b>	<b>183</b>
<b>Paper E</b>	<b>199</b>
<b>Paper F</b>	<b>211</b>
<b>Paper G</b>	<b>227</b>
<b>Paper H</b>	<b>249</b>



**Part I**  
**Background**





# Introduction

A longstanding goal of computer graphics is to synthesize, or *render*, images on a computer that are indistinguishable from real photographs. *Photorealistic rendering* has found many applications over the last decades and is today a key component in the entertainment industry's use of visual effects, as well as for computer aided design, product visualization, and virtual reality. An enabling factor driving these developments is the increased attention to physically accurate simulation of light propagation in elaborate mathematical models of our world. In this thesis, we use such *physically based* rendering methods to synthesize images of virtual objects so that they can be seamlessly composited into photographs of the real world. This computationally challenging problem demands both efficient and accurate simulation of the light transport between the virtual and real objects, as well as detailed modeling of the geometries, illumination conditions, and material properties in the scene. In this introductory chapter, we motivate and formulate the challenges found in each of these steps, and discuss the contributions presented in this thesis.

## 1.1 Towards virtual photo sets

An example application that illustrates the advantages of using photorealistic rendering is large scale photo production used for product catalogues, web stores, and other media. Traditionally, this process relies on the construction and maintenance of numerous physical photo sets. Figure 1.1 shows an example of a real photo set constructed at IKEA Communications AB, the creators of the most widely distributed print publication in the world - the IKEA Catalogue.



Figure 1.1: A physical photo set constructed at IKEA Communitons AB, the creators of the most widely distributed print publication in the world - the IKEA Catalogue.

The catalogue is printed in more than sixty different version and in more than forty regions in the world<sup>1</sup>, where each region has individually designed details such as choice of color schemes, regional symbols, and placement of furniture. There are also often differences in standard measures of, for example, stoves, sinks, and refrigerators. This makes it necessary to keep the physical sets over a long period of time and change them according to artistic and standard requirements. In many cases it is also necessary to rebuild entire sets weeks after they have been disassembled in order to reshoot certain details. The potential cost savings and design freedom obtained when avoiding the construction of physical photo sets have led to a rapidly increasing use of computer graphics in these production environments. Instead of using completely virtual scenes, which require tedious and elaborate modeling of the complete photo set, it is often desirable to mix virtual objects with entire real photo sets or parts of real sets. This is also preferable as interior designers and traditional photographers are accustomed to working with physical scenes, and not with relatively complicated 3D modeling software. Furthermore, photo sets are often located outside the studio, for example in someones home. As the time available at such locations is usually limited, exact modeling of the real scene is often impractical as it requires accurate modeling of the reflectance properties of materials in the scene and a detailed geometric model.

The methods discussed in this thesis contribute towards rapid photorealistic rendering of virtual objects that can be seamlessly placed into real photo set; figure 1.2b shows an example taken from paper D included in this thesis.

---

<sup>1</sup> Numbers given for the 2013 version. In total, approximately 208 million copies of the IKEA catalog were printed in 2013, more than double the number of Bibles estimated to have been printed in the same year.



(a) Photograph of the real set      (b) Virtual furniture placed in the real set

Figure 1.2: Virtual photo sets provide a flexible alternative to traditional photo sets by allowing virtual objects to be seamlessly integrated into existing environments. **a)** Photograph of the physical photo set, **b)** Rendering of virtual furniture composited into the photograph of the real set shown in **a**. The example is taken from paper D included in this thesis.

## 1.2 Photorealistic rendering

Physically based approaches to rendering have become practical both due to the widespread availability of sufficient computational power and advances in rendering algorithms. Accurate simulation of the physical processes underlying visual phenomena enables not only increased realism but also provides predictable and consistent results. Since the pioneering work during the late 60's, 70's and 80's [13, 30, 72, 101, 224], the capabilities and efficiency of light transport simulation algorithms have evolved dramatically, with increasingly impressive results. The increasing realism of computer generated images has, for example, enabled wide-spread adoption of these techniques for generating visual effects in movies, where it is now often difficult or impossible to distinguish real from simulated results. However, despite this rapid progress, rendering photorealistic images is still a complex task. Even for simple scenes, most physically based light transport algorithms require extensive processing

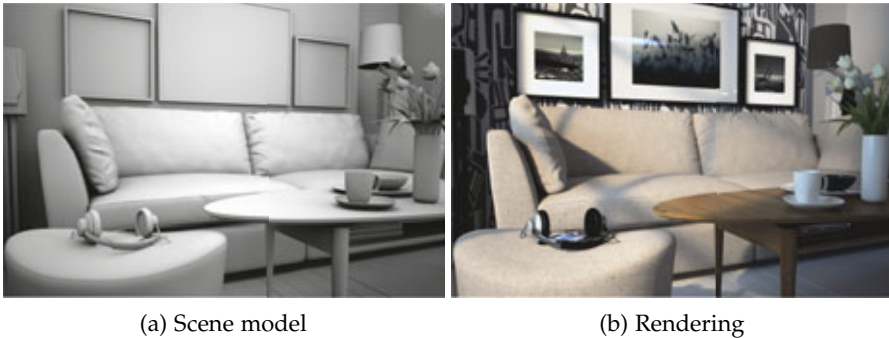


Figure 1.3: Photorealistic rendering requires not only accurate simulation of the propagation of light in the scene, but also detailed models of the scene geometry, surface reflectance and illumination.

power, limiting their use in real time applications. For complex scenes, containing, for example, caustics, heterogeneous participating media, and glossy materials, rendering a single frame can easily take hours, or days, on high-end computer hardware. This not only requires costly computer resources but also impedes the use of physically based rendering in applications such as interactive design and virtual reality.

Another challenging aspect of photorealistic rendering is that of obtaining accurate models of scene geometry, illumination, and surface reflection. The results of even a perfect light transport simulation is only as accurate as the input model permits. Direct measurement of visual attributes such as illumination [46] and surface reflectance [220] in the real world is one of the most accurate ways to obtain high-quality models. The geometry of objects in real scenes can also be captured using, for example, range scanning techniques [132] and image based modeling approaches [50]. Basing the modeled scene on measurements from real objects is one of the key factors that has enabled the rapid transition towards true photorealism in computer graphics renderings. However, detailed measurement of visual attributes is a time consuming endeavor, where there is often a clear trade-off between accuracy, data size, and capture time. Editing, and effective storage of direct measurements can also be challenging, as this often requires other representations of the captured data. For example, to efficiently represent reflectance measurements, parametric models [15, 27, 202, 220] are typically fitted to the data. These should not only be accurate but also provide intuitive parameters, such as diffuse color, specularity, and glossiness.

An important application of photorealistic image synthesis is to render virtual objects into photographs and videos of real world scenes. For consistent results, it is not only required to accurately simulate the light transport among the virtual objects, but also to model and simulate the interactions between the

virtual and real objects. While a completely digital model of the real scene would enable synthetic rendering to be used in place of the real photograph, this requires extremely elaborate modeling of the objects in the real scene. For complex scenes, accurate renderings also demand large amounts of computing power. Furthermore, in many applications, for example on a movie set, it is often desirable to work with the design of a scene in the real world directly and not with the virtual models. Instead of generating a complete digital model of the real scene, the incident illumination from the real scene onto the virtual objects can be measured and used during rendering [46]. This technique is referred to as *Image Based Lighting* (IBL), and is a key component in modern photorealistic rendering.

A central aspect of IBL is the capability to perform accurate measurement of the the incident lighting in the scene. In particular, in order to use the captured illumination information for physically based rendering, it is necessary to use radiometrically calibrated measurements capturing the full range of intensities in the scene illumination, from direct sunlight to dark shadows. Traditional digital cameras are limited to capturing around 12-14 bits per pixel, only capable of representing a ratio in the order of 10,000 : 1 between the largest and smallest distinguishable value, this ratio is often referred to as the *dynamic range* of the sensor. The limited dynamic range of traditional digital photography has led to the development of *High Dynamic Range* (HDR) imaging, which is a set of techniques for capturing and representing the full dynamic range of the illumination in a scene using radiometrically calibrated linear-response measurements. After almost two decades of intensive research over the last years, HDR imaging has been adopted in almost all fields of digital imaging. Today, many consumer-level cameras offer a HDR image capture mode that offers more dynamic range than traditional photographs. An example showing the difference between a HDR image and a traditional photograph is shown in figure 1.4.

The most widespread method for capturing HDR images today is based on fusing photographs captured with different exposure settings [48, 73, 143]. These techniques work well for static scenes, in particular if a tripod or similar is used to stabilize the camera between the shots. However, when capturing dynamic scenes, in particular HDR video, using these techniques is difficult, as robust registration of the individual exposures is necessary to reduce ghosting artifacts, and motion blur artifacts can appear if not corrected for [191]. This has led to an ongoing development of more robust HDR imaging techniques that can handle dynamic scene motion and are suitable for capturing HDR video.

The use of IBL and HDR imaging has become a common practice for major special effects studios focusing on movie and television productions [28]. However, in practice the illumination environment is often assumed to be static



(a) Traditional panorama

(b) *Tonemapped* HDR panorama

Figure 1.4: IBL is based on using panoramic images to represent the incident illumination on virtual objects. **a)** Traditional photography can not capture the full dynamic range found in many real scenes, in this scene the interior of the house and the sky is not accurately represented. Using this representation in rendering results in images that look flat due to suppressed highlights and missing reflections. **b)** HDR images capture the full dynamic range of the scene enabling both specular and diffuse objects to be accurately rendered. In contrast to traditional photography, often representing images using non-linearly mapped 8-bit values, the HDR pixel values represent radiometrically calibrated measurements. The HDR image shown here has been *tonemapped* to a lower dynamic range, making it possible to show in print.

and only captured at a single point in the scene. This can result in artificially looking results when applying IBL techniques in scenes where the illumination is dynamic or includes spatial variations such as cast shadows etc. To overcome these limitations, IBL methods for representing temporally and spatially varying illumination conditions have been proposed [86, 209]. These techniques rely on the use of HDR video to efficiently capture dense representations of the illumination in the scene. However, previous methods are limited, as they often require substantial manual tweaking and user effort, and have been limited by the lack of robust HDR video cameras capable of capturing the full dynamic range in the scene.

## 1.3 Contributions

The main contribution of this thesis can be summarized as the development of models and algorithms for efficient and accurate photorealistic rendering of synthetic objects in real environments. In particular, the material contained in this thesis addresses four related areas : HDR imaging, IBL, reflectance modeling, and physically based rendering. Below we give a brief overview of the major contribution in each of these areas and list the novel technical developments:

### HDR Video

This thesis presents methods and algorithms for state-of-the-art HDR video capture using both custom built cameras with multiple sensors, and consumer cameras using a spatially varying ISO setting over the sensor. Our technical contributions in this area can be summarized as:

- A statistically motivated algorithm for HDR reconstruction from raw camera data, combining demosaicing, denoising, and HDR fusion in a single processing operation.
- A radiometric noise model adapted to HDR video cameras.
- Methods for improving the sharpness of HDR reconstructions based on adaptive filters.
- Demonstrations of state-of-the-art HDR video capture using multi-sensor and dual-ISO camera configurations.

### IBL

Enabled by the contributions in HDR video capture, the thesis also presents new methods for capturing and rendering with temporally and spatially varying illumination conditions. Specifically, the technical contributions in this area can be summarized as:

- Practical and robust methods for rendering with temporally varying illumination conditions captured using omnidirectional HDR video.
- Methods for reconstructing scene representations that allow for accurate and efficient rendering of virtual objects in scenes with spatially varying illumination conditions.

### Reflectance modeling

The thesis also presents new parametric reflectance models for modeling glossy surface reflectance. Compared to previous work, the proposed models provide



better fits to measured reflectance data, enabling more accurate and efficient renderings of materials with glossy reflectance, such as metals and coated plastics. The technical contributions in this area are:

- An empirical study of material reflectance and properties of different model parameterizations.
- Two new parametric reflectance models for surfaces exhibiting wide angle gloss.
- An extended model for materials exhibiting anisotropic wide angle gloss.

### Physically Based Rendering

Finally, the thesis presents a novel physically based rendering algorithm that is designed to work particularly well in scenes that traditionally have been very difficult to render, such as scenes containing participating media and glossy materials. The technical developments in the area of physically based rendering can be summarized as:

- A rendering algorithm based on Markov Chain Monte Carlo (MCMC) that allows unbiased approximations to be used in place of computationally expensive, or intractable, light transport models. This enables us not only to increase the generality and flexibility of MCMC based rendering algorithms, but also to improve their efficiency.
- Demonstration of how the proposed rendering algorithm enables efficient MCMC based rendering of scenes containing heterogeneous participating media and glossy transfer.

## 1.4 Publications

The published work of the author with direct relevance to this thesis is listed below in reverse chronological order. Papers marked with a "\*" are included in the second part of the thesis.

- \* J. Kronander, T. B. Schön, and J. Unger. Pseudo-Marginal Metropolis Light Transport. In *SIGGRAPH Asia Technical Briefs*, 2015
- \* S. Hajisharif, J. Kronander, and J. Unger. Adaptive dualISO HDR-reconstruction. *Submitted to EURASIP Journal on Image and Video Processing*, 2015
- \* J. Kronander, F. Banterle, A. Gardner, E. Miandji, and J. Unger. Photorealistic rendering of mixed reality scenes. *Computer Graphics Forum (Proc. of Eurographics STARs)*, 34(2):643–665, 2015



E. Miandji, J. Kronander, and J. Unger. Compressive image reconstruction in reduced union of subspaces. *Computer Graphics Forum (Proc. of Eurographics)*, 34(2), 2015

S. Hajisharif, J. Kronander, and J. Unger. HDR reconstruction for alternating gain (ISO) sensor readout. In *Eurographics Short Papers*, 2014

J. Kronander, J. Dahlin, D. Jönsson, M. Kok, T. B. Schön, and J. Unger. Real-time video based lighting using HDR video and Sequential Monte Carlo samplers. In *Proceedings of EUSIPCO'14: Special Session on HDR-video*, 2014

\* J. Kronander, S. Gustavson, G. Bonnet, A. Ynnerman, and J. Unger. A unified framework for multi-sensor HDR video reconstruction. *Signal Processing: Image Communication*, 29(2), 2014

\* J. Unger, J. Kronander, P. Larsson, S. Gustavson, and A. Ynnerman. Temporally and Spatially Varying Image Based Lighting using HDR-video. In *Proceedings of EUSIPCO'13: Special Session on HDR-video*, 2013

\* J. Kronander, S. Gustavson, G. Bonnet, and J. Unger. Unified HDR reconstruction from raw CFA data. In *IEEE International Conference on Computational Photography (ICCP)*, 2013

\* J. Unger, J. Kronander, P. Larsson, S. Gustavson, J. Löw, and A. Ynnerman. Spatially varying image based lighting using hdr-video. *Computers & graphics*, 37(7):923–934, 2013

E. Miandji, J. Kronander, and J. Unger. Learning based compression of surface light fields for real-time rendering of global illumination scenes. In *SIGGRAPH Asia Technical Briefs*, 2013

\* J. Löw, J. Kronander, A. Ynnerman, and J. Unger. BRDF models for accurate and efficient rendering of glossy surfaces. *ACM Transactions on Graphics (TOG)*, 31(1):9, 2012

Other publications by the author, loosely related to, but not included in this thesis, are:

J. Kronander and T. B. Schön. Robust auxiliary particle filters using multiple importance sampling. In *IEEE Statistical Signal Processing Workshop (SSP)*, 2014

J. Kronander, T. B. Schön, and J. Dahlin. Backward Sequential Monte Carlo for marginal smoothing. In *IEEE Statistical Signal Processing Workshop (SSP)*, 2014

A. Tsirikoglou, S. Ekeberg, J. Vikström, J. Kronander, and J. Unger. S(wi)ss: A flexible and robust sub-surface scattering shader. In *SIGRAD*, 2014

D. Jönsson, J. Kronander, T. Ropinski, and A. Ynnerman. Historygrams: Enabling interactive global illumination in direct volume rendering using photon mapping. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2364–2371, 2012

J. Kronander, D. Jönsson, J. Low, P. Ljung, A. Ynnerman, and J. Unger. Efficient visibility encoding for dynamic illumination in direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 18(3):447–462, 2012

S. Hajisharif, J. Kronander, E. Miandji, and J. Unger. Real-time image based lighting with streaming HDR-light probe sequences. In *SIGRAD*, 2012

S. Lindholm and J. Kronander. Accounting for uncertainty in medical data: A cuda implementation of normalized convolution. In *SIGRAD*, 2011

E. Miandji, J. Kronander, and J. Unger. Geometry independent surface light fields for real time rendering of precomputed global illumination. In *SIGRAD*, 2011

J. Kronander, J. Unger, T. Möller, and A. Ynnerman. Estimation and modeling of actual numerical errors in volume rendering. *Computer Graphics Forum (Proc. of Eurovis)*, 29(3):893–902, 2010

## 1.5 Thesis outline

The thesis is divided into two parts. The first part introduces background theory and gives an overview of the contributions presented in the thesis. The second part is a compilation of eight selected publications that provide more detailed descriptions of the research leading up to this thesis. Note that the first publication, Paper A, is a review article covering photorealistic rendering of synthetic objects in real scenes. Paper A should therefore be viewed as part of the introduction, complementing the material presented in Part I.

### 1.5.1 Outline of part I

The first part of the thesis is divided into several chapters, each discussing a specific topic. Apart from the second chapter presenting the fundamentals of light transport theory and the last chapter providing concluding remarks, each chapter first introduces the background of the related topic and then discusses how the contributions in the second part of the thesis address the limitations of current methods. Each chapter is also concluded with a short summary and a discussion of possible venues for future work in the topic.

To produce photorealistic renderings of digital models it is necessary to introduce appropriate measurements and mathematical models describing the physics of light transport. The models of light transport used in this thesis are described in chapter 2. In chapter 3 we then discuss the simulation of light transport, the basis of physically based rendering, using stochastic *Monte Carlo* methods. In chapter 4 we then discuss how virtual objects can be rendered so that they can be seamlessly integrated into real images using IBL techniques.

The simulation of the light transport in augmented scenes requires several components of the scene model to be specified, for example light sources, cameras, and reflectance properties of surfaces in the scene. A central part of performing accurate and efficient measurements of these properties in the real world is HDR imaging. In chapter 5 we present several techniques for accurate HDR capture. In chapter 6 we then discuss techniques for measuring and modeling the reflectance of real world surfaces. Finally, in chapter 7 some concluding remarks and directions for future work are discussed.

## 1.5.2 Outline of part II

This part consists of a collection of eight selected, previously published, publications as outlined below. Besides a short summary of the main content, a brief explanation of the background of the publication and the contributions of the author is provided.

### Paper A: Photorealistic rendering of mixed reality scenes

J. Kronander, F. Banterle, A. Gardner, E. Miandji, and J. Unger. Photo-realistic rendering of mixed reality scenes. *Computer Graphics Forum (Proc. of Eurographics STARs)*, 34(2):643–665, 2015.

This paper describes an overview and categorization of the state-of-the-art methods for rendering synthetic objects into real images and video. The survey paper provide an overview of the many facets of mixed reality rendering and connects the topics of the other papers in this thesis.

**Background and contributions:** When studying previous surveys on the topic published in the computer graphics and the augmented reality literature, the need for an up-to-date survey was identified. The survey includes work from both of these fields, as well as recent methods developed in the computer vision literature. The state-of-the-art report (STAR) was written in collaboration with other researchers working at Linköping University and Francesco Banterle from the visual computing laboratory located in Pisa, Italy. The STAR was presented at Eurographics 2015 in Zurich, Switzerland.

### Paper B: Pseudo-marginal metropolis light transport

J. Kronander, T. B. Schön, and J. Unger. Pseudo-Marginal Metropolis Light Transport. In *SIGGRAPH Asia Technical Briefs*, 2015.

This paper introduces a physically based light transport algorithm based on Markov Chain Monte Carlo methods that allows approximation to be used in

place of exact quantities, while still converging to the exact result. The method is closely related to the *pseudo-marginal* MCMC construction recently developed in statistics for inference in Bayesian models with intractable likelihoods. The paper shows that the proposed rendering algorithm allows for efficient rendering of scenes containing glossy transfer and participating media.

**Background and contributions:** The idea of using the pseudo-marginal MCMC approach for deriving new rendering algorithms came up when working on *Sequential Monte Carlo methods*, another class of Monte Carlo methods that has seen widespread use in statistics. The paper was written in close collaboration with Thomas B. Schön, professor of Automatic Control at Uppsala university. The paper was presented at SIGGRAPH Asia held in Kobe, Japan 2015 and as a poster at the 2015 Sequential Monte Carlo workshop located in Paris.

### Paper C: Temporally and Spatially Varying Image Based Lighting using HDR-video

J. Unger, J. Kronander, P. Larsson, S. Gustavson, and A. Ynnerman. Temporally and Spatially Varying Image Based Lighting using HDR-video. In *Proceedings of EUSIPCO'13: Special Session on HDR-video*, 2013.

This paper describes an IBL pipeline for capturing and rendering with temporally or spatially varying illumination using HDR video. Based on a dense set of captured video light probes synthetic objects can be composited into real world scenes, such that it appears that they were actually there in the first place, reflecting the dynamical and spatially varying character of the real world illumination in the scene.

**Background and contributions:** In 2011 a state-of-the-art HDR video camera was developed in collaboration between the computer graphics group at Linköping University and Spheron VR. This camera enabled the development of a system for temporally varying IBL. The author worked on all of the methods presented in the paper. Several of the renderings in the paper were generated in collaboration with Christian Bloch working at a visual effects studio located in California. Results from this work were featured in Blochs textbook on practical techniques for IBL and HDR imaging [28].

### Paper D: Spatially varying image based lighting using HDR-video

J. Unger, J. Kronander, P. Larsson, S. Gustavson, J. Löw, and A. Ynnerman. Spatially varying image based lighting using hdr-video. *Computers & graphics*, 37(7):923–934, 2013.

This paper presents a complete system, including capturing, processing, editing, and rendering with spatially varying IBL. The presented approach is based on extracting approximate geometry onto which captured HDR video data is projected and stored as light fields. Explicit extraction of direct light sources in the scene enables the user to edit the real world illumination and fit reflectance parameters of geometric surfaces in the recovered scene model.

**Background and contributions:** The main supervisor Jonas Unger was the main contributor to the development of an approximate scene reconstruction framework to represent spatially varying illumination. The author worked on methods for geometry extraction, light source recovery, light field projection, and the development of robust algorithms for representing HDR video data. He also helped to write the article. Many of the examples presented in the article are taken from the real production environment at IKEA Communications AB, located in Älmhult, Sweden.

#### Paper E: Unified HDR Reconstruction from raw CFA data

J. Kronander, S. Gustavson, G. Bonnet, and J. Unger. Unified HDR reconstruction from raw CFA data. In *IEEE International Conference on Computational Photography (ICCP)*, 2013.

This paper introduces a unified framework for reconstructing HDR images and video frames from raw sensor data captured with multiple exposures. Using local polynomial approximation filters, several low level image processing tasks such as realignment, color filter interpolation, HDR fusion, and noise reduction can be formulated as a single noise aware filtering operation. In the paper a radiometric camera model suitable for HDR video cameras is also introduced and used for improving the local polynomial approximations.

**Background and contributions:** The benefits of a unified reconstruction framework was identified when developing a reconstruction software for a new multi-sensor HDR video camera, designed by researchers at Linköping University and the German camera manufacturer Spheron VR. The idea of using local polynomial approximations was inspired by normalized convolution filtering [113], a technique the author came in contact with during a graduate course in multidimensional filtering. The paper was presented at ICCP 2013 held at Harvard, shortly after the intense police investigation to locate the Boston marathon bombers.

### Paper F: A unified framework for multi-sensor HDR video reconstruction

J. Kronander, S. Gustavson, G. Bonnet, A. Ynnerman, and J. Unger.  
A unified framework for multi-sensor HDR video reconstruction.  
*Signal Processing: Image Communication*, 29(2), 2014.

This paper extends the previous conference publication, paper E, with an anisotropic filtering operation that adapts the filter supports to the image structure. This results in sharper reconstructions around edges and corners, and less noise in homogenous image regions. Using a state-of-the-art multi-sensor HDR video camera, the paper shows how the proposed framework produces better results than previous multi-sensor HDR video reconstruction methods.

**Background and contributions:** A limitation of the previous unified reconstruction framework, presented in paper E, was that it did not include some of the desirable features of modern color filter interpolation and denoising algorithms. Inspired by the design of such algorithms, a natural extension of the previous framework was to consider anisotropic filtering supports to enable sharper reconstructions around edges, this reduces noise and provide less color artifacts in high-frequency regions

### Paper G: Adaptive dualISO HDR-reconstruction

S. Hajisharif, J. Kronander, and J. Unger. Adaptive dualISO HDR-reconstruction. *Submitted to EURASIP Journal on Image and Video Processing*, 2015.

This paper extends the HDR reconstruction framework presented in papers D and E to use statistically motivated adaptive window selection. The paper shows how high quality HDR frames can be reconstructed from a standard Canon DSLR camera running the *Magic Lantern* software in the *dual-ISO* configuration, where interleaved rows in the sensor are amplified with different ISO settings.

**Background and contributions:** The unified reconstruction framework was first developed with multi-sensor HDR video cameras in mind. However, we later discovered that it was useful for reconstructing other input data as well, such as data from dual-ISO. In an earlier publication, we showed that our unified reconstruction framework, presented in paper E, provided better results than other methods for dual-ISO capture [80]. The development of adaptive filtering supports that takes into account the statistical properties of the noise was performed in close collaboration between Saghi Hajisharif, the author, and Jonas Unger. The author contributed with ideas and theoretical foundations

for the design of the adaptive window supports. The author also helped with writing the article.

### Paper H: BRDF models for accurate and efficient rendering of glossy surfaces

J. Löw, J. Kronander, A. Ynnerman, and J. Unger. BRDF models for accurate and efficient rendering of glossy surfaces. *ACM Transactions on Graphics (TOG)*, 31(1):9, 2012.

introduces two new parametric BRDF models for modeling wide angle scatter, or gloss, inspired by the Rayleigh-Rice theory [193] for optical scattering from smooth surfaces. Based on an empirical study of material reflectance, two different parameterizations are used; the standard half angle parametrization, similar to previous models based on microfacet theory, and the projected deviation vector formulation.

**Background and contributions:** Joakim Löw was responsible for deriving the foundations of the new BRDF models. The author helped with the development of the new models and was responsible for deriving the theoretical foundations for importance sampling the developed BRDF models. The author of this thesis also made the practical implementation of the model in a renderer, and was responsible for generating the rendered images in the article and the supplementary material. The author also helped to write and edit the paper.





## Fundamentals of Light transport

To create photorealistic renderings – images depicting a virtual environment as seen by a virtual camera – it is necessary to specify a detailed three-dimensional model of the scene. The geometry of the scene consists of three-dimensional surfaces that are often described by simpler geometric primitives such as triangles or other surface patches. It is also necessary to specify properties, such as focus, viewing angle, and the position of the virtual camera. Finally, the light sources should be modeled as well as the material properties of surfaces in the scene, which describe their appearance and color. The rendered image is then computed by performing a detailed physically based simulation of how light propagates in the scene and finally reaches the virtual camera sensor. The rendering process is illustrated in Figure 2.1. The propagation of light emitted from the light sources and its interaction with materials on surfaces in the scene is described by light transport theory.

This chapter outlines the basic quantities, domains and equations that form the basis of light transport theory used in physically based rendering. More in-depth discussions about the theory of light transport can also be found in many excellent books, such as, [59] and [174].

### 2.1 Light transport model

Light transport can be modeled in different ways with varying levels of detail and complexity. At the most detailed level, quantum electrodynamics describes the interaction between light and matter at the quantum scale. Classical electromagnetic theory based around Maxwell's equations presents a somewhat

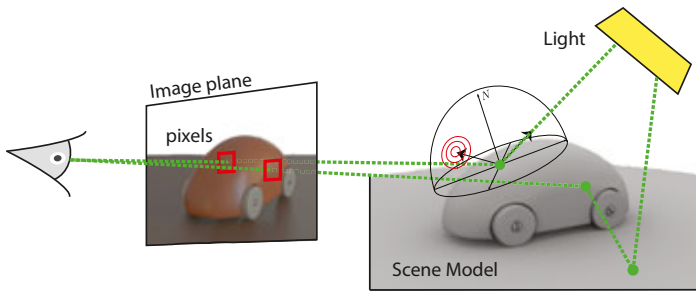


Figure 2.1: To create computer graphics renderings, a mathematical model describing the 3D geometry, light sources and material properties in the scene is first specified. The resulting image is then computed by simulating the amount of light reaching a virtual camera sensor. Light is emitted from light sources in the scene and may be reflected several times before reaching the camera sensor.

coarser model, which describes visible light as electromagnetic radiation with a wavelength from around 380 nm (blue) to 740 nm (red). Neglecting effects such as diffraction and interference, a simpler model of light transport is provided by geometric optics (also known as ray optics). In this model, light propagates along rays and it can be emitted, reflected, and transmitted. For computational efficiency, physically based rendering is almost always based on the geometric optics model of light transport, ignoring the speed of light and treating the energy transfer as instantaneous. It is also common to further approximate the classical geometrical optics model by ignoring the polarization of light as its visual impact is often negligible. Although this simplified model is usually sufficient, more advanced effects, such as diffraction on metallic surfaces, can often be used inside this model in a localized manner, for example to derive surface reflection models, see section 6 for more details. To produce colored renderings, the actual wavelength distribution of the simulated light is typically approximated by only considering a set of discrete wavelength bands. Often the three wavelength bands corresponding to the additive primary colors, red, green and blue (RGB) is sufficient, however, sometimes spectral renderings simulating more than 3 separate color channels produce more accurate results [174].

## 2.2 Radiometry

*Radiometric* quantities allows us to measure and quantify light transport in a structured manner. The central radiometric quantity of interest in physically based rendering is *radiance*,  $L(x, \omega)$ , which describes how much energy/light

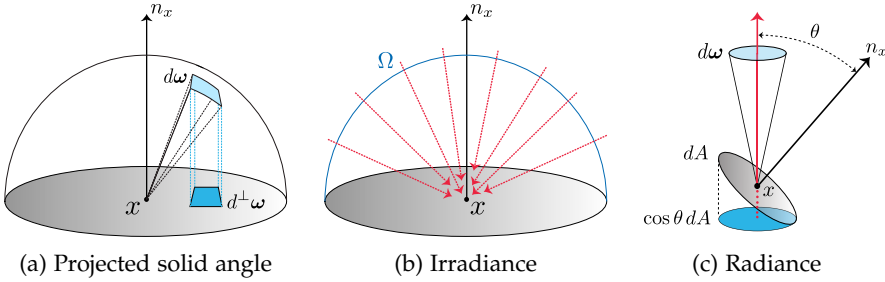


Figure 2.2: Illustration of common measures and radiometric quantities. The different quantities are described in detail in the text.

flows through the point  $x$  in the direction  $\omega$ . Intuitively, radiance can be thought of as the amount of energy/light arriving on a small surface patch  $dA$  at  $x$  perpendicular to the direction  $\omega$  in a small cone  $d\omega$  centered around  $\omega$ , see figure 2.2c. To precisely specify radiance and other related radiometric quantities, we first need to introduce appropriate domains and measures.

### 2.2.1 Domains and measures

Directions are represented by normalized vectors,  $\omega$ , on the unit sphere  $S^2$  in  $\mathcal{R}^3$ . To integrate a function defined on the unit sphere  $f(\omega)$ , we express the integration with respect to the *solid angle measure*,  $d\omega$ , as :

$$\int_{S^2} f(\omega) d\omega = \int_0^{2\pi} \int_0^\pi f(\theta, \phi) \sin \theta \, d\theta d\phi, \quad (2.1)$$

where,  $\{\theta, \phi\}$ , denote the spherical coordinates. To integrate the incident light at a point  $x$  on a surface with normal  $\mathbf{n}_x$ , the *projected solid angle* measure,  $d^\perp\omega$ , is used:

$$\int_{S^2} f(x, \omega) d^\perp\omega = \int_{S^2} f(x, \omega) |\mathbf{n}_x \cdot \omega| d\omega = \int_0^{2\pi} \int_0^\pi f(x, \theta, \phi) \cos \theta \sin \theta \, d\theta d\phi, \quad (2.2)$$

where the additional cosine factor represents the foreshortening effect due to the angle of incidence,  $\theta$ , and can be thought of as representing the projection of the differential solid angle onto the unit disk, see figure 2.2a for an illustration.

## 2.2.2 Radiometric quantities

We can now introduce some of the most common radiometric quantities, each of which is defined by measuring the energy of light with respect to different units. **Raidant power** (Flux) is defined as the energy,  $Q$ , per unit time,

$$\Phi = \frac{dQ}{dt}, \quad (2.3)$$

and has the unit Watt,  $W$ , (Joule per second). This quantity can for example be used to describe the total emitted power of a light source with finite area. A related quantity, **Irradiance** is the power per unit surface area, arriving at a point  $x$

$$E(x) = \frac{d\Phi(x)}{dA(x)}. \quad (2.4)$$

Finally, **radiance**, is defined as the incident or outgoing power at a surface per unit projected solid angle per unit area,

$$L(x, \omega) = \frac{d^2\Phi(x, \omega)}{d^\perp\omega dA(x)} = \frac{d^2\Phi(x, \omega)}{|\mathbf{n}_x \cdot \omega| d\omega dA(x)}, \quad (2.5)$$

where  $\mathbf{n}_x$  is the surface normal at the point  $x$ . It is also possible to define radiance as the power per unit solid angle per unit projected area,  $dA^\perp(x) = |\mathbf{n}_x \cdot \omega| dA(x)$ , see figure 2.2c for an illustration.

It is convenient to denote incident radiance that arrives at a point  $x$  from direction  $\omega$ , or from the point  $y$ , by  $L(x \leftarrow \omega)$  and  $L(x \leftarrow y)$  respectively. Similarly we let  $L(x \rightarrow \omega)$  and  $L(x \rightarrow y)$  denote the outgoing, scattered or emitted, radiance, receptively.

An important relationship in geometrical optics is the *radiance invariance law* which states that the radiance does not change along a ray in vacuum, that is

$$L(x \leftarrow y) = L(y \rightarrow x). \quad (2.6)$$

The irradiance at a point  $x$  can be computed by integrating the radiance with respect to the projected solid angle measure over the visible hemisphere,  $\Omega$ , centered around the normal,  $\mathbf{n}_x$ , i.e.  $\Omega = \{\omega \in \mathcal{S}^2 : (\mathbf{n}_x \cdot \omega) > 0\}$ ,

$$E(x) = \int_{\Omega} L(x \leftarrow \omega) |\mathbf{n}_x \cdot \omega| d\omega. \quad (2.7)$$

These radiometric quantities are summarized in figure 2.2.

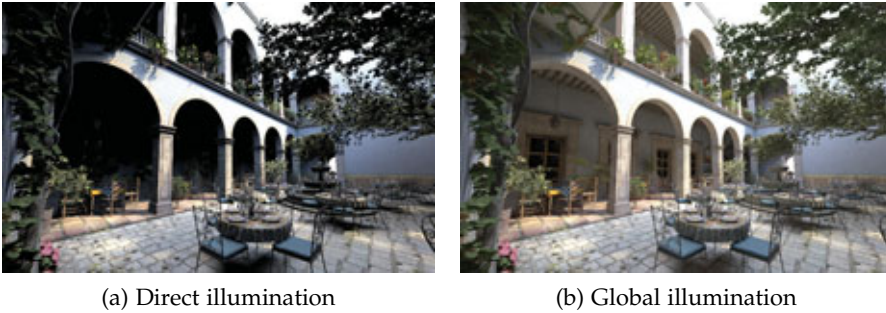


Figure 2.3: Renderings of a complex scene. rendered using, **a)** a single evaluation of the rendering equation (2.8) corresponding to *direct illumination*, only accounting for light that reflects once in the scene. **b)** recursive evaluation of the rendering equation, corresponding to *global illumination*, accounting for light that reflects multiple times in the scene. Scene modeled by Guillermo M. Leal Llaguno and rendered using PBRT [174]

## 2.3 Rendering equation

For any surface in the scene, the outgoing radiance,  $L(x \rightarrow \omega_o)$ , leaving a point  $x$  in a direction  $\omega_o$  can be described as the sum of the emitted radiance  $L_e(x \rightarrow \omega_o)$  and the reflected radiance  $L_r(x \rightarrow \omega_o)$ , at  $x$  towards  $\omega_o$ . For now we will assume that there is no participating media in the scene, ie we assume that light travels unobstructed between surfaces in the scene. The reflected radiance can then be computed by integrating the incident radiance over the visible hemisphere,  $\Omega$ , at  $x$ . This relationship is formalized by the *rendering equation* [101]:

$$L(x \rightarrow \omega_o) = L_e(x \rightarrow \omega_o) + \underbrace{\int_{\Omega} L(x \leftarrow \omega_i) \rho(x, \omega_o, \omega_i) (\mathbf{n}_x \cdot \omega_i)}_{L_r(x \rightarrow \omega_o) d\omega_i}, \quad (2.8)$$

where  $\rho(x, \omega_o, \omega_i)$  is the *bidirectional reflectance distribution function* (BRDF) describing the surface reflectance. Intuitively, the BRDF describes how much of the incident light from direction  $\omega_i$  is scattered into the direction  $\omega_o$ . Ideally smooth materials are characterized by having a specular reflection described by a Dirac delta distribution. Most real materials on the other hand usually have a smooth BRDF function. More details on the properties of the BRDF are provided in chapter 6. Note that for surfaces which are not modeling light sources,  $L_e(x \rightarrow \omega_o) = 0$ .

The rendering equation describes the interaction of light with surfaces in the scene. Light that only interact once with surfaces, often referred to as *direct illumination*, can be described by applying the rendering equation once.

However, to account for light that interact multiple times with surfaces in the scene, often referred to as *global illumination*, the rendering equation have to be evaluated recursively. The difference between direct and global illumination is illustrated in figure 2.3. In section 2.5 we will describe other formulations of light transport that allows us to express the radiance reaching the virtual camera in a more direct form that don't require recursive evaluation.

### Area formulation

The formulation of the rendering equation provided by equation (2.8) expresses the reflected radiance as an integral over the visible hemisphere at  $x$ . Sometimes it can be more convenient to describe the reflected radiance as an integral over the surfaces in the scene rather than over the visible hemisphere. This leads to the *area formulation* of the rendering equation, which expresses the reflected radiance at  $x$  as an integral over all other points,  $y \in \mathcal{M}$ , in the scene. Here  $\mathcal{M} \in \mathcal{R}^3$  denotes the set of 2-dimensional manifolds that constitute the surfaces of the scene.

The area formulation is based on performing a change of variables using the relation:

$$d\omega_i = \frac{(\mathbf{n}_y \cdot (-\omega_i))}{\|x - y\|^2} dA(y), \quad (2.9)$$

where  $\mathbf{n}_y$  is the surface normal at  $y$ . In order to change the integration from the hemisphere of directions to surface area it is also necessary to take into account if there is a clear line of sight form  $x$  to  $y$ . This relationship is expressed using a binary *visibility function*, defined by

$$V(x, y) = \begin{cases} 1 & \text{if } x \text{ and } y \text{ are mutually visible,} \\ 0 & \text{otherwise.} \end{cases} \quad (2.10)$$

Using these relations we can formulate the rendering equation as:

$$L(x \rightarrow \omega_o) = L_e(x \rightarrow \omega_o) + \int_{\mathcal{M}} L(x \leftarrow y) \rho(x, \omega_o, \omega_i) V(x, y) G(x, y) dA(y), \quad (2.11)$$

where

$$G(x, y) = \frac{(\mathbf{n}_x \cdot \omega_i)(\mathbf{n}_y \cdot (-\omega_i))}{\|x - y\|^2}, \quad (2.12)$$

is the *Geometry term* taking into account the relative differential areas at  $x$  and  $y$ .

## 2.4 Radiative transfer equation

In the previous section, we assumed that there was no participating media in the scene. This implies that the radiance leaving a surface remains unchanged

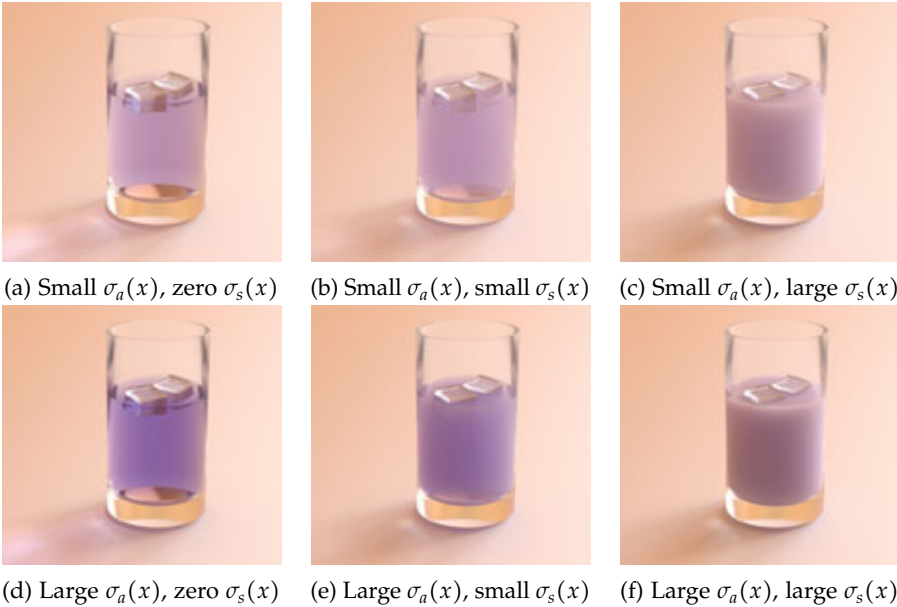


Figure 2.4: Renderings of a glass filled with liquid modeled using homogenous media with varying absorption coefficient,  $\sigma_a(x)$ , and scattering coefficient,  $\sigma_s(x)$ . **Upper row:** Renderings using a small absorption coefficient and **a)** no scattering, **b)** a small scattering coefficient, and **c)** a large scattering coefficient. **Lower row:** Renderings using a large absorption coefficient and **d)** no scattering, **e)** a small scattering coefficient, and **f)** a large scattering coefficient.

until it hits another surface. In reality, however, surfaces of interest are often located in different forms of participating media, such as air, water or fog. For optically thin media, such as clean air, the assumption that light travels unobstructed between surfaces serves as a reasonable approximation for short distances. However, over longer distances even clean air scatters light (the sky appears blue due to such scattering), and for photorealistic rendering of scenes with denser media such as water, smoke, fire etc, it is necessary to consider models that take into account the effects of how light interacts with the participating media in the scene.

In computer graphics, and in many other fields of science such as neutron transport [190] and medical physics [10], the media is modeled as a large number of microscopic scattering particles that the light can interact with. As the sheer number of these particles makes deterministic models infeasible, we instead make use of *Linear Transport Theory* that, similar to other statistical models used in physics [129], considers the aggregated behavior of a large

number of randomly distributed particles. The main insight in these approaches is that we do not need to represent the exact position of each individual particle as long as their average effect on the light propagation through the media can be accounted for. To further simplify the models, light-particle interactions in the media is assumed to be independent, that is if the light interacts with a particle in the media, this interaction is statistically independent from the outcome of subsequent interaction events (in other words, a random photon trajectory can be characterized by a Markov process).

In computer graphics we are interested in simulating the interactions between particles in the media and photons with relatively low energy (visible light). This allows us to model interactions using two types of events, either a photon is absorbed (for example converted to heat) or it collides with a particle in the medium and scatters in another direction. In other fields, considering photons with higher energy, such as radiation dosimetry [10], more complex collision events, such as *Compton scattering* and *pair production*, have to be considered as well [185]. The relative probability of a particle being absorbed or scattered per unit length is described by the *absorption coefficient*,  $\sigma_a$ , and the *scattering coefficient*,  $\sigma_s$ , respectively. These quantities generally depend on the density of particles in the medium, and are often allowed to vary spatially. Media where  $\sigma_a(x)$  and  $\sigma_s(x)$  are constant for all  $x$  is referred to as *homogeneous*, otherwise, if the coefficients vary spatially, the media is *heterogeneous*. The absorption and scattering coefficient can have a profound effect on the appearance of the media, an illustration is given in figure 2.4. The sum of  $\sigma_a(x)$  and  $\sigma_s(x)$  constitute the probability that an interaction takes place per unit length, and is described by the *extinction coefficient*  $\sigma_t(x) = \sigma_a(x) + \sigma_s(x)$ . Both absorption and scattering can reduce the radiance along a ray in the medium, as photons traveling along the ray can be absorbed or scattered into different directions, referred to as *out-scattering*. Similarly, the radiance along a ray can also increase due to emission of photons in the media, or from *in-scattering* of photons originating from other directions.

The net effect of the change of radiance along a ray in direction  $\omega$  from a point  $x$ , is modeled by an integro-differential equation known as the *radiative transfer equation* (RTE) [36] as:

$$(\omega \cdot \nabla)L(x \rightarrow \omega) = \underbrace{L_e(x \rightarrow \omega) + L_i(x \rightarrow \omega)}_{\text{net increase}} - \underbrace{\sigma_a(x)L(x \rightarrow \omega) + \sigma_s(x)L(x \rightarrow \omega)}_{\text{net extinction}}, \quad (2.13)$$

where

$$L_i(x \rightarrow \omega) = \sigma_s(x) \int_{S^2} \rho_p(x, \omega, \omega_i) L(x \leftarrow \omega_i) d\omega_i, \quad (2.14)$$



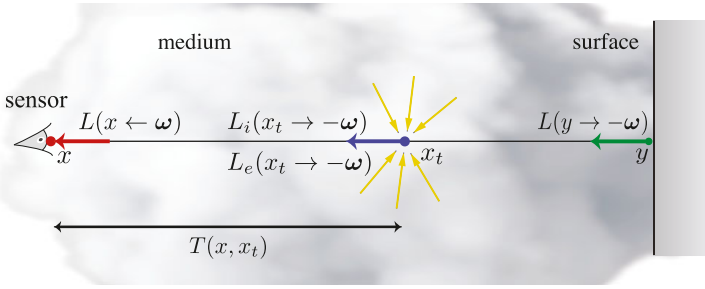


Figure 2.5: The radiative transport equation describes the radiance reaching a point  $x$  from direction  $\omega$ ,  $L(x \leftarrow \omega)$  as a sum of the attenuated radiance from the nearest surface,  $L(y \rightarrow -\omega)$ , and the accumulated (integrated) in-scattering,  $L_i(x_t \rightarrow -\omega)$ , and emission,  $L_e(x_t \rightarrow -\omega)$ , for points  $x_t$  along the ray in the medium.

describes the in-scattering, given by an integral over the unit sphere,  $S^2$ , defined using a *phase function*,  $\rho_p(x, \omega, \omega_i)$  that models the angular distribution of light scattering at a point  $x$  in the medium.  $L_e(x \rightarrow \omega)$  represents the radiance emitted in the direction of the ray from the medium, given in units of radiance per unit length.

Using the rendering equation (2.8) as a boundary condition, the RTE can be formulated in integral form [14, 95], describing the radiance reaching a point  $x$  from direction  $\omega$ , illustrated in figure 2.5, as:

$$L(x \leftarrow \omega) = \underbrace{T(x, y)L(y \rightarrow -\omega)}_{\text{radiance originating from closest surface}} + \underbrace{\int_0^d T(x, x_t) (L_e(x_t \rightarrow -\omega) + L_i(x_t \rightarrow -\omega)) dt}_{\text{radiance from accumulated in-scattering and emission in the volume}} \quad (2.15)$$

where  $y$  is the first point on a surface in the direction  $\omega$  from  $x$ ,  $d$  is the distance from  $x$  to  $y$ ,  $x_t = x + t\omega$   $t \in (0, d)$  are points along the ray, and,  $T(x, y)$  is the transmittance between the points  $x$  and  $x_t$  given by:

$$T(x, x_t) = \exp\left(-\int_0^{\|x-x_t\|} \sigma_t(x_t') dt'\right), \quad (2.16)$$

where  $\sigma_t(x)$  denotes the extinction coefficient at  $x$  describing the loss of light due to absorption and out-scattering per unit distance. The integral form of the RTE is also commonly referred to as the *volume rendering equation* [62].

## 2.5 Path integral formulation

To render an image with  $M$  pixels it is necessary to compute a set of *pixel measurements*,  $I_1, I_2, \dots, I_M$ , that expresses the value of the pixel as a function of the incident radiance,  $L(x \leftarrow \omega)$ , over the pixel. By defining a set of pixel response functions (pixel filters),  $W_j(x \leftarrow \omega), j = \{1, \dots, M\}$ , we can define the pixel measurements as

$$I_j = \int_{\mathcal{A}_s} \int_{\Omega} W_j(x \leftarrow \omega_i) L(x \leftarrow \omega_i) (n_x \cdot \omega_i) d\omega_i d\mathcal{A}(x), \quad (2.17)$$

where  $\mathcal{A}_s$  denotes the area of the pixel.

To simulate the light transport in the scene the RTE – or the rendering equation – has to be evaluated recursively, due to the fact that light can be reflected multiple times in the scene before reaching the virtual camera sensor, i.e. the outgoing radiance at a point in the scene affects the incident radiance at all other points visible from that point. This, *implicit*, definition of the incident radiance reaching a pixel in the virtual sensor makes it difficult to derive rendering algorithms that try to reason about the structure of possible light transport paths in the scene. An alternative formulation of the light transport is given by the *path integral formulation* [95, 172, 213]. In contrast to the RTE, the path integral formulation provides *explicit* expressions for pixel measurements as integrals over light paths in the scene. This explicit form allows for the design of rendering algorithms that can use more advanced methods for simulating possible paths in the scene. For example once a high contribution path has been found, similar paths can be found by small perturbations of the scattering events constituting the path.

The path space formulation is defined by considering light transport paths that connects a light source (or any surface/volume emitting light) in the scene to the sensor. A path,  $\bar{x}_k$ , of length  $k$  is defined by  $k + 1$  scattering events, or vertices,  $\bar{x} = \{v_0^x, v_1^x, \dots, v_k^x\}$ , located on surfaces or in participating media. The first vertex,  $v_0^x$  is always located at a light source (emitter) and the last vertex  $v_k^x$  is always located on the sensor.

Let  $\mathcal{P}_k$  denote the set of all paths of length  $k$ . The complete set of all paths of all lengths are then defined by the *path space*,

$$\mathcal{P} = \bigcup_{k=1}^{\infty} \mathcal{P}_k. \quad (2.18)$$

To define an integral over the path space, it is necessary to introduce appropriate measures over the path space. The *path space measure*,  $\mu$ , is defined by a product

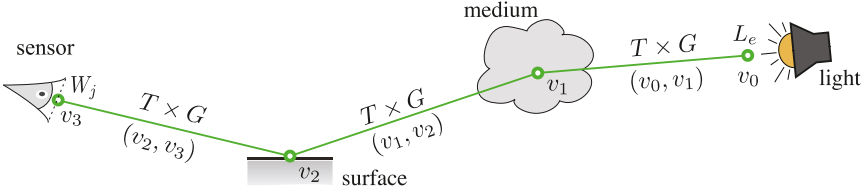


Figure 2.6: Illustration of the notation used in describing the light transport in a scene using the path integral formulation.

measure over path vertices as<sup>1</sup>

$$d\mu(\bar{x}_k) = \prod_{i=0}^k d\mu(v_i^x), \quad (2.19)$$

where  $d\mu(v_i^x)$  is defined as area integration,  $d\mu(v_i^x) = d\mathcal{A}(v_i^x)$ , for path vertices on surfaces and as volume integration,  $d\mu(v_i^x) = d\mathcal{V}(v_i^x)$  for path vertices in a media.

Equipped with a definition of the path space and an accompanying measure we can now formulate a pixel measurement  $I_j$ , describing the value of pixel  $j$ , using the path integral formulation [172, 213] as:

$$I_j = \int_{\mathcal{P}} W_j(\bar{x}) f(\bar{x}) d\mu(\bar{x}), \quad (2.20)$$

where,  $f(\bar{x})$ , is the the *measurement contribution function*, defined by

$$f(\bar{x}) = L_e \left[ \prod_{i=0}^{k-1} G(v_i^x, v_{i+1}^x) T(v_i^x, v_{i+1}^x) \right] \left[ \prod_{i=1}^{k-1} \rho(v_i^x) \right], \quad (2.21)$$

where  $L_e = L_e(x_0 \rightarrow x_1)$  is the emitted radiance at the light source,  $\rho(v_i^x)$  denotes the scattering distribution function defined at path vertices,  $G(v_i^x, v_{i+1}^x)$  is the *generalized geometry term*, and  $T(v_i^x, v_{i+1}^x)$  is the transmittance, defined by equation (2.16), on the segments between path vertices. The generalized geometry term is defined by

$$G(v_i^x, v_{i+1}^x) = V(v_i^x, v_{i+1}^x) \frac{D(v_i^x, v_{i+1}^x) D(v_{i+1}^x, v_i^x)}{\|v_i^x - v_{i+1}^x\|^2}, \quad (2.22)$$

where  $V(v_i^x, v_{i+1}^x)$  is the binary visibility function, defined by equation (2.10), and

$$D(v_i^x, v_{i+1}^x) = \begin{cases} |n_{v_i^x} \cdot \omega| & : \text{if } v_i^x \text{ is located on a surface,} \\ 1 & : \text{if } v_i^x \text{ is located in a medium,} \end{cases} \quad (2.23)$$

<sup>1</sup> A formal definition of the path space measure can be found in [172, 213]

where  $n_{v_i^x}$  denotes the normal at  $v_i^x$  and  $\omega$  is the direction from  $v_i^x$  to  $v_{i+1}^x$ .

Similarly, the scattering function,  $\rho(v_i^x)$  is defined to be equal to the BRDF if  $v_i^x$  is on a surface, and as the product between the scattering coefficient  $\sigma_s$  and a phase function if  $v_i^x$  is in a medium.

The measurement contribution function,  $f_j(\bar{x})$ , used in the path integral formulation can be derived by recursively expanding the RTE and accounting for the appropriate changes of measure, a detailed derivation can be found in [95].

## 2.6 Simulating light transport

The RTE is a Fredholm integral equation of the second kind. Such integral equations appear in many fields of science and their mathematical properties have been studied extensively. For all but trivial cases these integral equations can not be solved analytically [176]. Instead, different approximations, often based on stochastic Monte Carlo methods, are used in order to solve them. In chapter 3, we will introduce techniques based on Monte Carlo methods in more detail.

Deterministic methods for approximating the light transport in a scene have also been developed. *Radiosity* methods [9, 72] are based on *finite element methods* and compute the light transport by discretizing the scene into a population of small patches for surfaces, and voxels for media. The illumination distribution is then computed by simulating the interactions among these discrete elements. However, for scenes with complex geometries and glossy materials radiosity solutions require complex algorithms. In many cases this makes radiosity less practical to use compared to for example Monte Carlo methods.

## Monte Carlo rendering

Monte Carlo methods constitute a class of numerical methods based on averaging a large number of random samples to produce *estimates* of the true solution. The history of Monte Carlo methods is closely related to the development of computers, which have enabled complex and previously intractable problems to be solved using simulation based approaches. In comparison to more conventional, deterministic, numerical methods, which typically have error bounds that grow exponentially with the problem dimension, Monte Carlo methods are particularly well suited for high dimensional problems. Since the 1940s, when the name "Monte Carlo" was coined by scientists working on the nuclear weapons projects in Los Alamos [150], these methods have been used extensively in numerous fields, including the physical sciences, engineering, medicine, statistics/machine learning, and computer science. In fact, the foundational theory upon which modern Monte Carlo methods rests is often highlighted as one of the ten most important algorithmic developments in the preceding century [53].

The development of Monte Carlo methods for solving transport problems dates back to the work of John von Neumann and Robert Richtmeyer in 1947 [180], who proposed to use one of the first electronic computers, the ENIAC, to simulate neutron transport phenomena. To the authors knowledge, the first Monte Carlo simulation of light transport (photons) was presented by Hayward and Hubble in 1954 [87]. In their paper they discussed the simulation of 67 photon trajectories through a homogenous slab of participating media using a desk calculator. This can be contrasted to modern Monte Carlo based rendering methods, simulating millions of light paths through complex scenes constituted

by millions of geometrical primitives. Today it is not uncommon to trace hundreds of millions of rays to compute a single image. Entertaining historical accounts of the early developments of Monte Carlo methods can be found in [60] and [149].

In this chapter, we first present the theoretical foundations of Monte Carlo methods in section 3.1. We then, in section 3.2, discuss how these methods can be used to simulate light transport, for the purpose of rendering. In particular, the presentation in this chapter focuses on Markov Chain Monte Carlo (MCMC) methods, which present a powerful alternative to more traditional Monte Carlo based rendering algorithms. In section 3.3 we discuss the contributions of Paper B, which presents a new MCMC method for rendering scenes with analytically intractable measurement contribution functions, for example scenes containing heterogenous participating media. Finally in section 3.4 we present some concluding remarks and discuss some possible venues for future work in Monte Carlo rendering.

## 3.1 Monte Carlo estimators

In this section, we discuss the theoretical foundations of Monte Carlo methods. We do not aim at a formal treatment of the subject, but instead we try to emphasize the general ideas and provide some intuition for different sampling methods. Throughout the chapter the reader is assumed to have a basic knowledge of probability theory and statistics. A good review of the subject can be found in many excellent textbooks, see for example [58]. More in depth discussion about the theoretical properties of Monte Carlo methods can be found in the excellent book by Owen [171].

### 3.1.1 Constructing estimates from random samples

Monte Carlo estimators are based on expressing a quantity of interest,  $\gamma$ , as the expected value of some integrable function,  $g(X)$ , of a random variable  $X$ , i.e.  $X$  and  $g(X)$  are defined so that  $E[g(X)] = \gamma$ . Let  $p(x)$  denote the distribution of  $X$ . An estimate of the quantity of interest can then be formed by sampling a sequence of independent random variables  $\{X_i\}_{i=1}^N$ , distributed according to  $p(x)$ , and computing the average:

$$\hat{\gamma}^N = \frac{1}{N} \sum_{i=1}^N g(X_i) \quad X_i \sim p(x). \quad (3.1)$$

This estimator is motivated by the *strong law of large numbers*, which states that this estimate converges *almost surely* to the quantity of interest as  $N$  goes to

infinity,

$$\mathbb{P}\left(\lim_{N \rightarrow \infty} \hat{\gamma}^N = \gamma\right) = 1, \quad (3.2)$$

in other words the estimator is *consistent*.

When the quantity of interest is a deterministic integral,  $\gamma = \int_D f(x)dx$ , we note that we can express it as an expectation using the following relation:

$$\gamma = \int_D f(x)dx = \int_D \frac{f(x)}{p(x)}p(x)dx = \mathbb{E}\left[\frac{f(X)}{p(X)}\right] \quad X \sim p(x), \quad (3.3)$$

where  $p(x)$  is an arbitrary probability distribution such that  $p(x) > 0$  when  $f(x) \neq 0$ . The integral can then be estimated by drawing random samples according to  $p(x)$  and then using  $g(x) = f(x)/p(x)$  in equation (3.1) to form the following Monte Carlo estimator:

$$\hat{\gamma} = \frac{1}{N} \sum_{i=1}^N \frac{f(X_i)}{p(X_i)} \quad X_i \sim p(x). \quad (3.4)$$

The use of random samples to estimate an deterministic integral in this fashion is often referred to as *Monte Carlo integration*. In this thesis, we focus our discussion on estimators of this form, as they are the ones employed in standard Monte Carlo rendering algorithms.

From the strong law of large numbers, we know that the error resulting from using Monte Carlo will go to zero as the number of random samples is increased. The estimator is also *unbiased* as

$$\mathbb{E}[\hat{\gamma}] = \frac{1}{N} \sum_{i=1}^N \mathbb{E}\left[\frac{f(X_i)}{p(X_i)}\right] = \frac{1}{N} \sum_{i=1}^N \int_D \frac{f(x)}{p(x)}p(x)dx = \gamma, \quad (3.5)$$

in other words the expected value of the Monte Carlo estimator, regardless of  $N$ , is the quantity of interest. The perhaps most important practical concern is, however, how fast the error goes to zero when increasing  $N$ . This property can be studied by considering the variance of the Monte Carlo estimator, computing the variance of equation (3.4) gives:

$$\text{Var}[\hat{\gamma}] = \frac{1}{N^2} \sum_{i=1}^N \text{Var}\left[\frac{f(X_i)}{p(X_i)}\right] = \frac{1}{N} \text{Var}\left[\frac{f(X)}{p(X)}\right] \quad (3.6)$$

The root mean squared error (RMSE) of  $\hat{\gamma}$  is:

$$\sqrt{\mathbb{E}[(\hat{\gamma} - \gamma)^2]} = \sqrt{\text{Var}[\hat{\gamma}]} = \frac{\sigma}{\sqrt{N}}, \quad (3.7)$$

where  $\sigma = \sqrt{\text{Var}\left[\frac{f(X)}{p(X)}\right]}$ . This means that the RMSE decrease as  $\frac{1}{\sqrt{N}}$  when  $N$  increases. A quite remarkable property of this result is that it does not depend

on the dimension  $d$  of the expectation/integral. In stark contrast, deterministic numerical quadrature rules generally has a error rate that depends on the dimension, for example applying Simpson's quadrature in  $d$  dimensions gives an error rate in the order of  $\frac{1}{N^{4/d}}$ , making such methods perform very poorly in high dimensions. Another striking property of the Monte Carlo error rate is that it does not directly depend on the smoothness properties of the involved quantities, for example the integrand,  $f(x)$ , when using Monte Carlo integration. These features is what often makes Monte Carlo the method of choice for high dimensional or non-smooth problems.

### 3.1.2 Importance Sampling

To reduce the variance of the Monte Carlo estimator of the quantity of interest,  $\gamma = \mathbb{E}\left[\frac{f(x)}{p(x)}\right]$ , a common technique is to employ *Importance Sampling*, where  $p(x)$  is set so that more samples are generated in regions where the function  $f(x)$  is large and less samples in regions where  $f(x)$  is small. This can greatly reduce the RMSE of the Monte Carlo estimator in applications where  $f(x)$  is close to zero for many values of  $x$ . This is common in, for example, rendering applications where only a few of all the possible light paths in the scene transport the majority of the energy reaching the camera.

To get a better understanding of how  $p(x)$  should be chosen in order to minimize the variance of the Monte Carlo estimator it is informative to study the expression for the variance in the following form:

$$\begin{aligned} \text{Var}[\hat{\gamma}] &= \frac{1}{N} \text{Var}\left[\frac{f(X)}{p(X)}\right] = \frac{1}{N} \left( \mathbb{E}\left[\frac{f(x)^2}{p(x)^2}\right] - \mathbb{E}\left[\frac{f(x)}{p(x)}\right]^2 \right) \\ &= \frac{1}{N} \left( \int_D \left(\frac{f(x)^2}{p(x)}\right) dx - \gamma^2 \right) = \frac{1}{N} \left( \int_D \left(\frac{(f(x) - p(x)\gamma)^2}{p(x)}\right) dx \right). \end{aligned} \quad (3.8)$$

From equation (3.8), it is easy to see that the variance of the estimator will be low when  $p(x)$  is roughly proportional to  $f(x)$ . We can also conclude that a poor choice of the proposal density  $p(x)$  can lead to estimators with a very high variance. Specifically, if  $p(x)$  decrease towards zero faster than  $f(x)^2$  as  $x$  moves away from it(s) mode(s), the variance of the estimate will be infinite. In practice, the most difficult criteria to fulfill, however, is that  $p(x)$  should be easy to evaluate (point-wise) and generate samples from. Indeed, if the evaluation of  $p(x)$  is computationally expensive to evaluate or sample from, it might be a better choice to instead use a simpler proposal distribution and average more samples.



Importance sampling techniques constitute a fundamental building block of modern rendering algorithms, and considerable research effort has been directed towards finding good proposal functions for sampling light transport paths. For example, when evaluating direct illumination, light that has only been reflected once in the scene before reaching the camera, a common choice is to set  $p(x)$  proportional to the BRDF of the surface. Another common choice is to sample outgoing paths from the reflection point according to the distribution of light sources in the scene. However, neither of these choices work well for scenes with localized light sources and glossy BRDFs, i.e. the surface scatter most of the incident light towards a small set of directions and only a few outgoing directions correspond to high energy light sources. To address this problem, it is common to use multiple importance sampling distributions simultaneously, i.e. choose  $p_m(x) = \alpha p_1(x) + (1 - \alpha)p_2(x)$ , with  $\alpha \in (0, 1)$ , to ensure that some samples are distributed in directions corresponding to both high reflectance and bright light sources. This technique is also known as *defensive mixture sampling* in statistics [170]. A slight, but guaranteed, improvement over the random mixture sampling technique is to use a fixed, deterministic, number of samples from  $p_1(x)$  and  $p_2(x)$  and combine them using a so called *Multiple importance sampling estimator* [170, 215].

### 3.1.3 Independent Sampling methods

In this section, we discuss some commonly used methods for generating independent samples from a probability distribution  $p(x)$ .

#### Generating uniform random variables

In general all sampling methods are based on the use of random numbers uniformly distributed in the unit interval,  $U \sim \mathcal{U}(0, 1)$ . On a computer such numbers are generated using deterministic pseudo-random number generators. These generators are designed so that the resulting number streams pass statistical tests to ensure that the numbers are uncorrelated and uniformly distributed [171]. It is thus very unlikely that the subtle correlations between the values in the sequence due to their deterministic generation have a noticeable effect on the simulation results. In this thesis we therefore assume that we have access to an infinite stream of  $U$ :s uniformly distributed on the unit interval.

### Inverse transform sampling

A direct method for sampling a general random variable  $X$  with distribution  $p(x)$  is to make use of its cumulative distribution function:

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x p(x') dx'. \quad (3.9)$$

Now, if we generate  $U \sim \mathcal{U}(0,1)$  and set  $X = F^{-1}(U)$ , then:

$$\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x), \quad (3.10)$$

thus  $X$  is distributed according to  $p(x) = \frac{d}{dx}F(x)$ .

This method can directly be generalized to sample from some distribution  $p_Y(y) = \frac{d}{dy}F_Y(y)$  given samples from a distribution  $p_X(x) = \frac{d}{dx}F_X(x)$  by using the following relation:

$$Y = F_Y^{-1}(F_X(X)), \quad (3.11)$$

as  $F_X(X) \sim \mathcal{U}(0,1)$ .

For example, the inverse transform sampling method can directly be applied to sample from the exponential distribution, related to the transmittance between two points in a homogenous media,

$$p(x; \lambda) = \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right) \quad x \geq 0, \quad (3.12)$$

using:

$$X = F^{-1}(U; \lambda) = -\lambda \ln(1 - U) = -\lambda \ln(U), \quad (3.13)$$

since  $1 - U$  is distributed as  $U$ .

### Rejection Sampling

Sometimes one cannot obtain a computable transform that takes uniform random numbers  $U \sim \mathcal{U}(0,1)$  and produces a random variable with the desired distribution  $p(x)$ . For example, the inverse of the cumulative distribution function may not be known. In this case, an approach known as *acceptance-rejection sampling*, or *rejection sampling* can be used. To use this method, we specify another simpler distribution that we can draw samples from directly,  $g(x)$ , and a constant  $C$  such that

$$p(x) \leq g(x)C, \quad \forall x. \quad (3.14)$$

We can then generate candidate samples  $Y \sim g(y)$  and either accept them or reject them with the acceptance-probability  $a(Y) = \frac{p(Y)}{g(Y)C}$ . The accepted

candidates will then be distributed according to  $p(x)$ . The efficiency of rejection sampling depends on how many candidate samples that have to be generated for each accepted sample. To get good results, it is therefore important to choose  $g(x)$  so that it is a close approximation of  $p(x)$ . This can often be difficult in practice, especially for high-dimensional spaces.

### Sequential sampling

To sample a multivariate random variable,  $\mathbf{X}$ , following a distribution,  $p(x_1, x_2, \dots, x_d)$ , it is often practical to sample each dimension sequentially. In the simplest case, the joint distribution is separable and can directly be described by product of 1D densities  $p(x_1, x_2, \dots, x_d) = p(x_1)p(x_2), \dots, p(x_d)$ . In this case, each component can be sampled independently. However, many useful densities are not separable, so other factorizations have to be considered. A generally applicable factorization is to use the decomposition of the joint distribution,  $p(\mathbf{x})$  as

$$p(\mathbf{x}) = p(x_1) \prod_{i=1}^N p(x_i | \mathbf{x}_{1:i-1}), \quad (3.15)$$

where  $p(x_1) = \int p(\mathbf{x}) dx_{2:d}$  is the marginal distribution of  $X_1$ , and  $p(x_i | \mathbf{x}_{1:i-1})$  denotes the conditional distribution of  $X_i$  given  $X_{1:i-1}$ . To apply sequential sampling we first sample  $X_1$  from the marginal and then each subsequent  $X_i$  from the conditional distribution given the other sampled values.

Sequential sampling strategies have proven to be extremely useful in many applications, for example in computational statistics and related fields [188]. Most traditional Monte Carlo rendering methods are also based on this type of sequential sampling strategy.

### Transformation of random variables

Many practical sampling methods are based on transformations of random variables. This is a very useful concept in many rendering applications. Here, we review the basic relationships used to determine the distributions of transformed random variables [58]. Assume that we have a random variable  $X$  distributed according to  $p_X(x)$ , and some strictly monotonic function  $y(x)$ , then  $Y = y(X)$  is distributed according to

$$p_Y(y) = \left| \frac{d}{dx} y(x) \right|^{-1} p_X(x). \quad (3.16)$$

In the general case if  $X$  is a  $d$ -dimensional random variable, and  $T(x)$  is a bijective function, then  $Y = T(X) = (T_1(X), T_2(X), \dots, T_d(X))$  is distributed according to

$$p_Y(y) = |J_T|^{-1} p_X(x) \quad (3.17)$$

where  $|J_T|$  denotes the the absolute value of the determinant of the Jacobian matrix:

$$|J_T| = \det \begin{bmatrix} \frac{\partial T_1}{\partial x_1} & \frac{\partial T_1}{\partial x_2} & \dots & \frac{\partial T_1}{\partial x_d} \\ \frac{\partial T_2}{\partial x_1} & \frac{\partial T_2}{\partial x_2} & \dots & \frac{\partial T_2}{\partial x_d} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial T_d}{\partial x_1} & \frac{\partial T_d}{\partial x_2} & \dots & \frac{\partial T_d}{\partial x_d} \end{bmatrix}. \quad (3.18)$$

For example, in a variety of rendering applications we are interested in sampling from a cosine distribution on the unit hemisphere,

$$p(\omega) \propto (n \cdot \omega) = \cos(\theta). \quad (3.19)$$

An efficient technique to sample from this distribution is *Malley's method* which is based on first generating uniformly distributed points on the unit disk and then projecting these points up to the hemisphere above them. These points will then correspond to directions that are cosine distributed on the hemisphere. This method is directly motivated by considering the Jacobian of the transformation from polar coordinates on the unit disk to spherical coordinates on the unit hemisphere, a detailed account of the algorithm is given in [174].

A similar transformation is considered in the derivation of the importance sampling methods of the BRDF functions presented in paper H.

### 3.1.4 Markov chain Monte Carlo methods

A problem with the independent sampling methods discussed in the previous section is that they can be very inefficient in high dimensions. In the context of Monte Carlo integration this means that it can be difficult to sample from a good proposal distribution that is proportional to the integrand of interest. A powerful alternative to independent sampling strategies is Markov Chain Monte Carlo, or *MCMC*, methods. These methods enable sampling from arbitrary probability distributions,  $p(x)$ , that may be impractical to sample from with other techniques, by constructing a Markov chain which admits  $p(x)$  as its unique stationary distribution. These methods no longer produce independent samples from the target distribution but instead generate statistically *correlated* samples. However, by using carefully constructed Markov chains, empirical averages of the obtained samples can be shown to converge to the desired quantity. Quite remarkably, it is possible to construct such Markov chains using simple transition rules that only require that the target distribution can be evaluated up to a normalization constant. Specifically, it is not required to compute the inverse of the cumulative distribution function or find a suitable proposal

distribution with a known bounding constant, which both are operations that can cause problems for other sampling methods.

To define a Markov chain it is necessary to specify a *transition rule*: a stochastic rule for selecting a new state  $x_{i+1}$  given a current state  $x_i$ . The *Metropolis-Hastings* algorithm [151]<sup>1</sup>[85] is the gold-standard method to define such a transition rule so that the resulting Markov chain admits the target distribution  $p(x) = \frac{\tilde{p}(x)}{Z}$  as its stationary distribution. Here  $Z = \int \tilde{p}(x)dx$  denotes a possibly unknown normalizing constant. This algorithm requires the choice of a conditional density  $q(x_{i+1}|x_i)$  known as the *proposal* distribution. The transition from the current state  $X_i$  of the chain to the next is then computed as:

**Metropolis-Hastings algorithm:**

Given  $x_i$ ,

1. Sample  $y_i \sim q(y|x_i)$
2. Set

$$x_{i+1} = \begin{cases} y_i & : \text{with probability } r(x_i, y_i) \\ x_i & : \text{with probability } 1 - r(x_i, y_i) \end{cases}$$

where

$$r(x, y) = \min \left[ \frac{\tilde{p}(y)q(x|y)}{\tilde{p}(x)q(y|x)}, 1 \right]$$

Note that the normalizing constant of the target distribution does not need to be evaluated in order to compute the transitions. The same holds true for the proposal density as any factors that are shared between  $q(x|y)$  and  $q(y|x)$  cancel out.

Given an initial state  $X_1 = x_1$  the Metropolis-Hastings algorithm produces correlated samples  $X_2, \dots, X_N$  by sequentially applying the transition rule specified above. Under mild conditions, it can be shown that the resulting Markov chain admits a stationary distribution that is equal to the target distribution  $p(x)$  [85, 136], i.e. if the chain is run long enough the samples will eventually be distributed according to  $p(x)$ . As the first samples obtained from the simulated Markov chain will be strongly dependent on the initial state, a common practice is to disregard the initial samples. This initial set of samples are often referred to as the *burn-in*. However, the *ergodic theorem* [181] tells us that under some mild conditions regardless of the starting value  $x_1$ , the empirical average of the

<sup>1</sup> This 4-page article has over 30 000 citations according to Google Scholar

obtained samples

$$\hat{\gamma}_{MH}^N = \frac{1}{N} \sum_{i=1}^N g(X_i), \quad (3.20)$$

is a consistent estimator of  $\gamma = E_p[g(x)]$ , i.e.

$$\mathbb{P}\left(\lim_{N \rightarrow \infty} \hat{\gamma}_{MH}^N = \gamma\right) = 1. \quad (3.21)$$

In particular, the ergodic theorem assumes that the Markov chain is *ergodic*. An *ergodic* Markov chain should be able to reach any state from any other state, which can easily be assured in practice by using a proposal density that assigns a non-zero probability of sampling any state in the state space.

A more formal treatment on the properties of the Metropolis-Hastings algorithm and MCMC methods in general can be found in for example [199] and [181].

An important aspect of the performance of the Metropolis-Hastings algorithm is the choice of the proposal density. On one hand, we seek an algorithm that captures a representative subset of the relevant parts of the state space by moving around quickly in the space. On the other hand, we would also like an algorithm that can take small steps so that it can locally explore a region of the state space where  $p(x)$  takes on large values. Taking too large steps can cause the chain to move away from local maxima which leads to high rejection rates, and thus long repetitions of the current state. However, if the chain takes too short steps the chain can not move around quickly in the state space and might miss important regions completely. It is therefore vital to use suitable proposal densities that ensure that both larger and smaller moves are used to simulate the chain. In practice this is often accomplished by mixing local and global moves. In section 3.2.3, we discuss the application of MCMC methods to rendering, and show how such proposal densities can be constructed to explore the space of all valid light paths in a scene.

## 3.2 Monte Carlo light transport simulation

In this section, we provide an overview of standard Monte Carlo rendering algorithms. More in-depth descriptions of these algorithms can for example be found in the book by Pharr and Humphreys [174].

### 3.2.1 Path tracing methods

Monte Carlo methods have a long history in computer graphics, dating back to the work of Kajiya [101] and Cook [40]. In contrast to other techniques available

at the time, these algorithms enabled researchers and practitioners to simulate effects that had been difficult to generate previously, such as: depth-of-field, motion blur, glossy reflections and global illumination.

In terms of the path integral formulation introduced in section 2.5, the goal of a rendering algorithm is to compute the pixel values by estimating the corresponding integrals over path space, introduced in equation (2.20),

$$I_j = \int_{\mathcal{P}} W_j(\bar{x}) f(\bar{x}) d\mu(\bar{x}). \quad (3.22)$$

Specifically, the algorithms are designed to sample light paths  $\bar{X}_i$  according to some probability function  $p(\bar{x})$  defined on the path space, and then evaluating the resulting Monte Carlo estimator:

$$\hat{I}_j = \frac{1}{N} \sum_{i=1}^N \frac{W_j(\bar{X}_i) f(\bar{X}_i)}{p(\bar{X}_i)} \quad \bar{X}_i \sim p(\bar{x}). \quad (3.23)$$

Although not all Monte Carlo rendering algorithms were originally derived in this form, which is a relatively recent construction, it is often informative to study and compare them using this formulation [78, 213]. In the path integral formulation, a rendering algorithm can be viewed as a sampling mechanism with the goal of constructing a proper probability distribution  $p(\bar{x})$  that produces a Monte Carlo estimator with high efficiency. To ensure that the resulting estimator has low variance,  $p(\bar{x})$  should be roughly proportional to the integrand (see section 3.1.2). However, as the measurement contribution function,  $f(\bar{x})$ , is a complicated function consisting of many terms, and the domain of the integrand is generally infinite (arbitrary long light paths may exist), it is generally not an easy task to design a suitable  $p(\bar{x})$ .

### Path tracing

One of the most fundamental Monte Carlo rendering algorithms, *Path tracing*, constructs  $p(\bar{x})$  by sequentially sampling path vertices starting from the sensor. First, a starting location on the image plane is sampled. Then the algorithm proceeds by generating a random walk into the scene. To sample the next vertex along the path, a conditional distribution function is used to sample an outgoing direction from the current vertex. If the resulting ray intersects some participating media, a media vertex is sampled at a random distance proportional to the transmittance of the media, otherwise the next surface intersection is found along the ray. The conditional distribution function used to sample an outgoing direction from the current vertex is set to be proportional to the local BRDF if the current vertex is at a surface or the local phase function if the current vertex is in media. The random walk procedure proceeds by

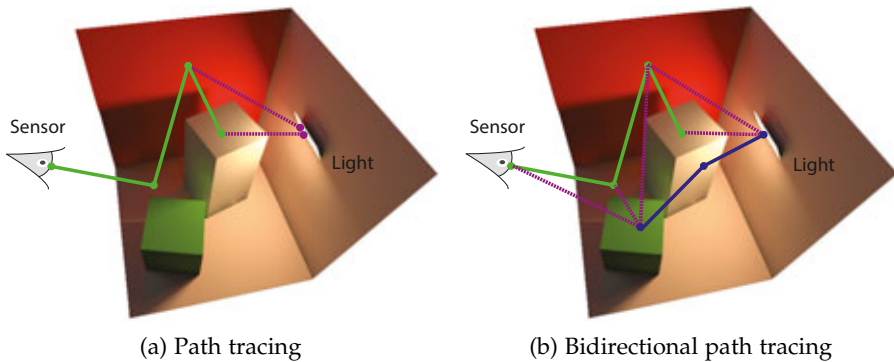


Figure 3.1: Illustration of path tracing and bidirectional path tracing. **a)** Path tracing constructs a random walk into the scene starting at the sensor (green line). At each intersection next event estimation tries to find a valid (visible) connection to a light source (dashed purple lines). **b)** Bidirectional path tracing, construct two random walks. One starting at the sensor (green line) and one starting at a light source in the scene (blue line). All possible connections between vertices on the two paths are then considered. Valid connection paths (dashed purple lines), not intersecting any scene geometry, are considered in the final estimator.

adding vertices in this fashion until a random termination criteria is met, also known as *Russian roulette* [102]. The random termination criteria is designed so that it ensures that shorter paths (say paths of 3 – 5 vertices) are sampled more often than longer paths, which are often less significant.

In practice, the basic path tracing algorithm is often improved by using a technique known as *next event estimation*. Using this method, the path tracer also creates additional connections to light sources in the scene at every sampled surface or volume intersection of the path. Each original path from the camera can thus produce many valid light path samples with varying length. As the explicit connections are only valid if there are no occluding objects between the vertex and the light source, path tracing methods generally only work well in scenes where the light sources are directly visible from sampled vertices of the random walk starting at the camera. This can be a major problem in many scenes which are indirectly illuminated, for example, an interior room lit indirectly through occluded light sources. The presence of specular or glossy materials can also be a problem, as they make it difficult to find light paths that ends up at light sources. For example, rendering a scene where the light source is located inside a glass enclosure often results in very noisy results using these methods.



### Bidirectional path tracing

An alternative to traditional Path tracing is to use Bidirectional path tracing [125, 214]. A bidirectional path tracer constructs  $p(\bar{x})$  by first generating two independent random walks. One starting from the camera and one starting from light sources in the scene. A large number of valid light paths are then generated by explicitly connecting each sampled interaction vertex along the two subpaths, creating a number of different lengths for the generated paths. This sampling strategy can lead to estimators with significantly less variance than for unidirectional path tracing, as it enables a larger variety of light paths to be sampled. Furthermore, by combining the sampled paths using multiple importance sampling, robust estimates can be obtained for scenes where one of the sampled connection lengths is much better than the others. However, bidirectional path tracing can still be inefficient when rendering scenes containing very glossy transfer (light paths reflected through glossy materials) or scenes that contain prominent *specular-diffuse-specular* reflection geometries in which the light path is scattered first by a specular reflection, then a diffuse reflection and then another specular reflection.

### 3.2.2 Caching methods

The path tracing methods discussed this far are *unbiased*, in other words they produce images which are correct on average. Biased Monte-Carlo methods relaxes this assumption and use different type of caching operations which assume that the incident (or outgoing) radiance at a point in the scene is similar to that of nearby points. Prominent examples of such methods are irradiance caching [97, 221], photon mapping [77, 98], and instant radiosity [43, 111] and their variants. While these methods can be extremely effective for some scenes, they generally produce biased results, which makes error analysis difficult. Similarly to traditional path tracing methods, they also have problems to handle glossy light transport paths [174].

### 3.2.3 Metropolis Light Transport

In contrast to other rendering methods, Metropolis Light Transport (MLT) [216] algorithms are based on using MCMC to sample light paths. Instead of using sequentially constructed approximations as in traditional path tracing, the use of MCMC enables these methods to sample from a probability distribution over light paths exactly proportional to their *contribution* to the image plane. Specifically, for a given *scalar contribution function*  $L(\bar{x})$ , which we will define shortly, MLT generates a sequence of samples  $\{\bar{X}_i\}_{i=1}^N$  from a probability distribution

proportional to  $L$ ,

$$p(\bar{x}) = \frac{L(\bar{x})}{\int_{\mathcal{P}} L(\bar{x}) d\mu(\bar{x})} \quad (3.24)$$

which allows us to estimate the value of a pixel,  $I_j$ , using a Monte Carlo estimator as:

$$\hat{I}_j = \frac{1}{N} \sum_{i=1}^N \frac{W_j(\bar{X}_i) f(\bar{X}_i)}{p(\bar{X}_i)} = \frac{Z}{N} \sum_{i=1}^N \frac{W_j(\bar{X}_i) f(\bar{X}_i)}{L(\bar{X}_i)} \quad (3.25)$$

where

$$Z = \int_{\mathcal{P}} L(\bar{x}) d\mu(\bar{x}) \quad (3.26)$$

is a normalizing constant that is straightforward to compute using traditional path tracing methods.

A major benefit of MLT algorithms is that by using MCMC to generate correlated light path samples they enable local exploration of the path space. Once a high contribution path has been found, the algorithm can use this knowledge and generate another path that is similar to the current path by using small perturbations. The use of the Metropolis-Hastings algorithm ensures that, under some mild conditions on the employed perturbations, the generated paths are in fact distributed as the target distribution  $p(\bar{x})$ . In contrast, traditional path tracing methods samples paths independently, this means that there is no straightforward way to reuse the knowledge of a high contribution path. This ability often makes MLT algorithms perform much better than traditional path tracing methods in scenes with complex visibility constraints or glossy materials. This is illustrated in figure 3.2.

To minimize the variance of the estimator, the contribution function,  $L(\bar{x})$ , should be chosen to be as close as possible to the integrand. However, to allow light paths to be distributed over all pixels, it is common to only consider the part of the integrand that is not dependent on the specific pixel, i.e. the measurement function  $f(\bar{x})$ . This allows us to sample light paths according to their contribution to the image plane, and then increment the pixel values whose corresponding pixel filters are positive. Using this approach one path can contribute to several pixel integrals. Furthermore, to efficiently generate renderings with color,  $f(\bar{x})$  is treated as a spectrally valued function, for example a common choice is to set  $f(\bar{x}) = [f_R(\bar{x}), f_G(\bar{x}), f_B(\bar{x})]$  to describe the three primary colors Red, Green and Blue. This allows us to compute all wavelength bands simultaneously for each sampled path. However, this prohibits us to set the scalar valued contribution function,  $L(\bar{x})$ , to  $f(\bar{x})$  directly. A common choice in practice is instead to set  $L(\bar{x})$  to be equal to the luminance of  $f(\bar{x})$ ,

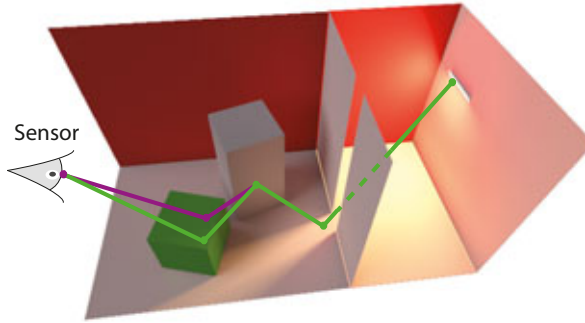


Figure 3.2: This scene represents a difficult case for traditional path tracing methods as, due to the complex visibility constraints, many paths have to be generated before finding one that successfully connects the sensor to the light source. In contrast to traditional path tracing methods, MLT generates correlated light path samples. This means that once a high contribution path (green line) has been found similar paths can be explored by simulating small perturbations (purple line) of the current path.

for example, when using the *sRGB* color space [2] this corresponds to setting  $L(\bar{x}) = f_L(\bar{x}) = 0.2126f_R(\bar{x}) + 0.7152f_G(\bar{x}) + 0.0722f_B(\bar{x})$ . Other choices of the scalar contribution function have also been explored, for example based on the properties of the Human Visual System [91].

To generate samples,  $\{\bar{X}_i\}_{i=1}^N$ , MLT uses the Metropolis-Hastings algorithm with a proposal function  $q(\bar{y}|\bar{x}_i)$ , in the context of MLT often referred to as a mutation or perturbation proposal. The initial state of the underlying Markov chain is normally generated from a path obtained using traditional path tracing methods, for example bidirectional path tracing. Usually a large set of *candidate paths* are first generated. To account for the ratio between the contribution function  $L(\bar{x})$  and the probability of sampling a candidate path,  $p(\bar{x})$  a sampling weight is computed for each path. An initial *seed path* is then sampled from the collection of all candidate paths with a probability that is proportional to the candidate weights (in other fields this procedure is also known as sequential importance resampling [136]). This ensures that the initial seed path is also distributed proportional to  $L(\bar{x})$  and increases the efficiency of resulting estimators, by effectively avoiding *burn-in*.

In practice, one seed path is chosen per available processing core. This enables several Markov chains to be run in parallel. The final image is then obtained

by averaging the results from the different cores. The set of initial candidate paths can also be used to estimate the normalizing constant,  $Z$ , using a standard Monte Carlo estimator:

$$\hat{Z} = \frac{1}{N} \sum_{i=1}^N \frac{L(\bar{X}_i)}{p(\bar{X}_i)} \quad \bar{X}_i \sim p(\bar{x}) \quad (3.27)$$

The algorithm then proceeds by repeatedly applying a tentative path mutation/perturbation to the current path. The current path is then updated to the proposed path according to the transition rule specified by the Metropolis-Hastings algorithm to ensure that the resulting samples are distributed according to  $p(\bar{x}) = \frac{L(\bar{x})}{\int_{\mathcal{P}} L(\bar{x}) d\mu(\bar{x})}$ . The updated, current, path is then used to increment the pixel estimators for which the pixel filter has a non-zero support, i.e.  $W_j(\bar{X}_{i+1}) > 0$ .

The complete MLT algorithm can be summarized as:

#### Metropolis Light Transport:

1. Generate a set of candidate paths using bidirectional path tracing
2. Compute  $\hat{Z}$  using the candidate paths, using equation (3.27)
3. Sample  $\bar{x}_1$  from the set of candidate paths
4. for  $i = 2 \dots N$
5. Sample a perturbed path  $\bar{y}_i \sim q(\bar{y}|\bar{x}_i)$
6. Set

$$x_{i+1} = \begin{cases} y_i & : \text{with probability } r(x_i, y_i) \\ x_i & : \text{with probability } 1 - r(x_i, y_i) \end{cases}$$

where

$$r(\bar{x}_i, \bar{y}_i) = \min \left[ \frac{L(\bar{y}_i)q(\bar{x}_i|\bar{y}_i)}{L(\bar{x}_i)q(\bar{y}_i|\bar{x}_i)}, 1 \right]$$

7. Update Monte Carlo estimates  $\hat{I}_j$  for  $j$  where  $W_j(\bar{x}_{i+1}) > 0$ .

An important aspect in practice, is the choice of proposal functions,  $q$ , used to mutate the current path. In the original implementation by Veach and Guibas [216] a mixture proposal consisting of five different tentative mutations were considered:

1. Bidirectional mutation: replaces a segment of an existing path using a bidirectional path tracing method. This mutation serves as a global proposal and has a non-zero probability of sampling any possible path in the space.

2. Lens subpath mutation: replaces a path segment between the lens and the first non-specular vertex with a new segment. This enables the end of the path to move locally on the image plane.
3. Lens perturbation: perturbs the outgoing direction at the camera and propagates the path until the first non-specular material is encountered. The modified part of the path is then explicitly connected to the rest of the old path.
4. Caustic perturbation: this proposal works like the Lens perturbation, but starting from a light source.
5. Multi-chain perturbation: is designed to handle mutations where there is a chain of specular reflections. The proposal uses an initial lens perturbation, but then a sequence of additional perturbations is performed until a connection to the rest of the old path is possible.

Recently, several other proposal functions have been proposed in the literature. Jakob and Marschner [96] introduced a *manifold exploration* proposal that improves the sampling of specular or glossy paths. This method is based on enforcing Fermat's principle as a constraint on specular path vertices after performing a random perturbation of some other vertex along the path. This perturbation strategy can be seen as a replacement of the Lens, Caustic and Multi-chain perturbations. In practice, it often provides considerable improvement in scenes with glossy transport including caustics. Other specialized proposals for glossy scenes has also been developed [82, 104].

Figure 3.3 shows a comparison between renderings of the classical scene by Veach [213], filled with a thin homogenous participating media, using unidirectional Path tracing, Bidirectional path tracing, Photon mapping and Metropolis Light Transport. The rendering generated using path tracing suffers from extensive noise. The Bidirectional path tracer performs slightly better, but has problems with handling the caustics from the glass egg. The Photon mapping algorithm is better at rendering the caustics but produces splotchy renderings that biases the result. The Metropolis Light Transport [216] algorithm with the manifold proposal [96] tends to produce significantly better result for scenes like this. However, note that for some simple scenes, where the light sources are directly visible from surfaces in the scene, path tracing methods can be more efficient than MLT methods as they do not require the computational overhead of path mutations and the computation of acceptance probabilities.

### Primary Sample Space MLT (PSSMLT)

A variant of MLT, which can be simpler to implement in practice is the *Primary Sample Space MLT* (PSSMLT) algorithm proposed by Kelemen et al. [110]. Instead of directly proposing changes to the path vertices, this method is based on

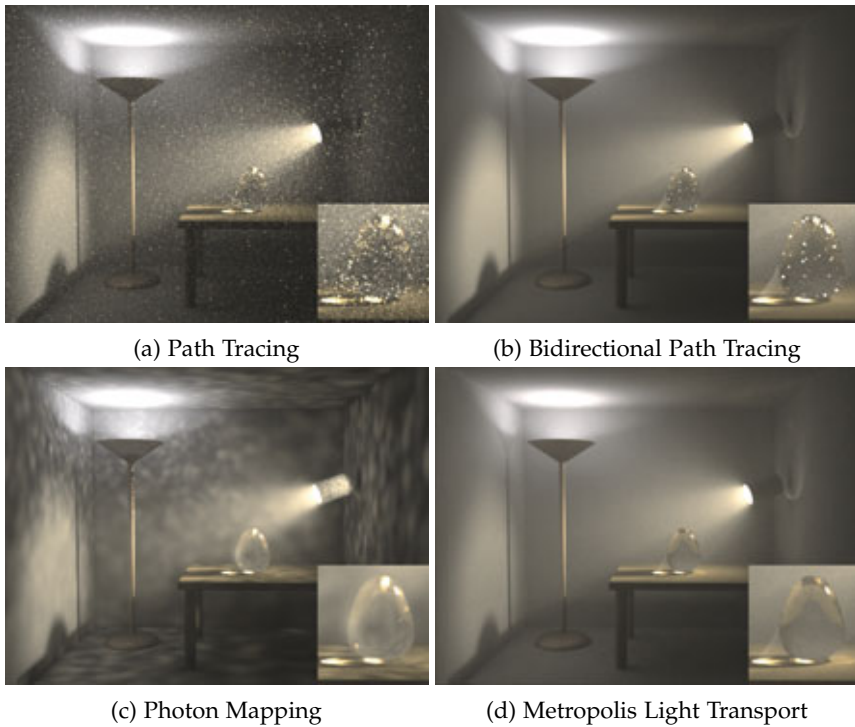


Figure 3.3: Renderings of the classical scene by Veach [213] filled with *homogeneous* (constant density) participating media. MCMC based rendering algorithms such as Metropolis Light Transport tend to result in lower image noise for this type of scene. The images were all rendered using the Mitsuba render [94] in five minutes using 64 Amazon EC2 vCPUs.

perturbing the sequence of uniform random numbers used by a path tracing type algorithm to sample a path in the scene. The key insight of this method is that a standard path tracing algorithm (all based on some form of sequential application of inverse transform sampling) can be viewed as a *deterministic* mapping from a sequence of uniform random numbers,  $\bar{U} = U_1, U_2, \dots, U_M$  to a valid light path,  $\bar{X}$  in  $\mathcal{P}$ . Denoting this mapping from uniform random numbers to a path with  $\psi : \mathcal{U}^M \rightarrow \mathcal{P}$ , where  $M$  can in theory be infinite but in practice limited by using random truncation, this enables us to rewrite the integral

defining the pixel values as

$$\begin{aligned} I_j &= \int_{\mathcal{P}} W_j(\bar{x}) f(\bar{x}) d\mu(\bar{x}) \\ &= \int_{\mathcal{U}^M} W_j(\psi(\bar{U})) \frac{f(\psi(\bar{U}))}{p(\psi(\bar{U}))} d\bar{u} \end{aligned} \quad (3.28)$$

where  $p(\bar{x})$  is the path sampling probability of the underlying path tracer, for more details we refer to [110]. This integral can now be approximated using MCMC by sampling proportional to  $\frac{f(\psi(\bar{U}))}{p(\psi(\bar{U}))}$ . The *crucial* difference compared to the original MLT algorithm is that the proposals are now performed on the *primary sample space*  $\mathcal{U}^M$  instead of directly in path space. This approach can be significantly easier to implement, as the algorithm can be implemented around existing path tracing code. It also allows for symmetric proposal distributions which makes the evaluation of the Metropolis-Hastings acceptance probability easier to compute.

The major limitation of all algorithms based on sampling in the Primary Sample Space is that due to the black-box nature of the mapping  $\psi : \mathcal{U}^M \rightarrow \mathcal{P}$  it can be hard to predict how small changes in one of the uniform random variables manifests itself in path space. This makes it hard to define specialized proposal distributions that take into account the global structure of the path. For example, a benefit of using the original perturbation methods proposed by Veach and Guibas [216] is that it is possible to change the length of a path by inserting or removing vertices, which can not directly be done in Primary Sample Space.

### Energy Redistribution path tracing (ERPT)

The methods discussed this far are all based on generating a few Markov chains, usually one per core, each generating a long sequence of samples. A problem with this approach is that it can be hard to stratify the samples over the image plane, which is often preferable for generating perceptually pleasing images. The lens subpath mutation, discussed above, tries to redistribute the samples over the image plane, but often lead to rejections. The *Energy Redistribution path tracing* (ERPT) algorithm [39] is instead based on running a large number of shorter Markov chains seeded from a initial set of samples that is well stratified over the image plane, for example one path per pixel. These samples can for example be obtained using bidirectional path tracing. This approach is similar to the *Generalized Importance Sampling* method developed independently in the statistics literature [141, 181], where MCMC methods are used to improve importance sampling estimators. As in the original MLT algorithm each Markov chain can use any perturbation rule, e.g. the lens, caustic and multi-chain perturbations. However, if the initial samples used to seed the Markov chains are distributed according to the target, the Markov

chains will preserve that distribution as their stationary distribution even if the transition rule is not ergodic (i.e. even if it can not sample some part of the path space). This enables the algorithm to generate Markov chains without using the costly bidirectional and lens subpath mutations. The ERPT algorithm thus explores paths in a more local fashion than other MCMC rendering algorithms, which often leads to better stratification of the samples on the image plane.

### 3.3 Contributions

A necessary requirement of the MLT algorithm is that the contribution function  $L(\bar{x})$  can be evaluated *exactly* for a given light path  $\bar{x}$ . This requirement follows directly from the use of the Metropolis-Hastings algorithm for constructing the transition rule of the underlying Markov chain. The main contributions of paper B is a new MLT variant that uses *approximations* of the contribution function,  $\hat{L}(\bar{x})$ , in place of the exact quantity,  $L(\bar{x})$ . In the paper we show that as long as the approximation  $\hat{L}(\bar{x})$  is a non-negative and unbiased estimator of the contribution function the simulated light paths are in fact *exactly* distributed according to the target distribution  $p(\bar{x}) = \frac{L(\bar{x})}{\int_{\mathcal{P}} L(\bar{x}) d\mu(\bar{x})}$ . The proposed method is inspired by *exact-approximation* methods, recently developed in the statistics literature for Bayesian inference problems [12, 44, 135]. The resulting algorithm lets us not only increase the flexibility and applicability of existing MCMC methods, but also improve their efficiency by using computationally cheaper approximations in place of the exact contribution function.

In section 3.3.1, we discuss the general form of the proposed algorithm derived using the *pseudo-marginal* MCMC approach recently developed in the statistics literature [11]. We then demonstrate, in section 3.3.2, how the algorithm can be used to render *heterogeneous* participating media, where the transmittance between two path is analytically intractable.

#### 3.3.1 Pseudo-marginal Metropolis Light Transport

The original MLT algorithm is based on simulating a Markov chain on the path space,  $\mathcal{P}$ , generating samples  $X_1, X_2, \dots, X_N$ . The Markov chain is specified by using the Metropolis-Hastings transition rule, which requires the evaluation of the acceptance probability, discussed in section 3.2.3, defined by:

$$r(\bar{x}_i, \bar{y}_i) = \min \left[ \frac{L(\bar{y}_i)q(\bar{x}_i|\bar{y}_i)}{L(\bar{x}_i)q(\bar{y}_i|\bar{x}_i)}, 1 \right]$$

where we need to be able to evaluate  $L(\bar{x})$  for any sampled path  $\bar{x}$ . In applications where  $L(\bar{x})$  is impossible, or very expensive, to evaluate exactly an intuitive idea is to replace  $L(\bar{x})$  with an approximation  $\hat{L}(\bar{x})$  when evaluating



the Metropolis-Hastings acceptance ratio. However, without further analysis it is unclear under what conditions the resulting Markov chain admits a stationary distribution, and what this stationary distribution would be.

In paper B, we propose to employ a pseudo-marginal MCMC construction [11, 24]. This allows us to generate a Markov chain that admits the desired distribution  $p(\bar{x}) = \frac{L(\bar{x})}{\int_{\mathcal{P}} L(\bar{x}) d\mu(\bar{x})}$  as its stationary distribution although we only have access to a *non-negative and unbiased* estimate of the contribution function  $\hat{L}(\bar{x})$ . The pseudo-marginal construction is based on the idea of introducing an auxiliary variable,  $V$ , representing all the random quantities generated to construct the estimator  $\hat{L}_v(\bar{x})$ . This random variable can be assumed to follow some, possibly unknown, probability distribution  $g(v|\bar{x})$  defined on some space  $\mathcal{V}$ . One can then show that a Markov chain on the extended space  $\mathcal{P} \times \mathcal{V}$  admits a stationary distribution which has a marginal distribution that is the desired distribution,  $p(\bar{x}) = \frac{L(\bar{x})}{\int_{\mathcal{P}} L(\bar{x}) d\mu(\bar{x})}$ . Furthermore, when generating this Markov chain on the extended space, using the Metropolis-Hastings algorithm, the acceptance probability simplifies to :

$$r(\bar{x}_i, \bar{y}_i) = \min \left[ \frac{\hat{L}_v(\bar{y}_i)q(\bar{x}_i|\bar{y}_i)}{\hat{L}_v(\bar{x}_i)q(\bar{y}_i|\bar{x}_i)}, 1 \right]$$

For more details the reader is referred to section 3 in paper B.

The *Pseudo-marginal Metropolis Light Transport* algorithm, proposed in paper B, is based on using this construction. This allows us to construct consistent estimators of the pixel values  $I_j$  where we only need to have access to a non-negative and unbiased approximation of the exact contribution function. By using the extended space, the resulting marginal samples,  $X_1, X_2, \dots, X_N$ , are generally more correlated compared to the original MLT algorithm, which can cause a slight increase of the variance of the resulting Monte Carlo estimators,  $\hat{I}_j$ . However, as we will show later, in practice the resulting increase in variance is often negligible compared to the gains obtained from being able to generate longer Markov chains in the same computational budget. By modifying the underlying Markov chains generated in PSSMLT, ERPT and other MLT variants, similar pseudo-marginal versions can also be derived for these methods.

The proposed Pseudo-marginal MLT algorithm can be summarize in the following algorithm, the lines marked in blue highlights the introduced changes compared to original MLT:

**Pseudo-Marginal Metropolis Light Transport:**

1. Generate a set of candidate paths using bidirectional path tracing
2. Compute  $\widehat{Z}$  using the candidate paths
3. Sample  $\bar{x}_1$  from the set of candidate paths
4. Sample  $v_1 \sim g(v|\bar{x}_1)$  and construct  $\widehat{L}_{v_1}(\bar{x}_1)$
5. for  $i = 2 \dots N$
6. Sample a perturbed path  $\bar{y}_i \sim q(\bar{y}|\bar{x}_i)$
7. Sample  $v_y \sim g(v|\bar{y}_i)$  and construct  $\widehat{L}_{v_y}(\bar{y}_i)$
8. Set

$$\{\bar{x}_{i+1}, v_{i+1}\} = \begin{cases} \{\bar{y}_i, v_y\} & : \text{with probability } r(\bar{x}_i, \bar{y}_i) \\ \{\bar{x}_i, v_i\} & : \text{with probability } 1 - r(\bar{x}_i, \bar{y}_i) \end{cases}$$

where

$$r(\bar{x}_i, \bar{y}_i) = \min \left[ \frac{\widehat{L}_{v_y}(\bar{y}_i)q(\bar{x}_i|\bar{y}_i)}{\widehat{L}_{v_i}(\bar{x}_i)q(\bar{y}_i|\bar{x}_i)}, 1 \right]$$

9. Update Monte Carlo estimates  $\hat{I}_j$  for  $j$  where  $W_j(\bar{x}_{i+1}) > 0$ .

**3.3.2 Rendering heterogeneous media using MLT**

Many visually interesting scenes contain some form of participating media, for example smoke, clouds, biological tissue, fluids, and hair. To evaluate the contribution function in these scenes requires the computation of the transmittance between vertices along the path, defined in equation (2.16) as

$$T(x, x_t) = \exp \left( - \int_0^{\|x-x_t\|} \sigma_t(x_{t'}) dt' \right) \quad (3.29)$$

For homogenous media, where  $\sigma_t$  is constant, the transmittance between two vertices along the path simplifies to

$$T(x, x_t) = \exp(-d\sigma_t) \quad (3.30)$$

where  $d$  denotes the distance between the vertices. However, for heterogeneous media the transmittance is often analytically intractable. Instead different approximation techniques have been developed.

Perlin and Hoffert [173] proposed to solve the inner integral,  $\int_0^{\|x-x_t\|} \sigma_t(x_{t'}) dt'$ , in equation (3.29) using numerical quadrature. This approach, which has

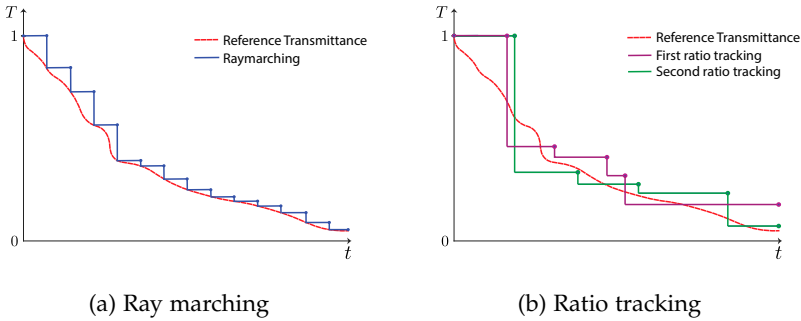


Figure 3.4: Illustrations of two techniques for evaluating the transmittance along a ray. **a)** Deterministic ray marching uses a fixed step size to traverse the ray. This method is biased and requires a large number of steps along the ray to reduce the error in the estimates. **b)** Ratio tracking is based on using a random step size and results in unbiased estimators regardless of the mean step size. This enables us to generate unbiased estimators of the true transmittance using only a few steps along the ray.

previously been used with MLT [172], is commonly referred to as *Ray marching*. While this approach allows us to estimate the transmittance in any media, it is biased. Even if the inner integral is approximated using an unbiased estimator the resulting transmittance estimator is still biased as,  $E[\exp(X)] \neq \exp(E[X])$ . In practice, the bias is often reduced by using a small step size along the ray. However, this results in expensive estimators and slow evaluation of the contribution function which leads to a significant performance drop compared to using MLT in homogenous media. Another important aspect is that when using biased approximations of the contribution function, the underlying Markov chain does not generally converge to the desired target distribution.

In paper B, we propose to render heterogenous media using the pseudo-marginal MLT algorithm, which enables us to approximate the exact contribution function by using recent unbiased transmittance estimators, know as *ratio tracking* [167]. In contrast to numerical quadrature rules, the ratio tracking algorithm generates stochastic estimates of the true transmittance using a random step size along the ray. The algorithm is based on a type of generalized rejection sampling, where distances between photon-media collisions are sampled according to a fictitious homogenous media with scattering coefficient  $\sigma_t = \mu$ . Random distances between collisions in the homogenous media can be easily sampled using a simple exponential distribution,  $\mu \exp(-d\mu)$ . The

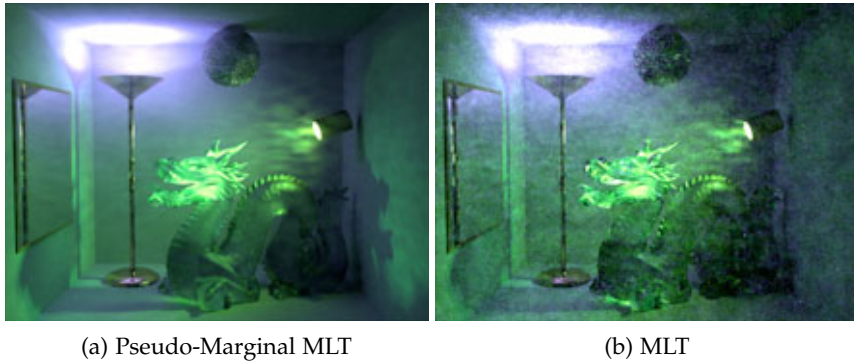


Figure 3.5: Renderings of a room with glossy and transparent objects, filled with a heterogenous smoke. This scene is very difficult to render using traditional path tracing methods. The pseudo-marginal MLT algorithm using ratio-tracking and the standard MLT algorithm using ray-marching where both run in 18 minutes on 64 vCPU:s on Amazon EC2 to render these images. The pseudo-marginal algorithm is able to simulate longer Markov chains which enables a better exploration of the path space than the standard MLT algorithm, leading to less noise.

transmittance along the ray in the heterogenous media is then evaluated by computing the product of the ratios  $r(x) = \frac{\sigma_t(x)}{\mu}$  at each sampled position,  $x$ , along the ray. More details on ratio tracking can be found in [167].

The ray-marching and the ratio tracking algorithm are illustrated in figure 3.4. A key advantage of the ratio tracking algorithm compared to traditional ray marching is that it enables us to obtain computationally cheap, unbiased estimates of the transmittance. Using the pseudo-marginal MLT algorithm with ratio tracking to evaluate the contribution function enables us to generate longer Markov chains, and leads to more efficient estimators. This leads to renderings with less noise compared to using biased ray-marching techniques. Figure 3.6 shows an equal render time comparison between the proposed method using ratio-tracking and a standard MLT algorithm using ray-marching to evaluate the transmittance between vertices along the ray.

We can also use the pseudo-marginal construction in other MLT variants, such as ERPT. Figure 3.6 shows two equal time renderings using pseudo-marginal ERPT with ratio tracking and ERPT with ray-marching. It is clear that the proposed approach leads to significant improvement in these types of scenes.

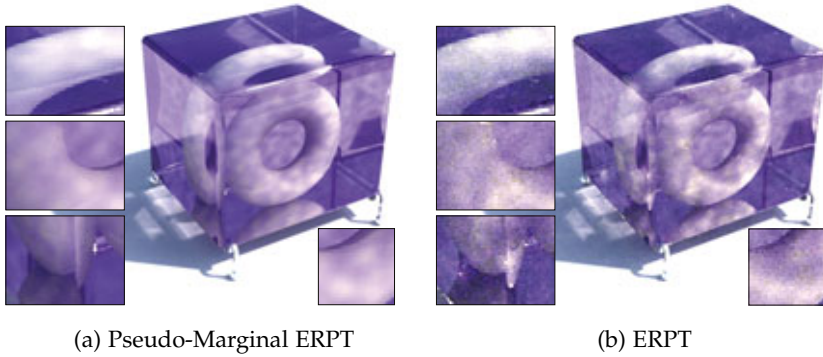


Figure 3.6: Equal time comparison between pseudo-marginal ERPT using ratio-tracking and ERPT using ray-marching to evaluate the transmittance. The pseudo-marginal algorithm enables faster exploration of the path space which leads to more efficient renderings.

### 3.4 Summary and future work

In this chapter we presented an overview of Monte Carlo rendering algorithms with an emphasis on methods based on MCMC sampling. We also showed how the pseudo-marginal MLT method presented in paper B increases the flexibility and efficiency of previous MLT algorithms. Here we discussed how the transmission term can be evaluated using unbiased estimators, such as ratio tracking. However, the algorithm is general and allows for any unbiased estimator to be used in place of the exact contribution function. An interesting venue for future work is to investigate other unbiased estimators that might be more efficient than using their exact counterparts. For example, approximating expensive production shaders with many texture lookups and procedural noise calls with unbiased estimates.

Another interesting venue for future work is to investigate how the developed method can be used in other applications dealing with Monte Carlo simulation of participating media, such as reactor design and radiation dosimetry. In these applications one has to include models of other types of scattering behavior that occur at higher energies than that of visible light. To the best of the authors knowledge, MCMC methods for simulating photon or neutron transport in these applications have not been investigated previously.



## Image Based Lighting

Lighting is one of the key components in the process of creating visually interesting renderings. This has led to development of a wide range of methods for building complex lighting environments in formats suitable for efficient image synthesis. Image based lighting (IBL) is now a commonly used technique for capturing and using the lighting conditions in real scenes as source of illumination in photorealistic rendering applications. In this chapter, we provide an overview of current state-of-the-art IBL techniques, and discuss what challenges remain to be solved. In this context, we also give an overview of the contributions presented in papers C and D included in this thesis.

First, in section 4.1, the foundations of light transport in mixed reality scenes are discussed. In section 4.2, we then present a brief overview of traditional techniques for capturing and representing the lighting conditions in real scenes. This overview is complemented by Paper A, which presents a comprehensive survey and classification of previous methods for photorealistic rendering of mixed reality scenes. The contributions presented in paper C for robust rendering of virtual objects lit with temporally varying illumination captured in a real scene is presented in section 4.3. Finally, in section 4.4, we discuss the methods presented in paper D for representing and rendering with captured spatially varying real world illumination.

## 4.1 Light transport in mixed reality scenes

When modeling the light transport in scenes where virtual objects are mixed with captured real environments, it is constructive to reformulate the rendering equation (2.8) by expanding it into several parts, modeling the reflectance from different types of incident illumination separately. The outgoing radiance at a virtual surface point,  $x$ , can then be expressed as,

$$\begin{aligned}
 L(x \rightarrow \omega_o) &= L_e(x \rightarrow \omega_o) \\
 &+ \int_{\Omega} L_i^r(x \leftarrow \omega_i) \rho(x, \omega_o, \omega_i) (\mathbf{n}_x \cdot \omega_i) d\omega_i \\
 &+ \int_{\Omega} L_i^v(x \leftarrow \omega_i) \rho(x, \omega_o, \omega_i) (\mathbf{n}_x \cdot \omega_i) d\omega_i \\
 &+ \int_{\Omega} L_i^{v,r}(x \leftarrow \omega_i) \rho(x, \omega_o, \omega_i) (\mathbf{n}_x \cdot \omega_i) d\omega_i
 \end{aligned}$$

The first term describes the emitted radiance from the virtual object. The second term describes the reflection of the incident illumination originating from the real scene which has not been reflected at surfaces of the virtual objects,  $L_i^r(x \leftarrow \omega_i)$ . The third term describes the response to incident illumination that has been emitted from the real or virtual scene and reflected one or more times on surfaces in the virtual scene before reaching  $x$ ,  $L_i^v(x \leftarrow \omega_i)$ . Finally, the fourth term represents the reflection from incident light that has interacted with both real and virtual objects in the scene,  $L_i^{v,r}(x \leftarrow \omega_i)$ .

To synthesise the appearance of the virtual objects, ideally all terms should be accounted for. However, the last term is often approximated, as a comprehensive model of the real scene would be necessary to simulate the complete interaction between the virtual and real objects. Instead, the simulation is often simplified by dividing the real scenes into two parts, a *distant* part which is not affected by the virtual objects, and a *local* part, in which the reflectance properties and geometry are modeled. The distant scene is, in contrast to the local scene, only represented by its emitted radiance or, equivalently, by the incident illumination on the virtual objects from the real scene. Techniques for capturing and estimating the incident illumination on the virtual objects from the real scene,  $L_i^r(x, \omega_i)$ , are discussed in section 4.2.

### 4.1.1 Differential rendering

A common IBL application is to use the captured illumination to create photo-realistic renderings of virtual objects and composite the renderings onto a background photograph or video sequences from the captured scene. To compute the final, composited image it is necessary to simulate both how the



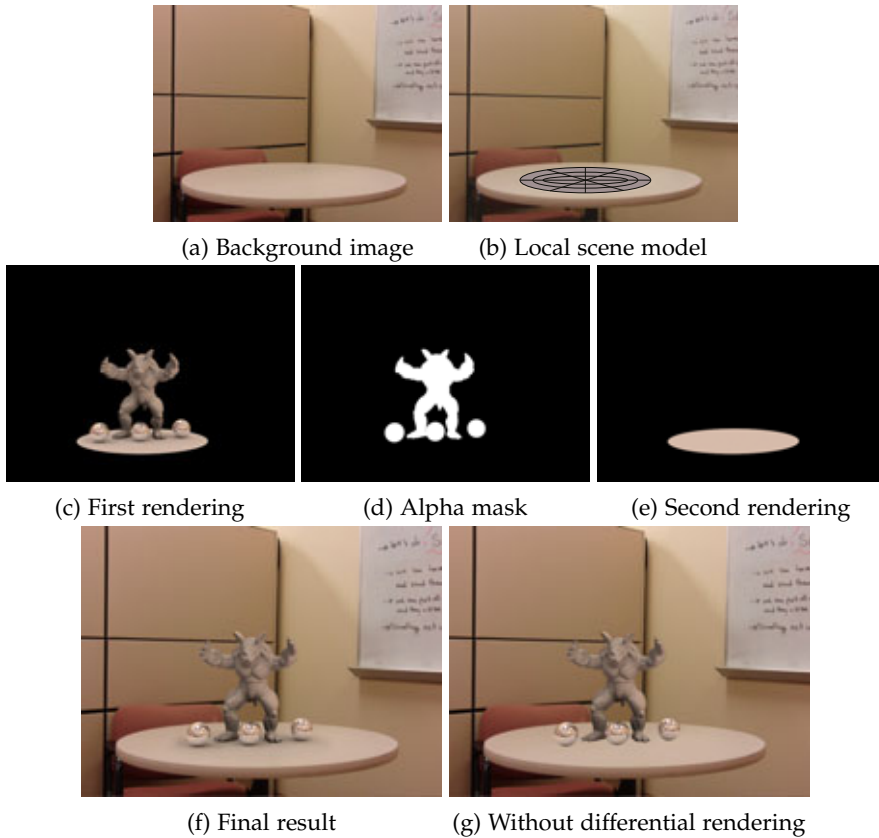


Figure 4.1: An example illustrating the steps of differential rendering [46], see the text for more details. The result generated with differential rendering (f) looks more realistic than the result obtained when directly compositing the virtual objects into the real photograph (g).

captured lighting affects the virtual objects and how the virtual objects affect the real scene, through effects such as cast shadows and color bleeding. In IBL, this is commonly accomplished using a technique known as *differential rendering* [46, 64], in which the captured scene is divided into a local scene and a distant scene as described in the previous section. The local scene describes a model of the geometry and reflectance properties of the real scene, and is used to simulate the light transport from the virtual objects onto the real scene. The distant scene is usually represented by the captured scene information only, and is only used as illumination during rendering of the virtual objects.

Differential rendering is outlined in figure 4.1. The first step is to render an image that includes the virtual objects and the local real scene model,  $I_{rpv}$ , with an associated *alpha mask*,  $\alpha$ . The alpha mask describes which pixels in the rendered image that correspond to the virtual objects by setting them to 1, and which pixels correspond to real objects (including the local scene model) to 0. Then, a second image,  $I_r$ , that only contains the real scene model is rendered. The final composited image,  $C$ , containing both the real scene and the virtual objects can then be computed as:

$$C = \alpha * I_{rpv} + (1 - \alpha)(B + I_{rpv} - I_r) \quad (4.1)$$

where  $B$  denotes the background photograph of the real scene. Pixels corresponding to virtual objects are taken from the first rendering and pixels corresponding to the real scene is computed by using the background image and adding the computed difference in appearance of the local scene model due to the virtual objects. Figure 4.1 shows an example of the steps of differential rendering.

## 4.2 Traditional IBL

While the incident illumination from the distant real scene can be modelled by a few manually placed light sources, this process is time-consuming and error prone. Traditional IBL methods instead represent the illumination using a panoramic HDR image, or *environment map*, that captures the angular distribution of incident radiance around a point in the scene [26, 74, 156], see figure 4.2 for an illustration. During rendering, the incident illumination can then directly be looked up in the environment map, where each pixel represents the radiance incident onto the virtual objects integrated over a small solid angle, corresponding to the pixel resolution in the HDR panorama. Previous methods for generating these environment maps can roughly be classified into two categories. The first category of methods aims at capturing a physically accurate model of the lighting conditions in the scene. These methods rely on measurements in the real scene, such as capturing omnidirectional images. The second category of methods instead try to avoid physical measurements in the scene by estimating the illumination directly from the backdrop image or video [112, 127], enabling for example object insertion in legacy footage [105, 106, 227]. Although these methods can provide increased speed and flexibility, they generally sacrifice physical accuracy and offer considerably less robust solutions making them less suitable for photorealistic rendering applications. A comprehensive survey of estimation based methods can be found in Paper A. In this chapter we focus the discussion on the measurement based methods.

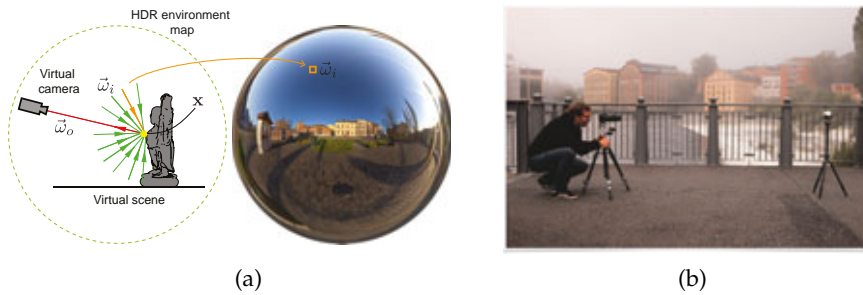


Figure 4.2: Traditional IBL represents the illumination from the distant real scene using an *environment map* (a), a spherical image describing the angular distribution of incident radiance around a point in the scene. A common method to capture environment maps is to photograph a mirror sphere using HDR imaging techniques to represent the full range of the incident radiance using calibrated linear-response measurements (b).

For physically accurate simulation of the light transport in the scene it is necessary to capture the full dynamic range of the lighting in the scene, from direct light sources to regions in shadow, using calibrated linear-response measurements. In the seminal work of Debevec [46], the lighting is captured as a  $360^\circ$  panoramic HDR image, a *light probe*, at a single point in the scene. Figure 4.3 shows a comparison of a virtual objects rendered using the scene illumination represented using 4.3a) a single extracted directional light representing the sun, 4.3b) a 8-bit low dynamic range environment map and 4.3c) a HDR light probe. The directional light fails to capture the detailed variations of the incident radiance, see e.g. the reflections on the specular spheres. Using only a low dynamic range environment map leads to fewer shadows and detailed reflections, as it does not capture the true intensity of the sun and the environment. Using IBL with an HDR light probe on the other hand, produces more realistic and detailed lighting effects as the full dynamic range of the incident illumination from the real scene is represented.

Illuminating virtual objects with HDR light probes is a key component in most current visual effects production workflows [28], and IBL is supported in the vast majority of commercial rendering packages. A more detailed description of the practical aspects of capturing and rendering with light probes can be found in [179] and [20]. Alternative approaches to capturing light probes in the scene have also been developed. For example, using other, more robust, panoramic imaging systems [51, 217] and specially designed light probes that can capture the radiance of strong light sources in the scene using standard low dynamic range imaging [35, 49]. More details on these methods is given in section 4.1 in

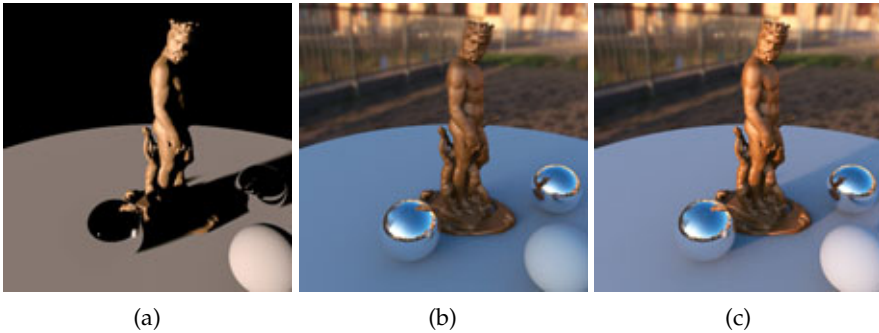


Figure 4.3: Synthetic objects rendered with incident illumination from a real scene represented using **a)** a single directional light source representing the sun **b)** a low dynamic range environment map **c)** a HDR light probe. The scene model was acquired by the Visual Computing Laboratory, ISTI-CNR, and the renderings are courtesy of Francesco Banterle.

paper A.

## 4.3 Video based lighting

Standard IBL techniques capture real world lighting conditions using a single panoramic HDR image at a fixed instant in time. In many applications it is desirable to allow the lighting conditions to change over time. For example, when the lighting conditions in the real scene change over time, or when rendering virtual objects that move through the real scene. In this section, we discuss methods for capturing the lighting as a panoramic HDR video sequence, henceforth referred to as a *video light probe*.

### 4.3.1 Previous approaches

Capturing video light probes is a challenging task that requires methods other than those developed for traditional light probe capture. For example, standard HDR imaging techniques used to capture light probes often rely on capturing several separate images with different exposure settings over time. Without further processing, this process introduces visual artifacts in light probes captured in dynamic illumination, such as misalignments and excessive motion blur. It is also necessary to achieve a temporally consistent illumination without perceptually disturbing flickering and other temporal artifacts. While some IBL systems have been developed for capturing video light probes, they provide very limited spatial resolution or limit the dynamic range that can be captured [86,

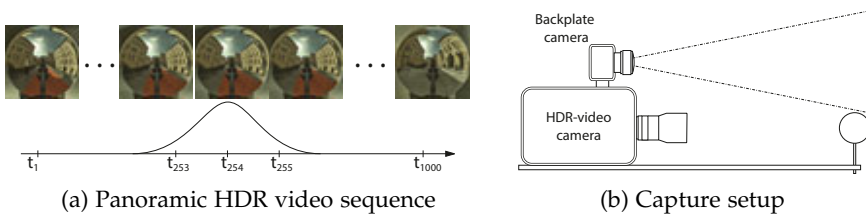


Figure 4.4: a) An example panoramic HDR-Video sequence. Each rendered frame use an individual lighting measurement. To avoid flickering, it is useful to apply temporal filtering over a few frames in the HDR light probe sequence. b) For each frame, our setup synchronously captures both a panoramic lighting measurement and a high resolution backplate.

218]. A more detailed overview of previous methods is given in section 4.2 in Paper A.

### 4.3.2 Contributions

In paper C, we present a simple and robust system for photorealistic rendering of synthetic object placed into real video footage. The illumination capture is based on using HDR video light probes to measure the temporal variations in the illumination. The capture setup, shown in fig 4.4b, consists of two synchronized cameras. One high-resolution HDR video camera capturing the spherical lighting distribution by imaging a mirror sphere, and one separate camera for capturing high resolution backplate footage. To keep the mirror ball and the supporting mount out of view of the backplate camera, the HDR video rig is placed slightly below or over the primary camera. While this can introduce some inconsistencies in the captured scene illumination when used to drive the rendering of virtual objects inside the field of view of the backplate camera, in many scenes this effect is negligible. The HDR video camera we use is based on simultaneous capture of several exposures of the incident light with synchronized exposure times and is discussed further in section 5. As the mirror ball is mounted on a supporting structure, some wobble and high frequency vibrations are often introduced in the HDR video footage of the mirror ball. However, as the mirror ball is symmetric, a simple tracking of the camera lens centre, visible in the captured frames of the mirror ball, can be used to stabilize the light probe footage, effectively centering the mirror ball in the field of view of the HDR video camera. Practical details of this procedure can be found in [28].

For each frame in the rendered video, we use a single frame of the captured video light probe. This can lead to high frequency changes in the time domain,

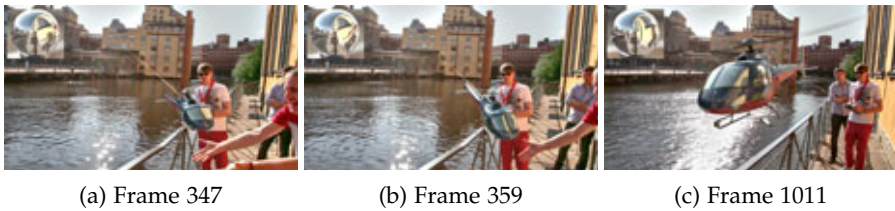


Figure 4.5: Rendered video frames of a virtual RC helicopter illuminated by captured temporally varying illumination. The captured light probe used for each frame is shown in the upper left corner.

which in some cases manifests themselves as flickering artifacts. To counteract such artifacts we smooth out rapid changes in the temporal domain by filtering the captured light probe frames, illustrated in Figure 4.4a. A gaussian or triangular filter over 3-5 frames usually provides a reasonable smoothing effect, however more advanced filters can also be applied.

In figure 4.5, four separate frames from a rendered sequences of a virtual helicopter moving through temporally varying illumination conditions is shown (renderings produced in collaboration with Christian Bloch at Eden FX).

The proposed process provides a rapid workflow for photorealistic rendering of virtual objects in real video footage. As the background video footage is synchronized with the video light probe by default, there is no need to track the camera motion or match illumination and reflections. One only needs to animate the synthetic object and render it with the provided video light probe, a simple task in most commercial modeling and rendering packages. Efficient Monte Carlo rendering of such light probe sequences is possible using Sequential Monte Carlo algorithms, a topic we explored in [118]. Another approach for photorealistic real-time rendering applications is to use precomputed radiance transfer techniques, which we explored in [79].

## 4.4 Rendering with spatially varying illumination

Capturing the incident illumination at a single point in the real scene, as with traditional IBL, provides accurate representations when the synthetic objects are small in comparison to the distance to the real scene. However, in many scenes there are considerable spatial variations in the incident illumination over the extent of the synthetic objects, e.g. cast shadows or light shafts. Failing to capture these variations in the illumination, which often are important parts of a visually interesting lighting design, generates less plausible renderings. Figure 4.6 shows a comparison between a) a traditional IBL rendering using



(a) Traditional IBL

(b) Spatially Varying IBL

Figure 4.6: Traditional IBL using a single light probe fails to capture the important spatial variations in the scene illumination. Using techniques developed for capturing, representing and rendering with spatially varying illumination, presented in paper D, the resulting renderings display more visually interesting complexities and subtleties in shadows, reflections, and shading.

a single light probe to capture the incident illumination and **b**) a rendering generated using methods for capturing and representing spatially varying illumination presented in paper D .

#### 4.4.1 Previous approaches

To accurately represent the spatially varying incident radiance in the scene, it is necessary to capture the full 5D spatial and angular dimensionality of the illumination. To this end, two main approaches can be used. The first approach is based on representing the illumination by a large number of spatially distributed angular radiance samples, e.g. a dense grid of light probes [92, 146, 159, 206, 208]. The incident light at a point on a synthetic object can then be computed by interpolating the angular radiance distribution of spatially adjacent light probes [138]. However, for photorealistic results these techniques require a large number of light probes to be captured in the scene, leading to tedious measurement procedures and large memory requirements. Such direct representations are also difficult to edit according to artistic needs. An alternative approach is to reconstruct a rough model of the scene geometry and representing the emitted illumination from the surfaces of the model. In the simplest case, the position of area light sources in the scene can be reconstructed using two light probe images using stereo matching techniques [42, 186]. Recent work has also proposed methods for extracting area light sources from a single light probe guided by user input to recover depth information [21]. Time-of-flight laser scanning and inverse



rendering techniques have also been used to recover diffuse reflectances of surfaces in the scene [28, 47]. However, these methods have previously been constrained to small scenes in controlled environments or required extensive manual processing to produce high quality results. Furthermore, they are often limited to Lambertian scenes and/or recover very rough models of the scene geometry and thus fail to capture detailed spatial variations in the incident illumination. A detailed overview of previous techniques for capturing and representing spatially varying illumination is given in section 4.3 in paper A.

## 4.4.2 Contributions

In paper D, we propose a framework for efficient capture and representation of spatially varying real world illumination that can be edited and re-used. The framework is based around HDR video capture, which enables rapid capture of the incident illumination in the scene. Directly storing the HDR video sequences requires tens to hundreds of gigabytes and lead to inefficient editing and rendering. To meet these challenges, we propose to reconstruct a model of the real scene that represents the radiometric properties of the incident illumination. The model requires substantially less storage, is straightforward to edit, and enables efficient photorealistic rendering of synthetic objects placed into the scene.

### Reconstructing the scene model

In a first step omnidirectional HDR video is used to capture the incident radiance in the part of the scene where synthetic objects will be placed. A set of high resolution images and/or video sequences are also captured. Based on the captured images a rough outline of the scene geometry can be reconstructed using standard structure from motion (SfM) techniques [67, 160, 189]. However, this geometry is often not complete and can suffer from noisy results in regions with specular materials or complicated geometry. To complement the SfM reconstruction a semi-automatic approach is used where the user models the scene using shape from focus cues. To provide the user with an intuitive interface for modeling based on shape from focus, a set of 3D, voxel-based representations, *focal volumes*, of the captured lighting environment is created. To compute the voxel values of the focal volumes, the average radiance captured by each light probe in the direction of the voxel is computed, as illustrated in Figure 4.7a. Conceptually, a focal volume is a synthetic aperture [212] projection using a 3D aperture of the size and shape of the capture region. A surface placed inside the volume can be thought of as a focal surface. Points on a focal surface that coincide with a surface in the real scene will be in focus, but will rapidly move out of focus with the distance to the real surface, see



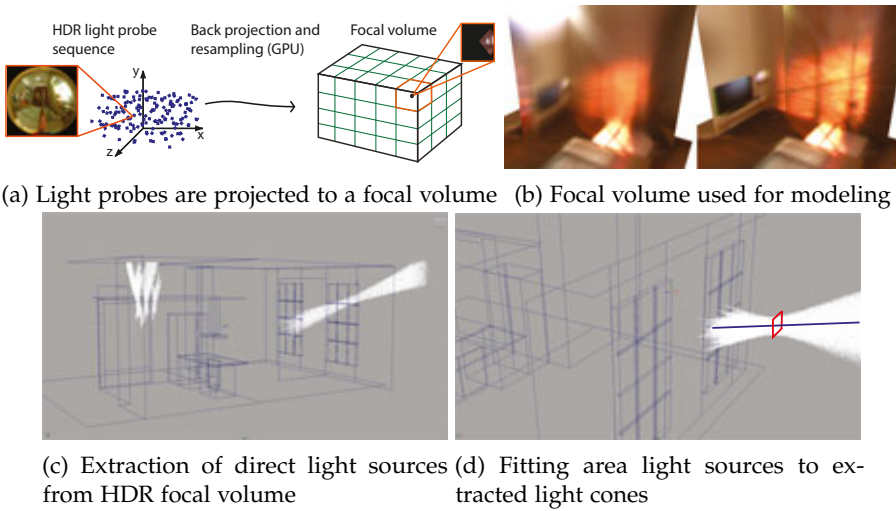
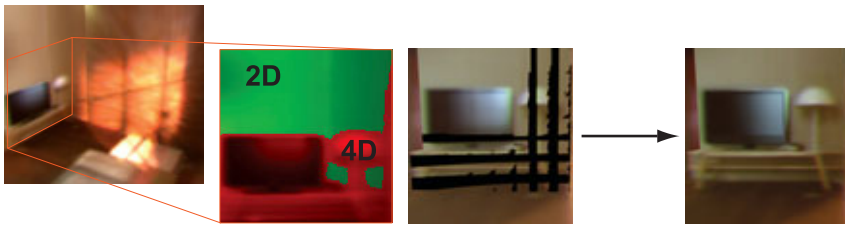


Figure 4.7: **a)** The captured radiance in the light probes are back projected from each capture position to a 3D focal volume, contributing to various statistics computed at each voxel they intersect. **b)** Focus volumes can then be used for rapid user assisted modeling of the approximate scene geometry by e.g. placing planar surfaces, represented by slices through the volume, such that they appear to be in focus. **c), d)** Using the focal volume, light sources can easily be extracted and fitted to represent the emitted radiance.

Figure 4.7b. Secondly, due to the HDR nature of the data, light sources will be represented by distinct high intensity voxels. By thresholding the focal volume, these light sources can easily be extracted and robustly fitted by using e.g. principal component analysis, see figure 4.7c and 4.7d.

Provided with a rough, piecewise planar, model of dominant geometry in the scene the next step is to project the captured radiance data from the omnidirectional HDR video onto the reconstructed surfaces. To represent the excitant radiance from surfaces we use 4D view dependent textures, also known as surface light fields [226]. Using surface light fields allows us to accurately represent the incident radiance on the virtual objects in the scene even though we only have a rough approximate model of the geometry, and in general no knowledge of the reflectance properties of the surfaces in the scene. For diffuse surfaces with accurate geometry, such as for example walls and other larger structures, the 4D representation can be collapsed into a standard 2D texture as the excitant radiance does not depend on the viewing angle.

A problem with directly using the mean radiance value of the omnidirectional



(a) Representation of emitted radiance      (b) Reconstruction of missing values

Figure 4.8: **a)**The emitted radiance from reconstructed surfaces in the scene are represented using 4D light fields describing both the spatial and angular dimension, however for accurately reconstructed diffuse surfaces, the emitted illumination can be represented using 2D textures directly. **b)**To reconstruct missing values in the 4D structures we use interpolation techniques based on e.g. fitting a analytical BRDF model to the data.

light probes, is that small miss registration errors manifest as blurred textures. To avoid this, we use only one light probe to determine the value of a pixel. To obtain high-resolution texture maps without visible seams we assign each pixel to a light probe by minimizing a cost function defined by a Markov Random Field. The cost function is designed so that light probes with a good resolution in the direction of the pixel are preferred and uses smoothness constraints to avoid visible seams in the resulting textures. The approach is similar to previous methods for high quality texture projection [68, 131], but extended to handle omnidirectional source images with considerable distortion in angles close to the backward facing direction.

The data stored in the 4D-light fields on surfaces in the scene model are often not complete after the projection step, for example a point on the surface may not have been observed from all possible viewing angles. This is due to the limited number of panoramic HDR images that is practical to capture in the scene, even using HDR video light probes. However, the data can be completed by interpolation and extrapolation of the data that is measured. For surfaces that coincide with real geometry in the scene, it is possible to fit standard reflectance distribution functions to interpolate over viewing angles. This can be done by fitting a parametric BRDF model to the measured excitant radiance and the incident radiance reaching the point from the extracted light sources in the scene. The fitted BRDF parameters can then be used to interpolate or extrapolate the excitant radiance that was not captured in the direct measurements, see figure 4.8 for an example.

## Rendering

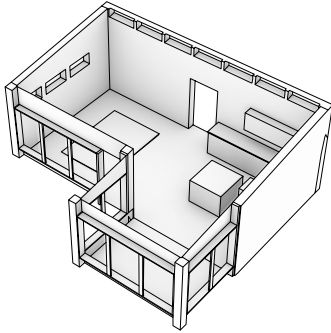
During rendering, we employ a standard differential rendering procedure as detailed in section 4.1.1. To represent the incident radiance from the distant real scene, we use the reconstructed scene model. By illuminating the virtual objects with area lights, corresponding to both the extracted direct light sources and the excitant radiance from the surfaces in the scene, accurate spatial variations in the scene illumination can be represented.

An example application of the developed methods is shown in Figure 4.9, demonstrating virtual furnishing of a living room. Including setup time, and time for capture of high resolution HDR video sequences for structure from motion computations and backdrops, the total time spent on-site was about 2.5 hours. In comparison, a real refurbishing of the house and ordinary photo-shoots could easily take a day or longer, without the option of adjusting the results at a later stage. The final proxy geometry shown as an ambient occlusion rendering in Figure 4.9a was recovered using structure from motion to get the point cloud data and then modified and modeled using the focal volume modeling approach. Figure 4.9b shows the virtual furniture placed in the scene, and Figures 4.9c and 4.9d shows examples of final renderings.

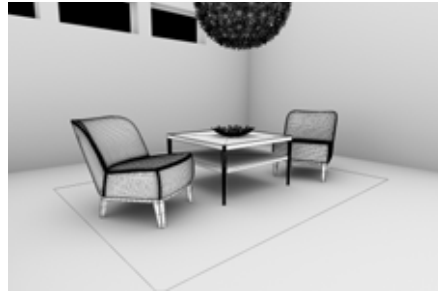
## 4.5 Summary and future work

In this chapter we discussed how traditional IBL techniques can be extended to capture dynamic and spatially varying lighting conditions. These approaches not only enables photorealistic renderings of virtual objects in video footage but can also add important visual effects in scenes with dappled lighting and cast shadows for example. We showed that state-of-the-art HDR video cameras enables rapid and accurate measurement of such illumination environments. We also presented two practical systems for rendering with captured HDR video data, using temporally and spatially varying illumination respectively.

An interesting venue for future work is to develop systems that can handle spatially varying illumination conditions that also change over time. Such systems could for example be based on using stereo HDR video light probes.



(a) Reconstructed geometry



(b) Virtual furniture models



(c) Final rendering



(d) Final rendering

Figure 4.9: Example application of the methods proposed in paper D, demonstrating virtual furnishing of a living room.

## HDR Imaging

Current commercial digital camera sensors restrict the information that can be captured in a single image. In particular, the sensor noise and limited well capacity of modern image sensors such as complementary metal-oxide-semiconductors (CMOS) and charge-coupled devices (CCD) limits the *dynamic range* that can be captured in a single image. A sensor's dynamic range describes the ratio between the largest measurable signal, before the sensor is saturated, to the smallest distinguishable signal level. In many applications it is desirable to capture images with a higher dynamic range than that of modern image sensors, as it improves exposure and post-production flexibility and expands artistic possibilities in dramatic and unevenly-lit scenes. For example, when using IBL in outdoor scenes, with areas in shadow and parts of the scene directly lit by the sun, a dynamic range in the order of 10,000,000 : 1 or more can be required for accurate capture of the illumination conditions in the scene. This is far more than the limited dynamic range of modern digital camera sensor, which is roughly limited to 12-14 bits per pixel, capable of representing a dynamic range of about 10,000 : 1. In contrast to traditional imaging, HDR images captures a much wider dynamic range and allows for radiometrically accurate capture of both direct light sources and shadowed regions of the scene.

The traditional approach for capturing of HDR images is to use *exposure bracketing* where a sequence of low dynamic range (LDR) images are captured with varying exposure times and then merged to an HDR image. However, for dynamic scenes or excessive camera shake this approach can suffer from ghosting and motion blur artifacts, especially in video applications. An alternative to these methods is approaches that use a single, global, exposure time. These

approaches are, in contrast to exposure bracketing, based on modifying the original camera design, using multiple-sensors, spatially varying neutral density (ND) filters or/and varying the ISO/gain over the sensor. Reconstructing HDR images from such systems is a challenging task. Multiple, possibly saturated, measurements of the scene illumination need to be fused to a single estimate. Furthermore, traditional digital camera processing steps such as debayering (color filter interpolation), realignment and denoising often need to be performed. In section 5.1 we give an overview of current methodologies used for HDR image and video capture.

The main contributions of paper E, F, and G are then discussed in section 5.2. In paper E we present a novel, statistically motivated, HDR reconstruction framework that unifies all the processing operations into a single filtering operation. We also present extensions to the original formulation allowing for adaptive filtering supports, presented in papers F and G. We show several examples where the novel reconstruction algorithms improve the quality and efficiency compared to previous methods for HDR video reconstruction using a multi-sensor system and a standard DSLR camera using *dual-ISO* capture. Finally, in section 5.3, we provide a summary of the material described in the chapter and discuss future work.

## 5.1 State-of-the-art HDR capture

In this section we give an overview of current methods for capturing HDR images and video. Before we discuss specific HDR capture techniques, we first introduce some background on traditional digital camera pipelines and their limitations.

### 5.1.1 Digital cameras and raw sensor data

Modern off-the-shelf digital cameras use CCD or CMOS sensors that convert incident photons into digital output values. These sensors are silicon based integrated circuits that consist of a dense matrix of photo-diodes. When the sensor is exposed to incident light, photons interact with the silicon atoms and, due to the photoelectric effect [61], generate electrons that are stored in a potential well in the photo-diod. The accumulated photo-induced electrons are then converted to an output voltage, amplified, and quantized to digital values. CCD sensors generally convert the charge to voltage one pixel at a time, or a few pixels at a time, using a few output nodes located outside of the photosensitive part of the sensor. CMOS sensors on the other hand allows additional electronics to be placed at each photo-diod which in turn enables the charge-to-voltage conversion to be performed per pixel, however the readout

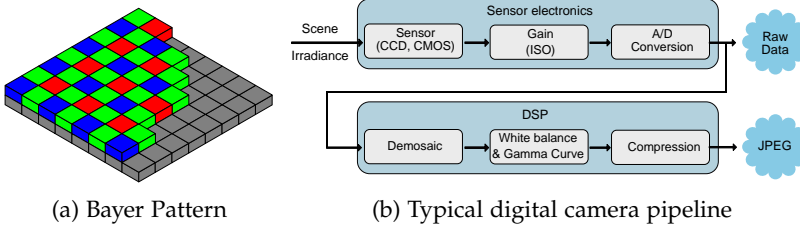


Figure 5.1: **a)** The standard method for capturing color images is to place a color filter array over the sensor. Illustration **a)** by Colin M.L. Burnett, copyright CC BY-SA 3.0 **b)** A typical model of image capture using a digital camera.

is often limited to one column of pixels at a time [168]. Both CCD and CMOS sensors has a roughly linear response to the incident light intensity up to a saturation threshold, where the potential well is full, and no further electrons can be stored.

Due to the limited bandwidth and storage capacity the output from most traditional off-the-shelf digital cameras are limited to 8-bit JPEG compressed images. These images are in general mapped through non-linear tone curves that distort the linear sensor response. However, modern DSLR cameras often offer a RAW capture mode, which allows direct access to the linear sensor response. If the raw camera data is not accessible directly, it is possible to use techniques for estimating the inverse camera response curve, mapping pixel values in the JPEG image to actual sensor responses [48, 75].

The photo-diodes on standard imaging sensors can not by themselves distinguish between different wavelengths of the incident light. Instead, to capture color images the most common way is to introduce a color filter array (CFA) over the image sensor. The CFA is a mosaic of different color filters, often red, green and blue, that only transmit a part of the visible spectrum onto each pixel, as shown in figure 5.1a. Full resolution color images can then be recovered by interpolation, often referred to as *demosaicing* [76]. The most common CFA is the Bayer pattern [23], consisting of 50% green, 25% red and 25% blue filters, motivated by the fact that the human visual system is more sensitive to the green part of the visible spectrum.

An overview of the typical operations performed in a digital camera is shown in figure 5.1b. First the analog signals recorded by the photo-diodes are amplified by a *gain* factor proportional to the ISO setting of the sensor. After amplification, the analog signal is quantized into digital values using A/D conversion electronics. After the A/D conversion many modern camera sensors allow the data to be accessed as RAW images. For JPEG images additional processing

steps are carried out inside the camera electronics, such as gamma mapping and image compression. More details on imaging sensors and the traditional camera pipeline can be found in many excellent books, see for example the book by Lukac [140].

The measured pixel values are affected by several uncertainties inherent in different parts of the acquisition process. Due to the discrete nature of light, the number of photons arriving at a photo-diode in a given time follows a Poisson distribution. The observed noise in the digital image sensor due to this uncertainty is often referred to as *photon shot noise*. Several other noise sources also affect the measured digital pixel values [5, 178]. These include noise due to thermal effects inherent to the readout circuitry, charge-to-voltage conversion, and the A/D conversion. The effect of quantization can also be modeled as stochastic noise [225]. The combined effect of all these noise sources are often referred to as the *readout noise* of the sensor.

### 5.1.2 Dynamic range

The dynamic range of an image sensor is fundamentally limited by the bit-depth of the output, often limited to 12-14 bits for RAW image readout. For example, a standard image sensor, with a linear response, using 12 bits to represent the resulting digital values is inherently limited to a dynamic range of  $1 : 2^{12} = 4096$ . In the HDR literature, the dynamic range is often described in units of *f-stops*, corresponding to the logarithm in base 2 of the ratio between the largest to the smallest representable value. For example, a 12-bit output corresponds to  $\log_2(2^{12}) = 12$  f-stops. While this definition of dynamic range serves as a general description of the fundamental limit of imaging sensors, in reality, all imaging sensors are affected by noise that can make the lower representable values extremely noisy. The *Signal-to-noise* ratio (SNR) is an alternative, more precise, definition of dynamic range that instead describes the ratio of the largest representable value (before saturation) to the pixel noise.

An often disputed topic is the dynamic range that can be perceived by the human eye. From a purely physical standpoint, the scattering of light inside the eye limits the maximum contrast that can be perceived by the eye to about  $1 : 3000$  [144]. However, as the eye rapidly changes gaze and adapts locally, it is commonly believed that we can perceive scenes with a dynamic range of  $1 : 40000$  or more [144], far more than the capabilities of a traditional camera sensor. In fact, with adaptations over time, relevant for video applications, the human visual system is believed to be capable of perceiving a dynamic range of about  $1 : 10^{14}$ , from a moonless night to bright daylight [63]. However, note that human perception is a complicated topic and the actual perceived dynamic range and color perception can in fact vary between scenes.



### 5.1.3 Exposure bracketing

The most widespread approach for capturing HDR images using off-the-shelf digital cameras involves taking multiple images, each captured using a different exposure time. Each image in the sequence is exposed to capture a limited part of the full dynamic range of the illumination. To combine the images into a single HDR image, the images are first normalized to a common radiometric space by accounting for the exposure time and the possibly non-linear camera response curve. The final result is then computed by a weighted average over the pixel values in the different exposures [48, 143, 157]. Modern HDR fusion is based on adapting the weights according to the statistical properties of the camera noise [7, 73] to minimize the resulting variance of the HDR pixel estimates. The minimum number of exposures required to accurately capture a scene given an acceptable noise level has also been studied [69, 84]. These techniques have also been used for HDR video capture using separate exposures for each frame [103] or rows [205].

Without explicit registration of the images this approach often leads to ghosting and motion blur artifacts when directly fusing the different exposures. In particular, for dynamic scenes, non-rigid registration of the individual exposures is necessary. The development of general and robust registration algorithms suitable for HDR imaging is still an open problem that has attracted much work during the last decade, a good survey of current methods is given in [191]. However, current approaches are still limited, in particular for complex scenes with many dynamic objects.

### 5.1.4 Single Shot Techniques

In this section we discuss HDR capturing systems that do not vary the exposure time, but instead use optical setups, neutral density (ND) filters or ISO variations, so that the different sensors or pixels capture different exposures. A major advantage of these systems is that using a common exposure time for all sensors enables robust capture of dynamic scenes. These, *single-shot*, HDR capture designs can be divided into two main categories: The first uses optical elements (e.g. beamsplitters) to project the incident optical image onto multiple image sensors. To achieve varying exposures on the different sensors, optical filters can be inserted in front of the individual sensors or different ISO/gain settings can be used for each sensor. These optical designs also make it possible to vary the percentage of incident light that is projected onto each sensor. The second category is based on using a single sensor where the response to incident light varies over the sensor. The spatially varying pixel response is often achieved by placing an optical filter array in front of the sensor, but there are also approaches in which e.g. the ISO/gain response is varied over the sensor.

This design avoids the need for more than one sensor, and allows for robust HDR video capture in dynamic scenes. This is similar to color imaging via a color filter array (commonly a bayer pattern).

Traditionally, a major disadvantage of these single-shot techniques compared to exposure bracketing is that they have relied on optical filters to achieve different exposures. This prevents a large part of the incident light to reach the sensor. However, recently several systems using both multiple sensors and spatial multiplexing have been proposed that utilizes close to all the incident light.

Another approach for single shot HDR capture is to use specialized image sensors, e.g. designed to have a *logarithmic response* to incident light. These sensors can capture a higher dynamic range compared to standard CMOS and CCD sensors [197]. However, in most applications they are still not accurate enough due to problems with low image resolution, excessive image noise in darker regions, and that the entire dynamic range is usually quantized to a 10-12 bit output.

### Multi-sensor systems

By using optical elements such as beamsplitters it is possible to project the optical image incident to the camera system onto several sensors. Different exposures are achieved by placing neutral density filters in front of the sensors, using different ISO/gain settings, or splitting the light unevenly onto the sensors. To eliminate registration and motion blur problems the exposure time of the sensors is often synchronized.

A simple setup is to use a common beamsplitter that projects the incident light onto several cameras with separate lenses, each receiving the same optical image. Froehlich et al. [65] used such an setup with two commercial Arri Alexa cameras to capture a dynamic range of up to 18 f-stop. This procedure can also be applied recursively to construct so called *optical splitting trees* where the light is projected onto multiple cameras. Mcguire and Hughes [148] presented a framework for optimal selection of components for such optical splitting trees given specific target goals, such as dynamic range, spectral sensitivity and cost budgets. However, these systems are limited as the separate lenses must be perfectly matched, and zoom and focus settings can be difficult to maintain between them. In addition, placing the beamsplitter in front of the camera lens often puts limitations on the field of view. This type of setups tend to be quite bulky, often prohibiting the design of a single handheld unit.

Alternatively the beamsplitters can be placed behind the camera lens, inside the camera housing. Aggarwal and Ahuja [4] presented one of the earliest multi-sensor systems for HDR video capture. In Figure 5.2c a modern multi-sensor

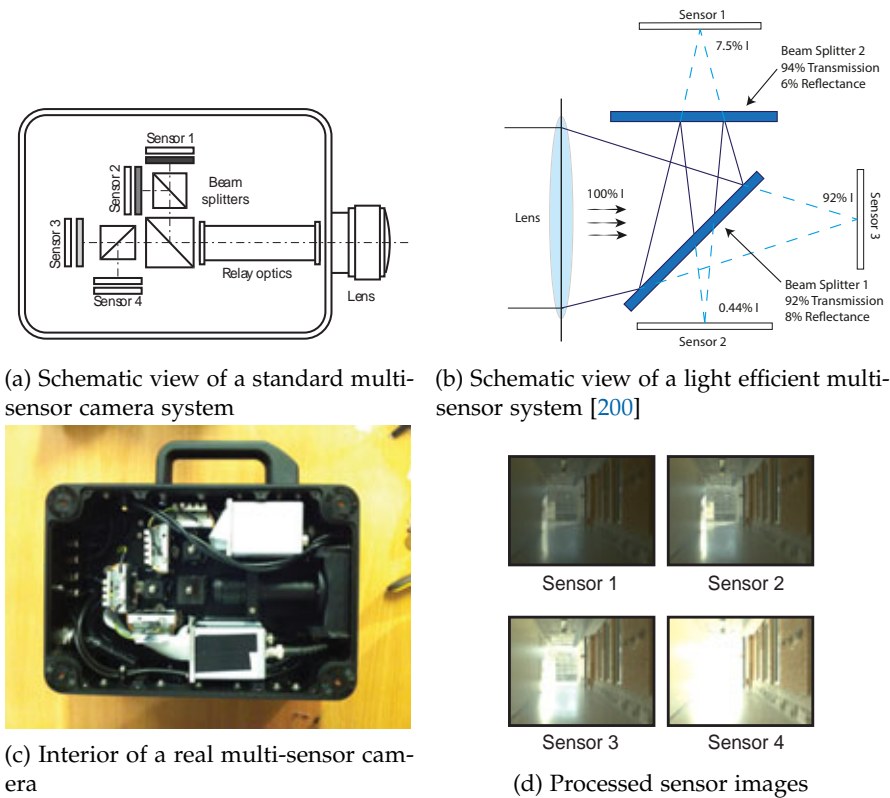


Figure 5.2: Examples of multi-sensor HDR capture systems

camera developed by Spheron VR and Linköping University is shown. The camera uses four sensors with different ND-filters introduced in front, and can capture a dynamic range of up to 24 f-stops.

While traditional setups often waste part of the incident light due to the ND-filters, Tocci et al. [200] recently presented a light efficient compact multi-sensor HDR video system utilizing up to 99.96% of the incident light. Instead of splitting the light equally and relying on ND-filters to achieve different exposures, the optical system is designed so that the light is split unevenly by the beamsplitters onto the sensors. A compact system can be designed by directing some of the light passing through the first beamsplitter back through the beamsplitter a second time, thus reusing the optical elements, see Figure 5.2b.

The use of traditional beamsplitters often result in custom built setups with

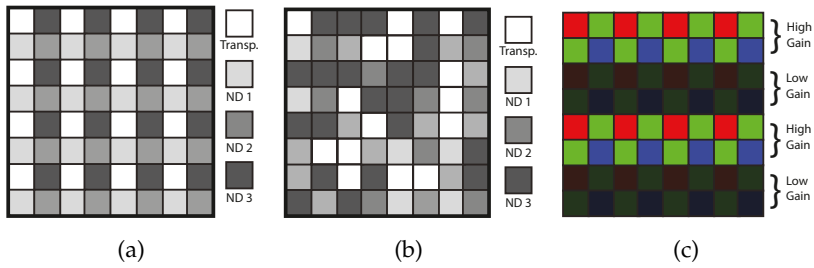


Figure 5.3: HDR images can be captured in a single shot by using filter masks placed in front of the sensor. **a)** spatially varying neutral density filter mask. **b)** spatially varying neutral density filter mask with random ordering. **c)** Varying iso/gain in combination with a bayer filter mask.

specialized camera bodies. To enable HDR-capture using off-the-shelf cameras Manakov et al. [142] recently proposed to insert a small optical element between the lens and body of a standard commercial DSLR camera. This element splits the incident light into four identical images, optically filters each image with a separate ND-Filter, and finally projects each image to a quarter of the same sensor. This setup thus introduces a tradeoff between spatial resolution on the sensor and the number of sub-images/exposures used, but can be used directly with standard cameras.

### Spatially varying sensor exposure

One of the simplest optical designs for achieving single-shot HDR imaging is to introduce an optical mask with spatially varying transmittance over the image sensor, allowing the amount of light reaching the pixels to change over the sensor. The mask can be introduced just in front of the camera sensor similar to a traditional Bayer pattern used to capture color or in the lens element. This introduces a trade-off between the spatial resolution, noise level and dynamic range that can be captured. This approach was first proposed by Nayar and Mitsunaga [162] who introduced a neutral density filter mask with four different optical densities (transparent, highly transmissive, moderately transmissive, and almost opaque) in a regular grid over the image sensor, see figure 5.3a.

When designing spatially varying exposure systems, two main design criteria have to be considered. The first is the number of optical filters to be used in the mask, i.e. how many different exposures of the scene are to be captured. The more filters/exposures are included the higher dynamic range can be captured. However, using a large number of exposures could lead to a low spatial resolution and introduce image noise due to excessive blocking of

incoming light. Secondly, the spatial distribution of the different filters will affect the acquired data. Both regular patterns and stochastic (random) patterns can be used. The choice of pattern is important as pixels with highly transmissive filters will most likely be saturated in high intensity regions and interpolation will be required to reconstruct these values. If the sampling patterns are regular, aliasing artefacts may appear in the interpolation. On the contrary, if a random or pseudo-random pattern is used aliasing can be avoided or significantly suppressed [6, 187]. To capture HDR color images the spatially varying exposure patterns can be combined with Bayer filter designs, for example by using a spatially varying exposure and color filter array [161, 228] instead of a traditional bayer filter.

One disadvantage with approaches introducing a filter array over the sensor is that some of the incident light is blocked in the filters and never reaches the image sensor. This can be a major concern in darker scenes, and in general leads to increased image noise. An alternative solution using all the incident light is based on recent (commercial) imaging sensors where the pixel gain can be varied over the sensor, [3, 80, 205, 207]. The analog pixel gain is proportional to the *ISO* setting found on most cameras. A low gain setting leads to a high saturation threshold but a lower signal to noise ratio in dark regions compared to a high gain setting. This approach can also be combined with traditional Bayer patterns resulting in a multi-gain RAW sensor image where color is captured using a *color filter array* (CFA), e.g. a Bayer pattern. Varying sensor gain can be achieved on existing commercial digital cameras with a customized firmware update. For example, the *Magic Lantern* firmware for Canon cameras. Using the *dual-ISO* module with this firmware allows the sensor to set two ISO settings in a single frame by modifying the CMOS registers flags. This results in images captured with varying gain/ISO for every second row. Figure 5.3 illustrates the spatially varying exposure pattern that is achieved by using the dual-ISO module.

### 5.1.5 Previous methods for HDR reconstruction

In order to reconstruct HDR images and video frames, it is necessary to perform several steps such as: *demosaicing* - reconstruction of full resolution colour images from colour filter array (CFA) data [76], *realignment* - correcting for possible geometric misalignments between multiple sensors [194], *HDR assembly* to fuse the input low dynamic range (LDR) images into the output HDR image [179], and *denoising* to reduce excessive image and color noise [33].

Most existing HDR reconstruction techniques, e.g. [8, 48, 73, 200], treat these steps in a pipeline fashion and perform demosaicing (and realignment) either

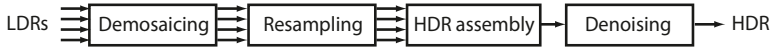


Figure 5.4: The vast majority of previous HDR reconstruction methods have considered *demosaicing*, *resampling*, *HDR assembly* and *denoising* as separate problems.

before or after HDR assembly, see Figure 5.4. This pipeline approach introduces several problems. Demosaicing before HDR assembly as in [48], causes problems with bad or missing data around saturated pixels. This is especially problematic as color channels usually saturate at different levels. Demosaicing after HDR assembly as in [8, 200] causes problems with blur and ghosting unless the sensors are perfectly aligned. For multi-sensor systems using high resolution sensors, it is problematic, costly, and sometimes even impossible to match the sensors so that the CFA patterns align correctly. Tocci et al. [200] report misalignment errors around the size of a pixel despite considerable alignment efforts. This means that overlapping pixels from different sensors might not have the same spectral response. Also, treating demosaicing, resampling and HDR assembly in separate steps makes it difficult to analyze and incorporate sensor noise in the HDR reconstruction in a coherent way.

Recently, after the publication of papers F and E, Heide et al [89] proposed a framework for joint demosaicing, denoising and HDR assembly. Their method is based on solving an inverse problem specified with different global image priors using convex optimization methods. However, their method does not incorporate a statistical model of sensor noise and despite the use of GPUs, their implementation is still computationally expensive due to the expensive global optimization problem. Instead, our contribution is based on a localized approach, enabling rapid parallel processing, while also incorporating a well founded statistical noise model.

## 5.2 Contributions

The main contribution of this thesis in the area of HDR imaging is a unified framework for efficient and robust HDR reconstruction, first presented in paper E and later extended in papers F and G. The reconstruction framework described in these papers unifies the demosaicing, sensor registration, HDR assembly and denoising steps performed in traditional HDR reconstruction algorithms into a single processing operation.

The reconstruction is formulated as an inverse problem where the goal is to estimate the irradiance reaching a simulated HDR image sensor placed

somewhere in the focal plane of the imaging system. Based on a general image formation model, which is applicable to both captured multi-sensor and spatially multiplexed sensor data, the reconstruction takes into account the entire set of non-saturated low dynamic range pixel responses available in a neighborhood around the spatial output location of each HDR pixel. To estimate the irradiance reaching the simulated HDR sensor, we fit local polynomial approximations (LPA) [18] to measured observations in each local region of the estimates. We also present methods for adapting the size of the considered local neighborhood to ensure a good fit for the truncated local polynomial models. The foundation of our approach is to include a statistical noise model describing the heterogeneous sensor noise. The noise model enables us to express the reconstruction as a statistical estimation problem, fitting the local polynomial approximations to the data using a local maximum likelihood approach [198], to minimize the variance of the resulting HDR pixel estimates. Before we discuss the details of the reconstruction algorithm, we first describe the radiometric camera model, that allows us to convert the observed pixel measurement in the input data into a common radiometric space and estimate their variance.

### 5.2.1 Radiometric camera model

Using a radiometric camera model each measured pixel value, given from the input data, can be transformed into a common radiometric space. In this space a transformed pixel response,  $f_i$ , at location  $i$  is proportional to the irradiance at that location. However, as the number of photons that is actually converted to electrons in the photo-diode is dependent on a generally unknown *quantum efficiency* constant, the exact irradiance at the pixel is often difficult to estimate. For convenience, we instead express the pixel response as the number of photo-induced electrons collected per unit time at the image sensor. In paper F a comprehensive radiometric camera model is presented, which takes into account the major sources of noise in the image formation process. We show that the poisson distributed signal dependent photon shot noise and the sensor read-out noise can be accurately described for non-saturated pixel values as realizations of a random variable  $Y_i$  with a normal distribution

$$Y \sim \mathcal{N}(g_i a_i t f_i + \mu_R, g_i^2 a_i^2 t f_i + \sigma_R^2(g_i)) \quad (5.1)$$

where  $g_i$  is the analog pixel gain (inverse of the ISO setting),  $a_i$  is a pixel non-uniformity factor (corresponding to for example varying pixel sizes and other manufacturing errors),  $t$  is exposure time and  $\mu_R$  and  $\sigma_R^2$  are the readout noise mean and standard deviation respectively.

Some previous radiometric models [84] set the standard deviation of the read-out noise to be linearly dependent on the gain setting. However, we have found

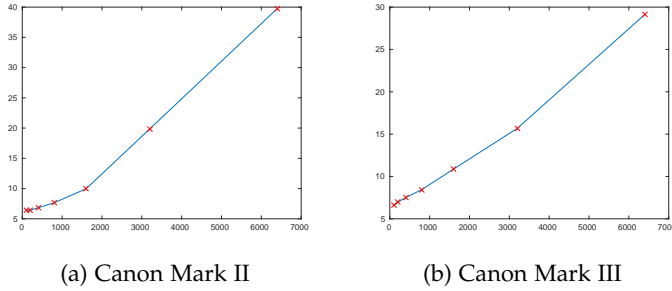


Figure 5.5: The dependence on the readout noise standard deviation,  $\sigma_R$ , on the ISO (gain) setting for **a)** a Canon Mark II sensor and **b)** a Canon Mark III sensor.

that such models cannot accurately describe the observed dependence for some modern camera sensors. Instead we calibrate the readout noise standard deviation,  $\sigma_R(g)$  for each considered gain/ISO setting individually. The observed standard deviation of the readout noise for varying gain settings using a Canon Mark II and a Canon Mark III sensor is shown in figure 5.5.

Dark current noise generated from thermal generation of electrons in the sensor over time is not included in the model as, especially for modern imaging sensors, it has a negligible effect at normal temperatures for exposures under 1-2 seconds [7].

A raw digital value,  $y_i$ , can be transformed to an estimate of the irradiance at the pixel as,

$$\hat{f}_i = \frac{y_i - \mu_R}{g_i a_i t}. \quad (5.2)$$

The signal dependent variance of  $\hat{f}_i$  is given by

$$\sigma_{\hat{f}_i}^2 = \frac{g_i^2 a_i^2 t f_i + \sigma_R^2(g_i)}{g_i^2 a_i^2 t^2}, \quad (5.3)$$

which can be estimated by using the noisy estimate  $\hat{f}_i$  in place of the exact signal  $f_i$ , as this serves as a good approximation to the true variance [7]. An in-depth overview of the radiometric camera model and procedures for calibrating the camera parameters,  $(g_i, a_i, \mu_R, \sigma_R^2)$ , are given in Paper E.

## 5.2.2 Unified reconstruction

After radiometric calibration of the input data, corresponding to the sensor responses of the capture system used, the HDR reconstruction is performed in



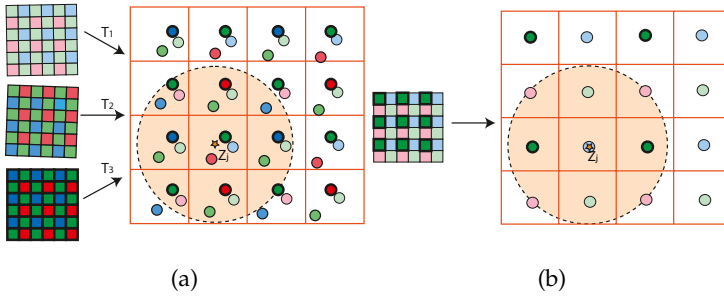


Figure 5.6: HDR reconstruction is performed in a virtual reference system with arbitrary resolution and mapping. A pixel  $Z_j$  in the HDR image is estimated using a local polynomial approximation of nearby samples (within the shaded circle). Saturated and defective pixels are discarded, and samples near the black level, local outliers and samples farther from the reconstruction point contribute less to the radiance estimate. a) For multi-sensor systems the sensor samples are mapped to the reconstruction grid using transformations  $T_s$ . a) For single sensor setups using spatially varying sensor responses, the virtual sensor is usually chosen to coincide with the real sensor.

a virtual reference space, corresponding to a virtual HDR image sensor placed somewhere on the focal plane, see figure 5.6. The virtual sensor dimensions are chosen to reflect the desired output frame, unconstrained by the resolution of the input frames. For single sensor setups using spatial multiplexing, the virtual sensor is usually chosen to coincide with the real sensor. For multi-sensor systems, it is often necessary to register the sensors. Assuming that the sensors,  $s$ , lie in the same virtual focal plane in the split optical path, the relations between the sensors are described using affine transformations,  $T_s$ , mapping sensor pixel coordinates,  $x_{s,i}$ , to the coordinates of the reference output coordinate system,  $X_{s,i} = T_s(x_{s,i})$ . For multi-sensor systems, in general the transformed coordinates,  $X_{s,i}$ , are not integer-valued in the output image coordinates, and for general affine transformations the sample locations  $X_{s,i}$  will become irregularly positioned.

To reconstruct the HDR output frame, we estimate the relative irradiance at each pixel location in the virtual HDR sensor separately using the transformed samples  $\hat{f}_{s,i}(X_{s,i})$ . For each output pixel  $j$  with integer valued, regularly spaced coordinates  $X_j$ , a local polynomial model is fitted to observed samples  $\hat{f}_{s,i}(X_{s,i})$  in a local neighborhood, see figures 5.6 and 5.7. The local polynomial can then be used to predict the irradiance at the HDR pixel. Over the last two decades similar adaptive local polynomial models have seen an increased popularity for applications in e.g. image interpolation, denoising and upsampling. Exam-

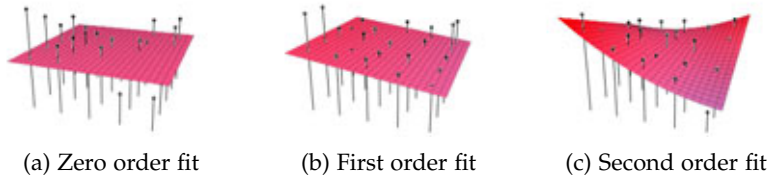


Figure 5.7: For each HDR pixel, nearby noisy pixel measurements (black) are used to fit a local polynomial model (red surface) to the data. The fitted polynomial model can then be used to estimate the irradiance at the HDR pixel location. Note that each output HDR pixel use a separate polynomial expansion centered at the location for that pixel.

ples, of similar approaches, include normalized convolution [113], the bilateral filter [201], and moving least squares [128]. Recently, deep connections have been shown [155, 195] between these methods and traditional non-parametric statistics [137].

To fit the local polynomial models we maximize the *localized likelihood* [198], of the observed samples  $\hat{f}_{s,i}(X_{s,i})$ , defined by a window function that weighs samples based on their distance from the reconstruction location. The following sections present the details of our local polynomial model and the localized likelihood function. For now, we limit the presentation to the reconstruction of a single color channel,  $c = \{R|G|B\}$ , independently of the other channels using an isotropic filtering support. In section 5.2.5 we then shown how the basic algorithm can be improved by using window functions that adapt their size iteratively based on the statistical properties of the noise and the reconstructed signal values, the main contribution in paper G. We also present another approach based on using anisotropic window supports that aligns with images features, the main contribution of paper F, in section 5.2.5. Finally, in section 5.2.6, we show some results and comparisons of our method with previous state-of-the-art methods for HDR reconstruction based on using input data from multi-sensor and dual-ISO data.

### 5.2.3 Local polynomial model

To estimate the irradiance,  $f(x)$ , at an output HDR pixel, we use a generic local polynomial expansion [18] of the irradiance around the output pixel location  $X_j = [x_1, x_2]^T$ . Assuming that the irradiance  $f(x)$  is a smooth function in a local neighborhood around the output location  $X_j$  an  $M$ -th order Taylor series

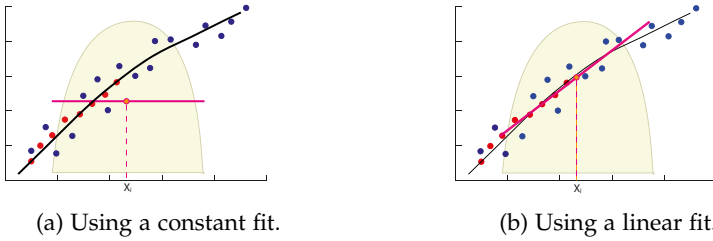


Figure 5.8: The effect of the polynomial order,  $M$ , can be especially significant in areas affected by sensor saturation. The red circles represent a sensor saturating at lower values than the sensor represented by the blue circles. **a)** Using a constant fit,  $M = 0$ , introduces significant bias as more unsaturated measurements are available to the left of the reconstruction point. **b)** Using a linear fit,  $M = 1$ , the bias is drastically reduced.

expansion is used to predict the irradiance at a point  $X_i$  close to  $X_j$  as:

$$\begin{aligned} \tilde{f}(X_i) = & C_0 + C_1(X_i - X_j) \\ & + C_2 \text{tril}\{(X_i - X_j)(X_i - X_j)^T\} + \dots \end{aligned} \quad (5.4)$$

where *tril* lexicographically vectorizes the lower triangular part of a symmetric matrix<sup>1</sup>, and where

$$C_0 = f(X_j) \quad (5.5)$$

$$C_1 = \nabla f(X_j) = \left[ \frac{\partial f(X_j)}{\partial x_1}, \frac{\partial f(X_j)}{\partial x_2} \right] \quad (5.6)$$

$$C_2 = \frac{1}{2} \left[ \frac{\partial^2 f(X_j)}{\partial x_1^2}, 2 \frac{\partial^2 f(X_j)}{\partial x_1 \partial x_2}, \frac{\partial^2 f(X_j)}{\partial x_2^2} \right] \quad (5.7)$$

Given the fitted polynomial coefficients,  $C_{1:M}$ , we can predict the irradiance at the output location  $X_j$  by  $C_0 = f(X_j)$ , and the first order gradients from  $C_1$ . Examples of local polynomials of different order are shown in Figure 5.7. The effect of the polynomial order can be particularly significant in areas of sensor saturation, see figure 5.8 for an illustration.

---

<sup>1</sup> i.e.  $\text{tril} \begin{pmatrix} a & b & c \\ b & e & f \\ c & f & i \end{pmatrix} = [a, b, c, e, f, i]^T$

### 5.2.4 Maximum localized likelihood fitting

To estimate the coefficients,  $C$ , of the local polynomial model, we maximize a localized likelihood function [198] defined using a smoothing window centered around  $X_j$ :

$$\mathcal{W}_H(X) = \frac{1}{\det(H)} \mathcal{W}(H^{-1}X) \quad (5.8)$$

where  $H$  is a  $2 \times 2$  smoothing matrix that determines the shape and size of the window. In our work we use a truncated Gaussian window, however other window functions, such as e.g. Epanechnikov or Tricube windows, has also been considered in the literature for localized likelihood approaches [18].

Let us denote the observed samples in the neighborhood,  $\{\hat{f}_{s,i}(X_{s,i}) : X_{s,i} \in \text{supp}(\mathcal{W}_H(X))\}$ , by  $f_k$  with a linear index  $k = 1 \dots K$ . Using the assumption of normally distributed radiant power estimates  $f_k$ , see Section 5.2.1, the log of the localized likelihood for the polynomial expansion centered at  $X_j$  is given by

$$\begin{aligned} L(X_j, C) &= \sum_k \log(N(f_k | \tilde{f}(X_k), \hat{\sigma}_{f_k}^2) \mathcal{W}_H(X_k)) \\ &= \sum_k \left[ f_k - C_0 - C_1(X_i - X_j) \right. \\ &\quad \left. - C_2^T \text{tril}\{(X_i - X_j)(X_i - X_j)^T\} - \dots \right]^2 \frac{\mathcal{W}_H(X_k)}{\hat{\sigma}_{f_k}} + R \end{aligned}$$

where  $R$  represents terms independent of  $C$ .

The polynomial coefficients,  $\tilde{C}$ , maximizing the localized likelihood function can then be found by the weighted least squares estimate

$$\begin{aligned} \tilde{C} &= \underset{C \in \mathcal{R}^M}{\text{argmax}}(L(X_j, C)) \\ &= (\Phi^T W \Phi)^{-1} \Phi^T W \bar{f} \end{aligned} \quad (5.9)$$

where

$$\begin{aligned} \bar{f} &= [f_1, f_2, \dots, f_K]^T \\ W &= \text{diag}\left[\frac{\mathcal{W}_H(X_1)}{\hat{\sigma}_{f_1}}, \frac{\mathcal{W}_H(X_2)}{\hat{\sigma}_{f_2}}, \dots, \frac{\mathcal{W}_H(X_K)}{\hat{\sigma}_{f_K}}\right] \\ \Phi &= \begin{bmatrix} 1 & (X_1 - X_j) & \text{tril}^T\{(X_1 - X_j)(X_1 - X_j)^T\} & \dots \\ 1 & (X_2 - X_j) & \text{tril}^T\{(X_2 - X_j)(X_2 - X_j)^T\} & \dots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (X_K - X_j) & \text{tril}^T\{(X_K - X_j)(X_K - X_j)^T\} & \dots \end{bmatrix} \end{aligned}$$

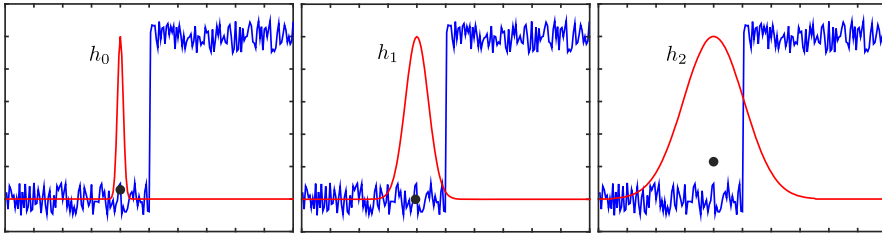


Figure 5.9: The statistically motivated adaptive filter selection rules are based on incrementally increasing the filter supports until the resulting change can no longer be explained by the expected signal noise. See the text for a detailed description.

### 5.2.5 Adapting the filtering window

The smoothing matrix,  $H$ , affects the shape of the window function. The simplest choice of the smoothing matrix is  $H = hI$ , where  $h$  is a global scale parameter and  $I$  is the identity matrix. This corresponds to an isotropic filter support. A large scale parameter setting will reduce the variance, but can lead to overly smoothed images (bias). Ideally, it is desirable to have large window supports in regions where the smooth polynomial model, used for the reconstruction, is a good fit to the underlying signal, while keeping the window size small close to the edges or important image features. The optimal choice of  $h$  is dependent on the sensor characteristics (noise) and the irradiance in the scene. In paper D, we treated the scale parameter  $h$  as a user defined quantity and set it to a global value used for the reconstruction of all HDR pixels. We refer to this method as Local Polynomial Approximation (LPA). However, to obtain a better trade-off between image sharpness and noise reduction, in paper F we propose to locally adapt the smoothing parameter  $h$  to image features and the noise level of observed samples. For this purpose we describe an iterative algorithm for selecting the locally best smoothing parameter,  $h$ , for each HDR pixel estimate,  $z_j$ , individually.

The intuition for the scale selection algorithm is that the scale parameters should be increased as long as the variance in the signal reconstructions can be explained by the underlying signal noise. Figure 5.9 illustrates how a signal value, the black point, is being estimated using a kernel with a gradually increasing smoothing parameter,  $h$ . When the smoothing parameter  $h$  is increased from  $h_0$  to  $h_1$ , i.e. a higher degree of smoothing, the variance in the estimated value can be explained by the signal variance. When the smoothing parameter is increased from  $h_1$  to  $h_2$ , the kernel reaches the step in the signal and the estimation at the black point can no longer be explained by the signal variance.

Smoothing parameter  $h_1$  thus produces a better estimate.

The adaptation of the smoothing parameter,  $h$ , is carried out iteratively. We start with a small  $h$  and try to increment it in small steps. In each iteration we estimate the signal value and its variance. An *update rule* is then applied, which determines if  $h$  should be increased further or not. This is repeated until the update rule terminates the recursion or the maximum  $h$  value,  $h_{max}$ , is reached. In paper G we presented two such update rules, described in more detail below.

**EVS - Update rule** The first update rule is built on the intuition that if the weighted mean reconstruction error is larger than the weighted mean standard deviation, the polynomial model does not provide a good fit to the underlying image data. The smoothing parameter,  $h_i$ , is iteratively increased with an increment  $h_{inc}$ . In each iteration,  $i$ , the EVS update rule computes the weighted reconstruction error  $e_i$  as

$$e_i = \sum_k W(k, k) |\tilde{f}(X_k) - \hat{f}_k|, \quad (5.10)$$

This estimate is compared to the standard deviation of the HDR pixel estimate that can be computed using the covariance matrix  $M_C$  for the fitted polynomial coefficients,  $\tilde{C}$ , given by

$$M_C = (\Phi^T W \Phi)^{-1} \Phi^T W \Sigma W^T \Phi (\Phi^T W^T \Phi)^{-1} \quad (5.11)$$

where  $\Sigma = \text{diag}[\sigma_{f_1}^2, \sigma_{f_2}^2, \dots, \sigma_{f_k}^2]$  is the variance of the observations. During the iterations, the smoothing parameter,  $h_i$ , is updated to  $h_{i+1} = h_i + h_{inc}$  as long as the weighted reconstruction error,  $e_i$ , is smaller than the standard deviation  $\tilde{\sigma}_i$ , i.e.  $e_i < \Gamma \tilde{\sigma}_{z_j, h_i}$ , where  $\Gamma$  is a user specified parameter controlling the trade-off between levels of denoising applied by the kernel.

**ICI - Update rule** The second update rule proposed in Paper F is based on the Intersection of Confidence Intervals (ICI) rule for adaptive scale selection, first developed for non-parametric statistics applications [71] and later also applied to other imaging problems such as denoising [107]. Using ICI, the smoothing parameter,  $h_{min} \leq h_i \leq h_{max}$ , is iteratively increased. For each iteration,  $i$ , the ICI rule determines a confidence interval,  $D_i = [L_i, U_i]$  around the estimated signal as:

$$L_i = \hat{z}_{j, h_i}(x) - \Gamma \tilde{\sigma}_{z_j, h_i} \quad (5.12)$$

$$U_i = \hat{z}_{j, h_i}(x) + \Gamma \tilde{\sigma}_{z_j, h_i} \quad (5.13)$$

where  $\hat{z}_{j, h_i}(x)$  is the estimated radiant power given the scaling parameter  $h_i$  and  $\tilde{\sigma}_{z_j, h_i}$  is the weighted standard deviation of this estimate computed using Eq.

5.11.  $\Gamma$  is a scaling parameter controlling how wide the intersection interval is. During adaptation,  $h_i$  is increased as long as there is an overlap between the confidence intervals, i.e.  $h_i$  is updated to  $h_{i+1} = h_i + h_{inc}$  if there is an overlap between  $D_i$  and  $D_{i+1}$ . In practice, we utilize  $\Gamma$  as user parameter enabling a intuitive trade-off between image sharpness and denoising. A detailed overview of the ICI rule and its theoretical properties can be found in [18].

### Anisotropic window supports

A problem with isotropic window supports is that they limit the adaption to the underlying image features. If, for example, an output pixel is located near a sharp edge, the neighboring samples cannot be represented accurately with a finite polynomial expansion, and thus a small isotropic window support needs to be used. However this limits the number of observed samples that are used to fit the polynomial and hence increases the noise in the estimate. It is therefore desirable to adapt the window support so that it can include several observations on the same side of the edge, but not the ones on the other side of the edge. Intuitively the shape of the window support should be: circular and relatively large in flat areas to reduce noise, elongated along edges, and small in textured areas to preserve detail. In paper F we propose to use a method inspired by previous work in image denoising [195] to adapt anisotropic window functions for this purpose.

To adapt the window functions we use a two step approach: First, we use a regular reconstruction with isotropic window support and  $M \geq 1$  to compute an initial estimation of the gradients,  $\left[ \frac{\partial f(X_j)}{\partial x_1}, \frac{\partial f(X_j)}{\partial x_2} \right] \forall j$ , in the image. In a second step we then adapt the smoothing matrix,  $H_j$ , for each HDR output pixel, to reflect the dominant directions of the estimated gradients in a neighborhood around  $X_j$ . The result of this process is elliptical window functions that elongate along edge structures in the image, as well as adapt their size based on the signal smoothness in the neighborhood.

To robustly estimate the window supports in areas with noisy gradient estimates we use a parametric approach, that describes the gradient covariance in the neighborhood using three parameters,  $\sigma_j$  describing an elongation along the principal directions,  $\theta_i$  describing a rotation angle and  $\gamma_i$  describing an overall scaling. These parameters are fitted to the local gradient structure using regularization to ensure that the window is not adapted to drastically, eg. shrinking to a single point. The details of the anisotropic window estimation is given in paper E, following the approach proposed by Takeda et al. [195], but with the difference that we adapt the window functions for each HDR output pixel instead of the kernel functions for each observed measurement.

### Color reconstruction

For isotropic window function we generally consider different scale parameters for the green and red/blue color channels. This as the standard bayer pattern uses more green samples per unit area. Generally, different color channels also saturate at different levels, which can require different minimum values of  $h$  to make sure that enough samples are used to fit the local polynomial model, ie to ensure that  $(\Phi^T W \Phi)$  is invertible in equation (5.9).

In paper F, we proposed to adapt anisotropic windows to the gradient information in the green channel first, and then use these adapted windows to also estimate the red and blue channels. This effectively forces the interpolation to be performed along the same signal features across all color channels. This often leads to less perceptually disturbing color artifacts and fringes in high frequency regions. We refer to this reconstruction method as *Color Adaptive Local Polynomial Approximation (CALPA)*.

## 5.2.6 Practical results and comparisons

In this section we provide a brief overview of how the developed HDR reconstruction framework performs in two applications. HDR reconstruction for multi-sensor systems and image data captured using spatially varying per pixel gain (ISO), also known as dual-ISO capture. We show results for both simulated data from virtual sensors and data obtained from real camera systems.

### Multi-Sensor HDR-video capture

Papers D and E evaluated the proposed unified framework for HDR video capture using multi-sensor imaging systems.

In Figure 5.10, LPA and CALPA are compared to the previous state-of-the-art multi-sensor reconstruction method recently proposed by Tocci et al. [200]. The reconstructions were generated from 3 virtual exposures of a simulated Kodak Kai-04050 sensor captured 5 f-stops apart. The lowest exposed and the highest exposed sensors are perfectly aligned, while the middle sensor has a translational offset of  $[0.4, 0.45]$  in pixel units. This is similar to the alignment errors reported by Tocci et al. for their camera system. The previous state-of-the-art reconstruction method of Tocci et al. suffers from noise. In contrast, the proposed methods LPA and CALPA which are based on a statistical sensor noise model incorporates information from all non-saturated sensor observations to obtain better estimates. CALPA effectively denoises the image by using larger filtering supports in flat image areas. Compared to the other methods, CALPA also reduces unwanted color artifacts in areas of high spatial frequency.



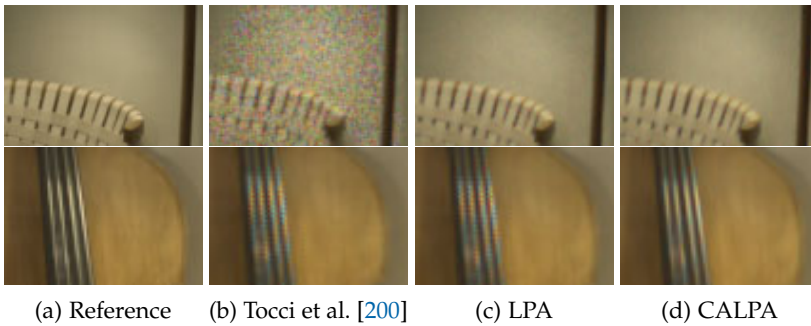


Figure 5.10: Magnifications from reconstructions of a scene virtually exposed with 3 Kodak Kai-04050 sensors 5 f-stops apart. a) Tonemapped HDR reference images. b) Reconstructions by the method proposed by Tocci et. al. [200], c)LPA d)CALPA. In areas of sensor transitions, the method by Tocci et. al. exhibits severe noise as only one sensor is used at a time, as can be seen in the magnification on the top row. In contrast LPA handles sensor transitions gracefully due to the incorporation of a sensor noise model and a statistical maximal likelihood fit to all sensor data. CALPA show even better results due to the inherent denoising using adaptive filtering supports. Using CALPA also help in reducing unwanted color artifacts.

In paper D and E we also evaluated our approach on data from a real, custom built, multi-sensor HDR camera system developed in collaboration between Spheron VR [1] and Linköping University. This camera is based on four Kodak KAI-04050 CCD video sensors that receive different amounts of light from an optical system with a common lens, a four-way beam splitter and four different ND filters. The sensors have 12 bits linear A/D conversion, and the exposure scalings cover a range of  $1 : 2^{12}$ , roughly equivalent to a dynamic range equivalent to  $12 + 12 = 24$  bits of linear resolution, commonly referred to as “24 f-stops” of dynamic range. The system allows for capture, processing and off-line storage of up to 32 frames per second at 4 Mpixels resolution, amounting to around 1 GiB/s per second of raw data. Figure 5.11 shows an example frame captured using this HDR-video camera and reconstructed with our method (LPA). Using our CUDA implementation with  $M = 0$  the four input frames are processed to an output HDR frame, with a resolution of  $2336 \times 1752$  pixels, at 26 fps on an Nvidia GForce 680 GPU.

Many more detailed comparisons between LPA, CALPA and previous methods performing demosaicing before or after HDR assembly are presented in papers D and E.



Figure 5.11: Left: A full resolution locally tonemapped frame from a video sequence captured with the HDR video camera. The image is reconstructed from four 4 Mpixel sensors at 26 fps. Right: Four gamma mapped images 2 f-stops apart from the same HDR-image.

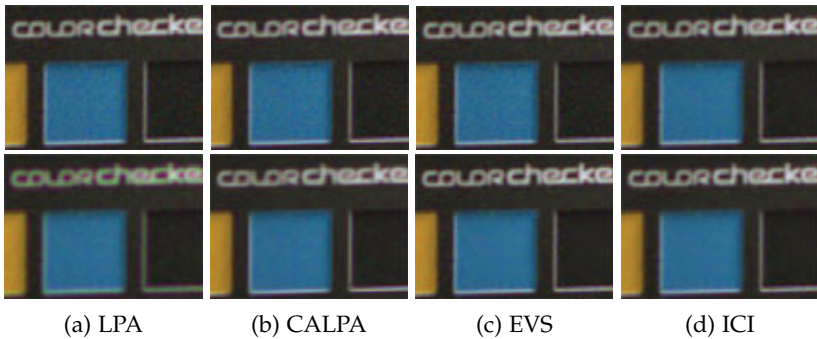


Figure 5.12: A comparison on simulated dual-ISO data from a virtual Canon sensor. (a) Local polynomial approximation (LPA) with isotropic windows using  $M = 2$ , top)  $h = 0.6$  and bottom)  $h = 1.4$  (b) CALPA  $M = 2$  top)  $h = 0.6$  and bottom)  $h = 1.4$ , (c) ICI  $M = 2$  top)  $\Gamma = 0.6$  and bottom)  $\Gamma = 1.0$  (d) EVS  $M = 2$  top)  $\Gamma = 0.6$  and bottom)  $\Gamma = 1.0$

### Spatial multiplexing using Dual-ISO Capture

The application of the unified reconstruction framework for sensors with spatially varying pixel responses is explored in Paper G. In particular we consider HDR reconstruction of sensor data obtained with a varying gain (ISO) set-

ting per pixel. This approach is motivated by recent developments in camera firmware, that allows setting the gain for different rows of an image sensor separately using, for example, a standard Canon DSLR camera with the Magic Lantern module Dual-ISO [3].

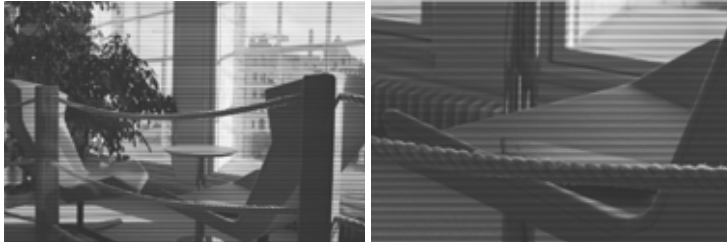
Figure 5.13c shows an example scene captured using a Canon Mark III camera using ISO setting 100 ( $g = 0.30$ ) and 1600 ( $g = 4.8$ ) for each other row. Using our reconstruction method with the ICI update rule we can reconstruct high-quality HDR images from this input data.

Figure 5.12 shows a comparison between the different variations of our proposed framework. In this example ICI performs best, however for some scenes with high-frequency edges using anisotropic window supports can reduce unwanted color artifacts around sharp edges in the image.

### 5.3 Summary and Future work

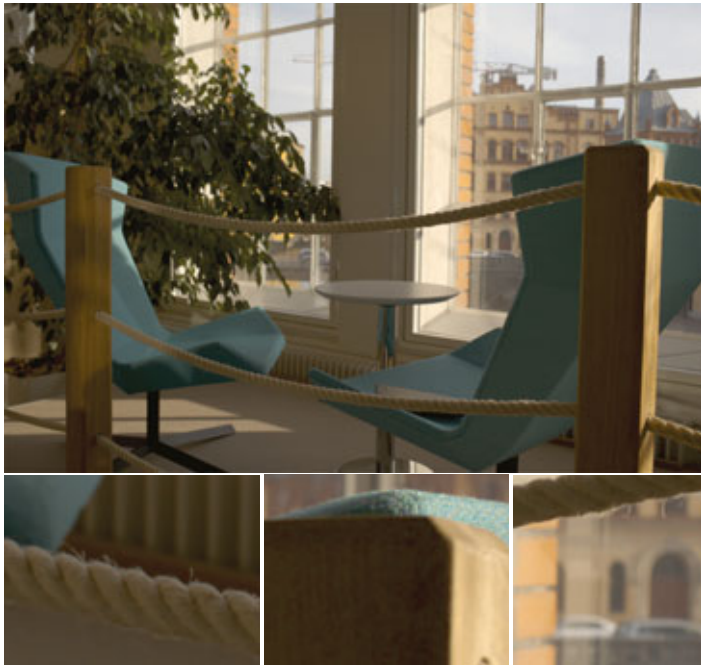
In this chapter we presented a novel unified HDR reconstruction framework based on a statistical model of the sensor noise. Using a non-parametric regression, based on local polynomial models, we estimate HDR pixels from noisy observations using localized maximum likelihood techniques. We also showed that by using adaptive window functions for defining the locality of the local polynomial model the estimates can be improved. The main applications of our framework discussed here were the reconstruction of multi-sensor data and image data obtained using spatially varying gain settings. However, the framework is general and should work in other scenarios as well. In future work it would be interesting to consider combinations of multi-sensor systems and sensors with spatially varying gain. For example, combining the approach of Manakov et al. [142] for multi-exposure imaging using optical splitting with a dual-ISO configured sensor [3]. Such a setup could also enable fewer sensors to be used in traditional multi-sensor system, thereby making them less bulky, while still maintaining a large dynamic range with low noise.

Possible developments of the reconstruction framework itself includes adapting the window functions not only spatially in each frame, but also extend them over several frames for video reconstruction. Using an approach similar to that of [196] for this purpose is an interesting venue for future work. It would also be interesting to investigate the combination of the scale selection rules for isotropic supports which are based on the statistical properties of the noise, presented in paper F, with the anisotropic window adoption methods presented in paper E. For example using the ICI rule as a last step to scale the adapted shape of the anisotropic window.



(a) Raw image

(b) Zoom in of Raw image



(c) Reconstructed from dualISO data with ISO 100-1600 using ICI,  $M = 2$ ,  $h \in [0.6, 5.0]$  and  $\Gamma = 1.0$

Figure 5.13: Top: Raw data from the sensor using the dual-ISO module in Magic Lantern running on a Canon 5D Mark III. Bottom: Tonemapped result image using the presented noise-aware reconstruction method.

## Surface Reflectance Models

The reflectance of an object, defined by its *material* properties, describes the way that it appears under varying viewing and illumination conditions. Different materials give rise to distinct visual appearances. For example, the reflectance of an object describes if it appears matte or shiny, plastic or metallic, smooth or rough. Figure 6.1, shows an example of how the visual appearance of a simple geometrical shape can be transformed by rendering it using different material properties. A detailed description of appearance is not only an essential requirement for realistic image synthesis, but can also help in image interpretation such as 3D reconstruction, view interpolation, and object recognition. Modeling reflectance, however, is a complex task. In general, an object can reflect a different amount of light at each position on its surface, for each wavelength, and for each possible direction of incident and excitant light. To completely characterize reflection, a seven dimensional function would be necessary to predict the reflected light for a given set of these parameters (neglecting a possible dependence on time). Yet, in practice, the complexity can often be reduced for many materials, for example by observing that the spatial variations only affect certain components of the reflectance. A comprehensive overview of appearance modeling can be found in many excellent books [52, 54] and surveys [124, 223].

In this chapter we limit the presentation to an important subset of a complete appearance model, namely the reflection at a single point at the surface described by the *Bidirectional Reflectance Distribution Function* (BRDF). Spatially varying appearance models can easily be derived from BRDFs by, for example, using textures to describe spatially varying parameters [52, 54, 223]. In sec-

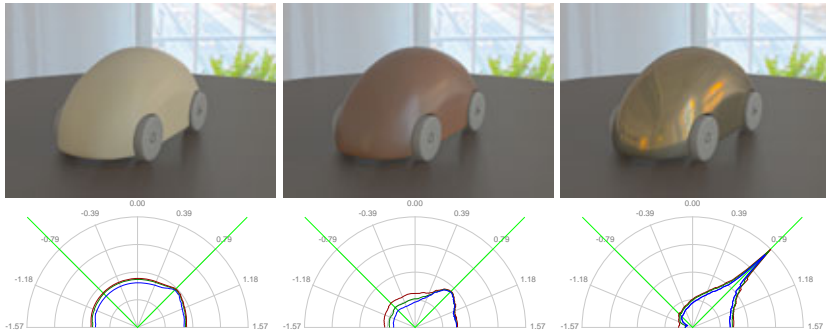


Figure 6.1: The visual appearance of objects is dependent on the reflectance properties of its materials. Upper row) A simple model of a toy car rendered with typical reflectance behaviors, corresponding to (from left to right), a diffuse reflectance, semi-glossy and glossy reflectance. Lower row) Corresponding scattering distributions (R,G,B color channels) in polar coordinates for a fixed incident illumination direction and azimuthal viewing angle.

tion 6.1 we provide an overview of previous methods for representing BRDFs, including parametric and data driven modeling techniques. We also discuss common methods for capturing the reflectance of real world materials. In section 6.2 we then discuss the contributions presented in paper H . These include a study of different BRDF model parameterizations and the derivation of two new parametric models that outperform previous models for glossy surface reflectance.

## 6.1 BRDF acquisition and representation

In this section we provide an overview of existing methods for representing BRDFs. We start by considering the theoretical foundations of BRDF models, including a discussion of the constraints that has to be fulfilled for physically based models. We then briefly discuss common techniques for capturing reflectance data from real world objects. Finally we discuss methods for representing BRDFs using both parametric and data driven models.

### 6.1.1 Theoretical foundations

The Bidirectional Reflectance Distribution Function (BRDF) [165] is formally defined as the relation between the reflected radiance in direction  $\omega_o$ , to the

irradiance arriving from direction  $\omega_i$ , at a surface point with normal  $n$ , as:

$$\rho(\omega_i, \omega_o) = \frac{dL(x \rightarrow \omega_o)}{dE(x \leftarrow \omega_i)} = \frac{dL(x \rightarrow \omega_o)}{L(x \leftarrow \omega_i)|n \cdot \omega_i|d\omega_i'} \quad (6.1)$$

where the last step was derived using the definition of irradiance in terms of radiance, see equation (2.7). Note that the BRDF is defined using differential quantities, the reason is that it makes little sense to talk about the amount of light arriving from exactly one direction and being reflected into exactly one outgoing direction. To get a better understanding of this definition, it is intuitive to consider the BRDF in the context of relating the incident and reflected radiance at a surface. By multiplying both sides of (6.1) by the denominator and integrating over incoming directions,  $\omega_i$ , we obtain an expression for the reflected radiance:

$$L(x \rightarrow \omega_o) = \int_{\Omega} \rho(\omega_i, \omega_o) L(x \leftarrow \omega_i) |n \cdot \omega_i| d\omega_i' \quad (6.2)$$

*Physically plausible* BRDF functions [133] are naturally positive and real valued, ie

$$\rho(\omega_i, \omega_o) \geq 0 \quad \forall \{\omega_i, \omega_o\}. \quad (6.3)$$

They are also energy conserving, i.e. a real surface cannot reflect more light than incident on it. More precisely the following property must hold:

$$\int_{\Omega} \rho(\omega_i, \omega_o) |n \cdot \omega_i| d\omega_i \leq 1 \quad \forall \omega_o. \quad (6.4)$$

Finally, they should obey the *Helmholtz reciprocity* principle, which states that the incident and outgoing directions can be swapped with no effect on the reflectivity, expressed by

$$\rho(\omega_i, \omega_o) = \rho(\omega_o, \omega_i). \quad (6.5)$$

The reflection of a perfectly smooth surface (mirror-like reflection) can be described by considering the solution to Maxwell's equations [29]. The solution give simple relationships for the intensity, spectral properties and directionality of the reflection. The reflection direction,  $\omega_r$ , has the same angle to the normal as the incident direction,  $\omega_i$ , and is located in the same plane as the incident direction and the surface normal. The fraction of light that is reflected is given by the *Fresnel equations* [29], which depends on material type (*metals* - conduct electricity, or, *dielectrics* - zero conductivity) and parameters. In general, for dielectrics the fraction of light that is reflected increases sharply for grazing angles (large incident angles,  $\theta_i$ ). For metals there is not such a dramatic increase

as the angle of incidence increases, as the reflectance is high for all angles. As the Fresnel equations depend on wavelength specific material properties, such as the index of refraction, the fraction of reflected light varies with wavelength, resulting in different color appearances. However, most real surfaces are not perfectly smooth. Instead the reflection is affected by the microstructure of the surface, and for some materials by scattering inside the object (for example paper). The implications of this is that, except for a few special cases such as mirrors, the reflections of real materials are instead distributed around the reflection direction in a *specular lobe*. Matte, *diffuse*, materials represents the opposite extreme to specular materials and reflect the incident light equally in all outgoing directions. Most real materials exhibit *glossy* reflection, which is neither ideally specular/mirror-like or diffuse. See figure 6.1 for an illustration.

### 6.1.2 Parameterizations and symmetry properties

The perhaps simplest parametrization of a BRDF is by considering it as a function of the incoming  $\omega_i$  and outgoing  $\omega_o$  directions, expressed in their standard spherical coordinates,  $\{\theta, \phi\}$ , defined with respect to the tangent/bi-normal/normal frame,  $\{\mathbf{t}, \mathbf{b}, \mathbf{n}\}$ , of the surface at the point of interest. The BRDF can then be written as  $\rho(\theta_i, \theta_o, \phi_i, \phi_o)$ . In the following, we will refer to this parametrization as the *standard parametrization*.

In general, the BRDF is a four dimension function. However, many materials only depend on the relative angle between the incident and outgoing directions, i.e. the reflectance remains the same if the incoming and outgoing vectors are rotated by the same amount around the surface normal. These materials can be modeled by simpler *isotropic* BRDFs, which are three dimensional functions, with the following property:

$$\rho(\theta_i, \theta_o, \phi_i, \phi_o) = \rho(\theta_i, \theta_o, |\phi_i - \phi_o|) \quad (6.6)$$

where  $\{\theta, \phi\}$  denote the standard spherical coordinates. The set of BRDF functions for which (6.6) does not hold are called *anisotropic* BRDFs, and include materials such as brushed metal, satin and velvet (generally any material with directionally dependent micro structures).

A problem with the standard parametrization is that it can make it difficult to study specular or highly glossy materials. A more practical alternative is provided by the *halfway vector* defined as  $\mathbf{h} = \frac{\omega_i + \omega_o}{2}$ . Many parametrical BRDF models make use of the *normalized halfway vector*,  $\hat{\mathbf{h}} = \frac{\mathbf{h}}{\|\mathbf{h}\|}$ , described in spherical coordinates by  $\{\theta_h, \phi_h\}$ , these models will be discussed in section 6.1.4. A complete, general, BRDF parametrization based on the normalized halfway vector was provided by Rusinkiewicz [184]. This parametrization is based on introducing a complementary vector, the *difference vector*,  $\hat{\mathbf{d}}$ , that describes the



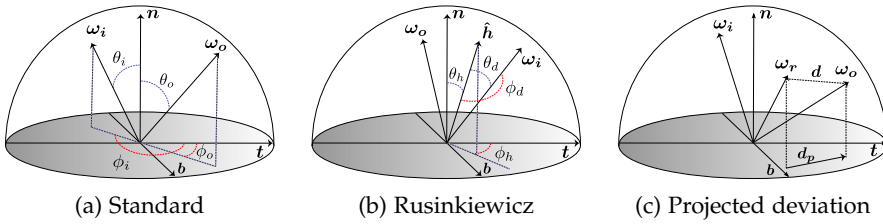


Figure 6.2: Illustrations of: **a)** The *standard parametrization* using the incident and outgoing directions. **b)** Rusinkiewicz *halfway/difference parametrization* using the halfway and the difference vector. **c)** The *projected deviation vector parameterization* using the projection of the deviation vector, defined as the difference between the reflection direction and the outgoing vector, onto the unit disk. See the text for details.

outgoing (or incident) direction  $\omega_o$  in a reference frame where  $\hat{h}$  is treated as the normal. It can thus be computed from the outgoing direction by applying the corresponding rotations,  $\hat{d} = \text{rot}_{\mathbf{b}, -\theta_h} \text{rot}_{\mathbf{n}, -\phi_h}[\omega_o]$ , see figure 6.2b. A general BRDF can be described in Rusinkiewicz *halfway/difference parametrization* as,  $\rho(\theta_h, \phi_h, \theta_d, \phi_d)$ , and for isotropic BRDFs by  $\rho(\theta_h, \phi_h, \phi_d)$ . Romeiro et al. [183] noticed that many isotropic BRDFs remain roughly constant for rotations of the input and output directions (as a fixed pair) about the halfway vector. In other words, the reflectance is approximately constant for varying azimuthal angles of the difference vector,  $\phi_d$ . Such BRDFs are said to be *bivariate* and can be expressed as a function of just two variables, ie  $\rho(\theta_h, \theta_d)$ .

Neumann et al. [163] proposed a class of BRDF models parametrized by the projected difference vector between the incident and outgoing directions, which can be computed by  $\mathbf{d}_p = \mathbf{h} - (\mathbf{h} \cdot \mathbf{n})\mathbf{n}$ , see figure 6.2c. We study the properties of this parametrization in section 6.2.1 and use it for deriving one of the BRDF models presented in paper H, discussed in section 6.2.2.

For more details on BRDF parameterizations see for example [192] and [22] for a recent summary and comparison. In section 6.2.1 we describe an analysis of different BRDF parameterizations for use in parametric models.

### 6.1.3 Acquisition

To measure the BRDF function on the surface of an object, one needs to measure the amount of light that reaches a sensor after being scattered on the surface. Various methods have been developed for measuring BRDFs, either controlling or measuring the incident illumination and the shape of the object so that the captured scattered light can be interpreted to provide reflectance data.

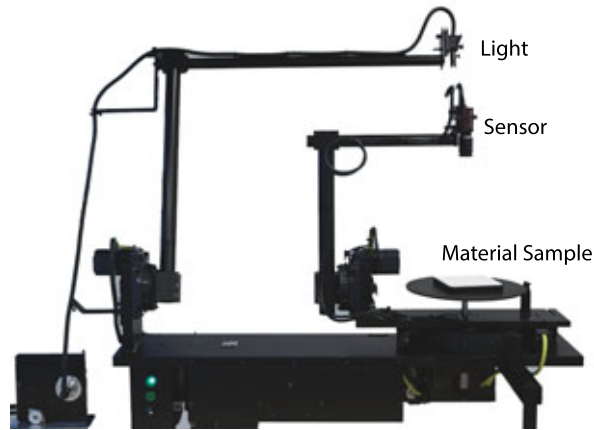


Figure 6.3: A multi-spectral gonioreflectometer used in production at IKEA Communications for measuring BRDFs, developed in collaboration between Rayspace AB and Linköping University.

A standard method for accurate reflectance measurement is to use a *gonioreflectometer*. The device consists of a (camera) sensor that captures the reflected light from a material sample that is illuminated by a controllable light source. Both the sensor and light source can be varied over the visible hemisphere of the material sample to capture the complete four dimensional BRDF function. An example of a multispectral gonioreflectometer developed in collaboration with Rayspace AB and Linköping University, that is now used in production at IKEA Communications AB is shown in figure 6.3. Gonioreflectometers have been used in many fields of science for characterizing surface reflectance for various applications, such as investigating properties of print paper, inferring microstructures on surfaces, and material design for military applications where it is desirable to reduce the reflectance footprint, see e.g. [99, 158, 177] for more details on the design of high-precision gonioreflectometers.

A problem with gonioreflectometers is that dense sampling over both the incident,  $\omega_i$ , and outgoing,  $\omega_o$ , directions can require long capture times. Some work has therefore considered adaptive sampling techniques [66] or methods to reconstruct complete BRDFs from just a few samples (e.g. only using 20 measurements) by using regularization based on prior information on the expected reflectance of common BRDFs [166]. Other more efficient measurements techniques using more complex setups with curved mirrors etc. have also been proposed [45, 70, 220]. Marschner et al. [145] proposed a particularly efficient technique to obtain a isotropic BRDF by capturing images of a sphere made out

of a homogenous material using a (1D) set of different illumination positions. This technique was used by Matusik et al. [147] to capture a large database (100 materials) of isotropic BRDFs that has since its publication seen widespread use in the computer graphics literature. Methods for estimating BRDFs with unknown illumination have also been proposed [182]. However, the inherent ambiguity between illumination and the reflectance is difficult to resolve.

### 6.1.4 BRDF Models

Over the last decades, a large number of BRDF models have been developed in computer graphics and other fields. While the goal of reproducing the behavior of real reflectance functions are the same, the methods to derive the models vary. *Empirical models* (or phenomenological model) are based on using relatively simple analytical functions to simulate the observed visual appearance of real reflectance without considering the underlying physical phenomena. On the other hand, *physically derived models* are based on deriving the BRDF function as a result of some underlying physical process, for example the statistical average of the reflection from small specular or diffuse microfacets. A third, more recent, approach is to use *data driven models*, where the BRDF model is represented explicitly by measured reflectance data, often stored using efficient representations.

#### Empirical models

Many BRDF models include a term to account for an ideal diffuse component of the reflectance. The reflectance of ideally diffuse materials can be modeled by the *Lambertian* BRDF:

$$\rho_L(\omega_i, \omega_o) = \frac{K_d}{\pi} \quad (6.7)$$

where  $K_d \in [0, 1]$  is the diffuse reflection (or *albedo*), describing the fraction of light that is diffusely reflected. This BRDF model is normalized so that the energy of the incident light is equal to the energy of the excitant light when  $K_d = 1$ , see equation 6.4. As Lambertian BRDFs don't account for the incident and outgoing directions, they appear, equally bright for all viewing angles. Many common surfaces has an approximatly diffuse reflectance, for example, matte white paper.

Most empirical models combine the Lambertian BRDF with a term to model specular (glossy) reflections. One of the earliest reflection models that was used in computer graphics, the *Phong* reflection model [175], is based on expressing the specular lobe as a function of the angle  $\theta_s$  between the view direction,  $(\theta_o, \phi_o)$ , and the angle of mirror reflection,  $(\theta_i, \phi_i + \pi)$ . While this model has

seen widespread use, it is a very crude model that often gives a unnatural plastic look. While the original model was not energy conserving, later work has derived physically plausible alternatives [126, 133].

The *Ward model* [220], models the specular component using an exponential function, and can be used to model both isotropic or anisotropic reflectance. The complete model for isotropic reflectance is given by:

$$\rho_W(\omega_i, \omega_o) = \frac{K_d}{\pi} + K_s \frac{1}{\sqrt{\cos(\theta_i) \cos(\theta_o)}} \frac{e^{-\frac{\tan^2 \theta_h}{\alpha^2}}}{4\pi\alpha^2} \quad (6.8)$$

where  $K_d$  and  $K_s$  are the diffuse and specular reflection scalings, and,  $\alpha$  is a *roughness* parameter, defining the shape of the specular lobe (smaller values gives sharper highlights). Alternative normalization terms has also been proposed for the Ward model [57].

Another popular empirical model is the *Ashikmin-Shirley model* [16] which includes an explicit term for Fresnel reflectance, modeling the increase in specular reflection as the angle of incidence increases. The model is constructed so that it is energy conserving by reducing the diffuse component proportionally to the increase in the specular reflection. It also fulfills the other requirements of a physically plausible BRDF model.

While many of the previous models have been derived manually, Brady et al. [31] proposed to use genetic programming techniques to automatically derive simple parametric models that provides good fits to measured data.

### Physically derived models

Another important class of BRDF models are derived by applying basic physical principles of light-surface interactions to a surface's microscopic structure to derive the bulk reflection properties. This approach has been inspired by other communities, for example, optical engineering, where there has been a long tradition of studying light scattering from rough surfaces and to derive surface statistics from measured reflectance data [193]. Rather than explicitly modeling the surface geometry at the microscopic level explicitly, statistical models of the surface height or slope is often used. Given a statistical distribution of the microstructures the expected reflection can then be computed by using either geometrical or wave optics considerations.

One of the earliest BRDF models used in computer graphics, the *Blinn-Phong model* [27], is based on evaluating a normal distribution function centered around the halfway vector. The method is motivated by considering the surface as a large number of, specular, microfacets which each has the strongest reflection (mirror reflection) between  $\omega_i$  and  $\omega_o$  for facets pointing towards the

halfway vector. The *Cook-Torrance model* [41], inspired by the Torrance-Sparrow model [203] developed in physics, also included the interactions between the microfacets on the surface, due to shadowing and masking effects. Their model for the interactions are based on principles of geometrical optics and thus assumes that the wavelength of light is small compared to the scale of the microfacets. An important insight in their model was that the introduction of a Fresnel term can be used to approximate color shifts for specular reflections (common for metals) and increasing reflectance towards grazing angles. Many variants of such microfacet based BRDF models derived from geometrical optics have appeared over the last decades, varying the distribution of microfacet orientations, shadowing approximations and the microfacet reflections [17, 19, 169, 219]. A good overview of the derivation of accurate shadow and masking functions is given in [90]. Some models have also been developed by constructing surface microstructure and then computing a table of reflectance values by running numerical simulations of the light scattering [34, 222].

He et al. [88] used physical principles based on wave optics to study reflections from microfacets of a size comparable to that of the wavelength of the incident light (geometrical optics is no longer a valid model for these cases). Their model is based on Kirchhoff diffraction theory. In paper H, we propose a new BRDF model inspired by the *Rayleigh-Rice* theory [193] for optical scattering from smooth surfaces (surfaces with structures smaller than the wavelength of light). Our model of the microfacet distribution is influenced by the ABC model first proposed by Church [38] to model measured surface profile statistics.

### Fitting parametric models to data

Using measured reflectance data, the parameters of analytical BRDF models can be fitted to simulate the appearance of specific materials. Ngan et al. [164] compared the accuracy of several fitted parametric models using the MERL database as input data [147]. As the parametric models seldom fit the measured reflectance exactly, important considerations for the fitting procedure are the choice of suitable metrics [19, 164], for example perceptually motivated metrics, and the choice of non-linear optimization techniques [56, 229].

### Data driven models

A simple approach for representing a BRDF is to explicitly store the values of measured reflectance for a dense set of sampled ingoing and outgoing directions [147]. However, this approach can require a very large amount of storage, especially if several wavelengths/color bands are considered. In addition, such naive representations are not straight forward to edit and interact with. Using more efficient representations such as the bivariate approximation

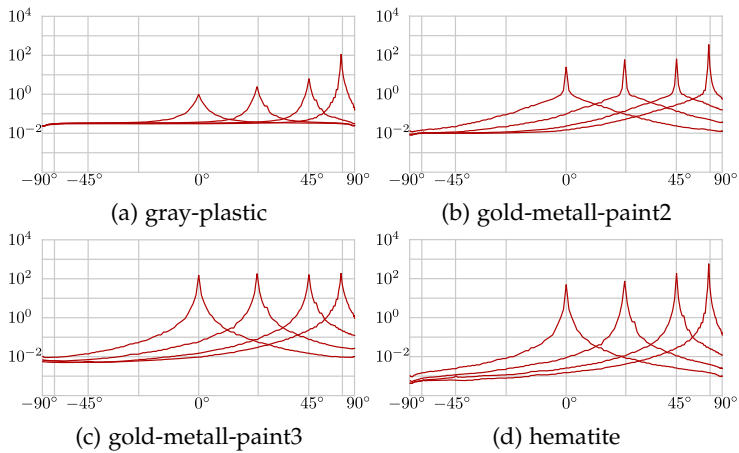


Figure 6.4: Reflection plane plots of BRDF scattering curves for measured data for a set of fixed  $\theta_i = \{0^\circ, 22^\circ, 45^\circ, 67^\circ\}$ . A clear symmetry of the specular lobes around the mirror reflection can be observed.

by Romeiro et al. [183] can help to reduce the data sizes. The measured data can also be approximated by using a finite truncation of some suitable analytical basis expansion, for example spherical harmonics [222]. However, in general these basis expansions require a large number of terms in order to accurately represent high-frequency, glossy, BRDFs. Another approach is to use factorization based methods, such as SVD/PCA or non-linear variants [108, 130, 147]. For efficient importance sampling of measured BRDFs, Lawrence et al. [130] proposed to factorize the data using non-negative matrix factorization methods. This idea was recently extended by Bilgili et al. [25] who used tensor decomposition methods to accurately represent BRDFs that also allows for efficient importance sampling.

## 6.2 Contributions

In this section we discuss the contributions of paper H. The paper propose two new BRDF models for modeling isotropic scattering. In section 6.2.3 we also show how anisotropic scattering can be modeled by extending one of the models presented in the paper.

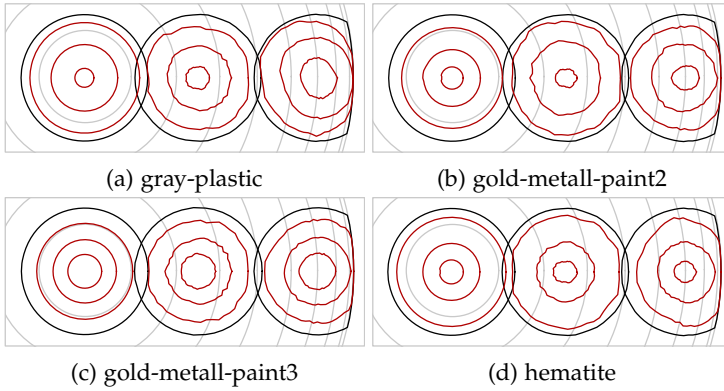


Figure 6.5: Contour line plots in polar coordinates on the unit disk, computed from spherical coordinates using  $(r, \varphi) = (\sin(\theta), \phi)$ , for a set of fixed  $\theta_i = \{0^\circ, 25^\circ, 60^\circ\}$ . A clear symmetry around the reflection direction can be observed. The distance to the contours are also approximately constant for different  $\theta_i$ .

### 6.2.1 Observations from measured data

Paper H studies the behavior of glossy materials from the MERL database [147]. The important feature of these materials is that they exhibiting significant wide angle scattering (broad specular lobes).

By plotting the BRDFs in the reflection plane, containing the incident direction and mirror reflection, using the scaling  $\sin \theta_o$  instead of  $\theta_o$  directly, a clear symmetry around the reflection direction can be observed, see Figure 6.4. Similar behaviours have also been noted and predicted by researches in the optics literature [83, 193]. The scaling of the viewing angle,  $\sin \theta_o$ , simulates the effect of projecting the viewing direction onto the unit disk. From the plots it is also clear that for higher incidence angles (grazing angles,  $\theta_i$  large) there is a clear increase in the specular reflection intensity, consistent with the Fresnel effect. In figure 6.5 display contour line plots of the measured BRDFs projected onto the unit disk, using polar coordinates where the radial coordinate is set as  $r = \sin \theta_o$ . In these plots, it is clear that there is a symmetry around the reflection direction, i.e. the symmetry observed in the reflection plots in figure 6.4 also extends to circular symmetry on the unit disk. In conclusion we observe a clear symmetry of measured scattering curves around the reflection direction when projecting the BRDF data from the hemisphere onto the unit disk.

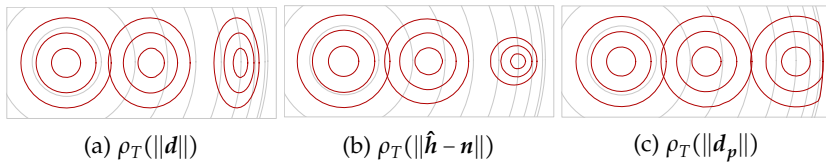


Figure 6.6: Contour line plots on the unit disk for the test function,  $\rho_T(f)$ . a) The Phong lobe does not preserve the circular symmetry observed for measured data. b) Using the Halfway vector preserves the circular shape around the projected reflection direction, but, distorts the radius of the the contour lines when varying  $\theta_i$ . c) The projected deviation vector parametrization exhibits similar properties as observed in the measured data.

### Properties of common model parameterizations

Given the observations from measured data, it is informative to consider the typical behaviors of common BRDF parameterizations used in parametric BRDF models. Do these parameterizations also describe the symmetries observed in figures 6.4 and 6.5? In paper H we set out to investigate this question by considering three different parameterizations commonly used in well-known parametric BRDF models:

1. The difference vector,  $d = \omega_o - \omega_i$ , inspired by its relationship to the specular lobe in the Phong model [175].
2. The difference between the halfway vector (normalized) and the normal,  $d = \hat{h} - \mathbf{n}$ , inspired by its relationship to microfacet based models.
3. The projected difference vector,  $d_p = \mathbf{h} - (\mathbf{h} \cdot \mathbf{n})\mathbf{n}$ , inspired by the relationship to the model of Neumann et al. [163] and the observations of measured data.

To compare the parameterizations we make use of a simple test function,  $\rho_T(f) = \max(1 - f^2, 0)$ , using the length of the vectors as argument, ie  $\rho_T(\|d\|)$ ,  $\rho_T(\|\hat{h} - \mathbf{n}\|)$  and  $\rho_T(\|d_p\|)$ . Figure 6.6 shows contour line plots, in polar coordinates on the unit disk of the test function computed using the three different parameterizations. The projected deviation vector exhibit similar properties as the measured data. On the other hand, the parametrization using the deviation vector does not preserve the circular shape of the reflection lobe for larger  $\theta_i$ , as observed in measured data. The parametrization based on the halfway vector distorts the radius to the contour lines when varying  $\theta_i$ , however this could in theory be compensated for in parametric models using an appropriate scaling term to counteract this effect.



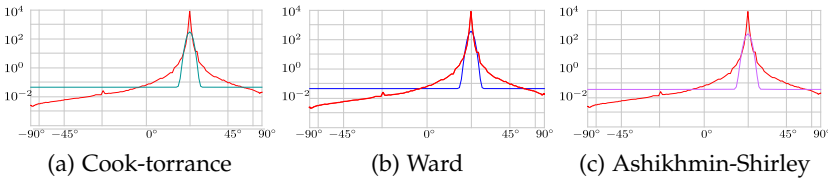


Figure 6.7: Reflection plane plots for three standard parametric BRDF models, Cook-Torrance (green), Ward (blue) and Ashikhmin-Shirley (purple), fitted to measured scattering distribution data for brass (red), for  $\theta_i = 22^\circ$ . Note that all of these models behave very similarly in this case and fails to accurately capture the inverse-power-law shape of the reflectance curves.

### Properties of common parametric distribution (scattering) functions

Another interesting observation from the reflection plane plots of the measured BRDF data for glossy materials shown in figure 6.4 is that the scattering curves do not resemble classical Gaussian or Phong like distribution functions used in standard parametric models. Instead they are much sharper, and have large tails exhibiting an inverse-power-law shape (the tails are mainly responsible for the observed glossy (or wide-angle scattering) behavior of these materials). A simple illustration of this phenomena can be seen in figure 6.7, showing reflection plane plots of standard parametric BRDF models fitted to the measured data. Clearly previous parametric models are not flexible enough to fit the measured data. While one could devise specialized optimization schemes (e.g. changing the fitting metric), the models can be made to either predict the specular peaks or the wide angle scattering reasonable well. However, both of these features cannot in general be captured at the same time using these models (many examples of this phenomena are presented in paper H).

### 6.2.2 New parametric BRDF models for glossy reflectance.

In paper H, we propose two new BRDF models that better reflects the observations of measured data presented in the previous section. One model uses a parametrization based on the halfway vector, inspired by micro-facet theory. The second model uses the projected deviation vector. Instead of relying on a Gaussian shape for describing the shape of the scattering distribution the proposed models use a simplified version of the ABC model [38], defined as

$$S(f) = \frac{A}{(1 + Bf^2)^C} \quad (6.9)$$

where,  $A$ ,  $B$ , and  $C$  are treated as free parameters. Using appropriate settings for the  $A$ ,  $B$  and  $C$  parameters this model allows us to model a variety of scattering

function shapes, better reflecting the observed inverse-power-law behavior of real data. The ABC model is closely related to the non-standardized form of the student-t distribution [93] (obtained from the standard student-t distribution by introducing a scaling of  $x$ , by  $\frac{x}{\sigma}$ ). As is well known, the student-t distribution is symmetric and bell-shaped, like a Gaussian distribution, but in general has heavier tails. For large degrees of freedom (corresponding to small B and large C in the ABC model) it becomes a close approximation to the normal distribution.

The *ABC-smooth surface* model uses the projected deviation vector and is defined as

$$\rho_{ABC-SS} = \frac{K_d}{\pi} + F(\theta_d, \eta) S(\|\mathbf{d}_p\|) \quad (6.10)$$

where  $F(\theta_d, \eta)$  is the fresnel factor proposed by Cook-Torrance [41], with extinction coefficient set to zero. The argument to the fresnel factor is computed as  $\theta_d = \arcsin\left(\frac{\|\omega_i^p - \omega_o^p\|}{2}\right)$ , where  $w_i^p$  and  $w_o^p$  are the surface tangent plane components (projected onto unit disk) of  $\omega_i$  and  $\omega_o$  respectively.  $\eta$  is treated as a free parameter of the model, note that it is generally not directly related to the physical quantity "the index of refraction" as we use the Fresnel factor as an approximation, and we treat  $\eta$  as a free parameter when fitting the model to the data.

The *ABC-microfacet model* is based on a modified version of the Cook-Torrance model [41], using the halfway vector to parameterize the model:

$$\rho_{ABC-M} = \frac{K_d}{\pi} + \frac{S\left(\sqrt{1 - (\hat{\mathbf{h}} \cdot \mathbf{n})}\right) F((\omega_i \cdot \hat{\mathbf{h}}), \eta) G}{(\omega_i \cdot \mathbf{n})(\omega_o \cdot \mathbf{n})} \quad (6.11)$$

where,  $G$ ,  $F$  are the geometrical attenuation and Fresnel factors respectively of Cook and Torrance [41]. Note that the  $A$  parameter is used to scale the specular lobe to fit measured data. Thus the S-factor does not directly represent a normalized facet distribution. Note that the Fresnel term is dependent on the polar angle between the incident direction and the halfway vector (there is an unfortunate typo in paper H where the Fresnel term is written as a function of  $\theta_h$  which is not correct).

**Physical plausibility** Both of the ABC models proposed in paper H was designed to reflect the observations made for measured data, and not modeled explicitly to fulfill energy conservation. However, both models fulfill Helmholtz reciprocity criteria. If strict energy normalization is a necessary requirement, one could use a numerical integration technique for a given set of parameters to find a suitable normalization coefficient. However, note that when fitting

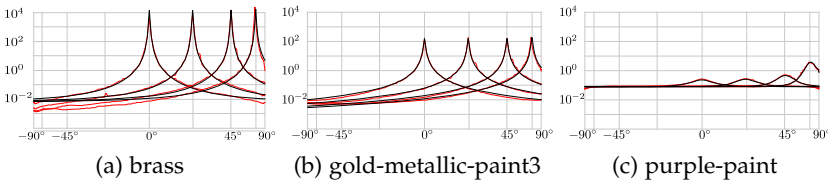


Figure 6.8: Reflection plane plots showing the proposed BRDF models (black) fitted to measured data (red), for a set of fixed  $\theta_i = \{0^\circ, 22^\circ, 45^\circ, 67^\circ\}$ .

the models to measured data, the resulting fitted parameters will roughly correspond to energy conserving BRDF as the measured materials fulfill this property.

**Importance Sampling** In order to use the models in modern image synthesis algorithms using Monte Carlo methods it is vital that the models allow for efficient importance sampling expressions. For both models, paper H describes expressions for sampling procedures for sampling from probability distributions that are roughly proportional to the model distributions. These expressions were derived by applying the inversion rule, and using suitable transformations to derive probability distributions over directions on the hemisphere. As the details are quite involved the reader is therefore referred to the paper for more information on the importance sampling procedures.

**Fitting to measured data and results** To fit the models to measured data we propose two different optimization criteria. The first uses an error metric that puts more emphasis on the central part of the specular lobe, and the second uses a log-mapping to put more emphasis on close fits for wide angle scatter (tails of specular lobes). Using the two error metrics, we fit the ABC models and other previous parametric models to the materials in the MERL database [147]. When comparing the mean fitting errors, it is clear that the new models provide better fits in both error metrics. An example of the fitted ABC parametric models are shown in figure 6.8. Figures 6.9 and 6.11 show example renderings of fitted materials using a HDR light probe as illumination. Comparing the proposed models to previous parametric models, it is clear that the new models provide a more accurate visual appearance of this glossy material. Details on the fitting metrics and many more results and comparisons are given in the paper.

### 6.2.3 Anisotropic model

In this section we show how the smooth surface model, ABC-SS, can be extended to model anisotropic scattering. A direct generalization of the model is given by replacing the metric for computing the length of the projected deviation

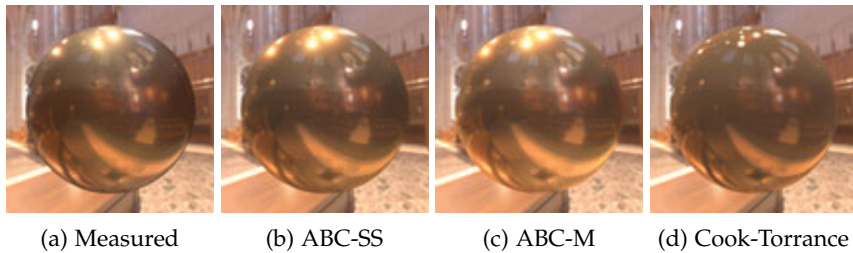


Figure 6.9: Renderings of a sphere made out of the ‘brass’ material. The new ABC models provide much more accurate visual approximations of the measured data.

vector with a generalized one. This general metric can be defined by specifying a quadric. Consider the positive semi-definite matrix,  $Q$ :

$$Q = \begin{bmatrix} B_t^2 & r_{tb}B_tB_b \\ r_{tb}B_tB_b & B_b^2 \end{bmatrix} \quad (6.12)$$

where  $B_t$  and  $B_b$  represents the scalings of the tangent and the bitangent component and  $r_{tb}$  describes the correlation between the scaling of the tangent- and bitangent-axis. Using,  $Q$ , we can define a scalar product and a corresponding metric on the unit disk as:

$$\|\mathbf{d}_p\|_Q = \sqrt{(\mathbf{d}_p^T Q \mathbf{d}_p)} \quad (6.13)$$

The anisotropic ABC-SS model is then given by:

$$\rho_{ABC-SS} = \frac{K_d}{\pi} + F(\theta_d, \eta)S(\|\mathbf{d}_p\|_Q) \quad (6.14)$$

where the  $B$  parameter of the ABC model,  $S(f)$ , can be premultiplied into  $Q$ . Two additional parameters are thus needed to specify the anisotropic model compared to the isotropic one: the ratio of  $B_t$  to  $B_b$  and the correlation of the scalings  $r_{tb}$ . The special case of  $B_t = B_b$  and  $r_{tb} = 1$  represent the standard isotropic form and other choices results in anisotropic scattering. In particular, if  $r_{tb} = 0$  the primary axis of the scaling is along the tangent and bitangent vectors. Figure 6.10 illustrates the effect of scaling the bitangent component of  $\mathbf{d}_p$ , here  $B_t \neq B_b, r_{tb} = 0$ , resulting in anisotropic scattering.

### 6.3 Summary and Future work

In this chapter we provided a summary of previous methods for BRDF representation and capture. We then showed how the new models proposed in

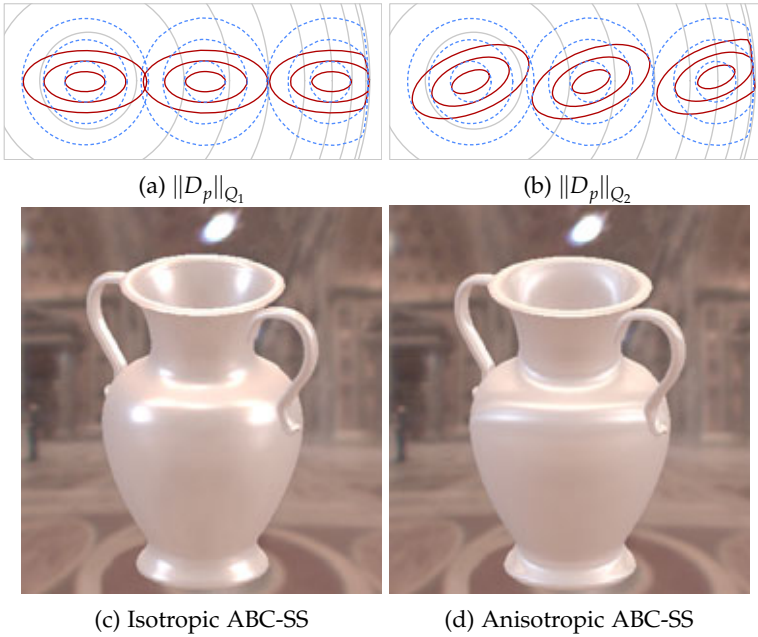


Figure 6.10: The smooth surface model extended model anisotropic scattering. a,b) Contour plots on the unit disk for a test function using the length of the projected deviation vector computed with a standard euclidian metric (blue dashed curves) and two generalized distance metrics (red curves). Using the euclidean metric gives rise to isotropic scattering, while the general metric produces anisotropic scattering. Rendering using c) Isotropic ABC-SS model d) Anisotropic ABC-SS model

paper H enables more accurate renderings of materials exhibiting wide angle scatter, or gloss.

A limitation of the proposed models is that they sometimes fail to predict observed scattering in grazing angles. In ongoing work we are investigating extensions to the proposed ABC-SS model that replaces the Fresnel term with other, more general, models of grazing angle scattering of rough surfaces [193]. We believe that this will enable more accurate modeling of grazing angle effects without introducing an excessive computational burden. Currently we are also investigating data-driven reflectance models based on using the projected deviation vector parameterization used by the smooth-surface model. This parametrization induces a lower dimensional space where the real data can be projected. For isotropic materials, we can also note a symmetry around

the reflection vector, which allows us to represent the measured data in a very compact 2D form instead of the original 3D form. Furthermore, we have observed that, similarly to the ABC-SS model, it is possible to use a separable representation that describes the scattering using two 1D functions, one depending on  $\|d_p\|$  and one on  $\theta_R$ . In contrast to the parametric model proposed here, in the data driven model these functions are represented directly, without fitting them to an explicit parametric model, such as the ABC curve.

We believe that investigations into more efficient parametric BRDF models have come to a point where further developments are hampered by the current lack of large publicly available BRDF databases. The standard database referred to in the vast majority of recent publications on the topic, the MERL database [147], is known to have several flaws. For example, scattering angles larger than approximately 70 degrees are known to be extremely noisy. We therefore believe it is highly important to generate large, open access, BRDF databases that can serve both as a reference in terms of accuracy, and also covers a wide range of common materials. Only then can one be sure to develop models that actually contribute towards the goal of accurately describing real reflectance behaviors, not the noise in the currently available datasets.



(a) Measured



(b) ABC-SS



(c) Ashikhmin-Shirley

Figure 6.11: Renderings using ) the original measured MERL BRDFs with 40000 samples/pixel, a) the new ABC-SS (isotropic) model presented in paper H, rendered using importance sampling at 1000 samples/pixel, b) the Ashikhmin-Shirley (isotropic) model with 1000 samples/pixel.





## Concluding remarks

This chapter provides a summary of the the first part of the thesis and discusses possible venues for future work. More detailed discussions of future work can also be found in the selected papers constituting the second part of the thesis.

### 7.1 Summary of contributions

In this thesis, we have discussed methods for accurate and efficient rendering of synthetic objects so that they can be seamlessly composited into real environments. In particular, four specific challenges have been identified and adressed in this thesis:

1. To enable robust and accurate HDR image and video capture we presented a new noise aware HDR reconstruction framework that unifies standard processing operations such as demosiacing, sensor realignment, denoising, and HDR fusion into a single filtering operation.
2. We showed how HDR video capture can be used to improve the generality and quality of traditional IBL methods, by capturing and rendering with temporally and spatially varying illumination conditions.
3. For realistic rendering of glossy surfaces we presented two new parametric BRDF models that provide a better fit to measured data than previous models.
4. Finally, we presented a generalization of previous rendering algorithms based on Markov Chain Monte Carlo methods that allow non-negative and

unbiased estimators to be used in place of exact light transport models, while still converging to the exact result. We showed that this not only extends the applicability of these methods, but can also improve their efficiency, for example in scenes with heterogenous participating media.

We believe that with the continued development of HDR video cameras, practitioners in visual effects, product visualization, and other related areas will benefit from using methods for temporally and spatially varying IBL similar to those presented in this thesis. These methods will enable rapid artistic workflows for rendering synthetic objects in real environments, providing accurate results on the first go without the need for tedious modeling and adjustments of the illumination in the real scene. We also believe that more accurate BRDF models, similar to those presented in this thesis, will enable artists at production environments to transition from complex hand modeled shaders to physically plausible models that lead to more predictable results. With the continued development of more accurate models for glossy materials and participating media, such as skin, fabrics, and other complex materials, we also believe that there will be an increasing need for efficient rendering algorithms suitable for such complex scenes. Rendering algorithms based on Markov Chain Monte Carlo methods are likely to play a key role in such developments as they enable a global view of the problem without restricting themselves to localized importance sampling methods.

## 7.2 Future work

There are many interesting areas for further development of the methods presented in this thesis.

In the area of HDR video, there is still much room for improvement. While spatially multiplexed imaging systems based on a single sensor are promising, they provide a tradeoff between spatial resolution and dynamic range. Simultaneous capture of high resolution and high dynamic range video is thus a challenging problem. While multi-sensor systems currently provide a better alternative for capturing scenes with a very high dynamic range, for example for IBL in outdoor scenes, they can be bulky. The development of combined approaches are therefore promising. Other specialized sensors can also be combined with multi-sensor setups, for example sensors with a logarithmic response to incident light [168]. In terms of HDR reconstruction methods, we believe that it would be interesting to develop methods based on exploiting the similarity of patches in the image. This is similar to modern denoising algorithms, such as *non-local mean* variants [32, 230].

Regarding IBL, we believe that a natural development is to use recent methods developed in computer vision for both more automatic and more accurate scene reconstructions, for example based on RGB-D video [37]. This will allow better modeling of the interactions between real and virtual objects, leading to more realistic results.

Rapid acquisition of reflectance data from the real world is another interesting topic for future work. Given the observed symmetries in real BRDF data that were discussed in this thesis, new rapid measurement techniques that exploit the lower dimension of the data could be developed. It would also be interesting to develop methods for editing measured BRDF data using parametric models, using a technique similar to differential rendering.

An interesting venue for future work in the area of efficient rendering algorithms is to investigate the use of Sequential Monte Carlo methods [55] for rendering of, for example, participating media. These methods are closely related to sequential importance sampling methods such as path tracing, but introduce a *resampling* step that can be seen as a generalization of the probabilistic termination criteria in path tracing known as *Russian Roulette*. Recently the combination of Sequential Monte Carlo methods with pseudo-marginal MCMC methods have shown to highly efficient for applications in computational statistics [12]. We believe that investigating such combinations for use in rendering is a promising venue for future work.

In the future, with the continued development of methods for photorealistic rendering of synthetic objects in real scenes, it is likely that it will be very difficult to distinguish the virtual from the real. While we believe that this technology will mostly be used with a good intent, performed for fun or for artistic value, we also note that it is likely that some image manipulations will be used for adversarial purposes, such as propaganda or misinformation campaigns. An interesting challenge in this context is that of image forensics, with the goal of understanding if an image has been manipulated or not. While recently proposed methods for this purpose are based on detecting inconsistencies in the lighting from shading and shadow constraints in the image [109], future image manipulation methods will likely be able to pass these tests.



# Bibliography

- [1] Spheron VR. [Online]. Available: <https://www.spheron.com/home.html> [page 89]
- [2] A standard default color space for the internet - sRGB. [Online]. Available: <http://www.w3.org/Graphics/Color/sRGB.html> [page 43]
- [3] alex. Dynamic range improvement for canon dslrs with 8-channel sensor readout by alternating iso during sensor readout. Technical documentation, url: [http://acoutts.com/alex/dual\\_iso.pdf](http://acoutts.com/alex/dual_iso.pdf), 2013. [pages 77 and 91]
- [4] M. Aggarwal and N. Ahuja. Split aperture imaging for high dynamic range. *International Journal of Computer Vision*, 58(1):7–17, 2004. [page 74]
- [5] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Musé. Study of the digital camera acquisition process and statistical modeling of the sensor raw data. 2012. [page 72]
- [6] C. Aguerrebere, A. Almansa, Y. Gousseau, J. Delon, and P. Musé. Single shot high dynamic range imaging using piecewise linear estimators. In *ICCP*, 2014. [page 77]
- [7] C. Aguerrebere, J. Delon, Y. Gousseau, and P. Musé. Best algorithms for hdr image generation. a study of performance bounds. *SIAM Journal on Imaging Sciences*, 7(1):1–34, 2014. [pages 73 and 80]
- [8] B. Ajdin, M. B. Hullin, C. Fuchs, H.-P. Seidel, and H. P. A. Lensch. Demosaicing by smoothing along 1-d features. In *CVPR*, 2008. [pages 77 and 78]
- [9] T. Akenine-Möller, E. Haines, and N. Hoffman. *Real-Time Rendering 3rd Edition*. A. K. Peters, Ltd., Natick, MA, USA, 2008. [page 28]
- [10] P. Andreo. Monte carlo techniques in medical radiation physics. *Physics in medicine and biology*, 36(7):861, 1991. [pages 23 and 24]

- [11] C. Andrieu and G. O. Roberts. The pseudo-marginal approach for efficient monte carlo computations. *The Annals of Statistics*, pages 697–725, 2009. [pages 48 and 49]
- [12] C. Andrieu, A. Doucet, and R. Holenstein. Particle markov chain monte carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 72(3):269–342, 2010. [pages 48 and 115]
- [13] A. Appel. Some techniques for shading machine renderings of solids. In *Proceedings of the April 30–May 2, 1968, Spring Joint Computer Conference, AFIPS '68 (Spring)*, pages 37–45, New York, NY, USA, 1968. ACM. [page 3]
- [14] J. Arvo. Transfer equations in global illumination. In *Siggraph*, 1993. [page 25]
- [15] M. Ashikhmin and P. Shirley. An anisotropic phong brdf model. *Journal of Graphics Tools*, 5(2):25–32, 2000. [page 4]
- [16] M. Ashikhmin and P. Shirley. An anisotropic phong brdf model. *Journal of Graphics Tools*, 5(2):25–32, 2000. [page 100]
- [17] M. Ashikhmin, S. Premože, and P. Shirley. A microfacet-based brdf generator. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 65–74. ACM Press/Addison-Wesley Publishing Co., 2000. [page 101]
- [18] J. Astola, V. Katkovnik, and K. Egiazarian. *Local Approximation Techniques in Signal and Image Processing*. SPIE Publications, 2006. [pages 79, 82, 84, and 87]
- [19] M. M. Bagher, C. Soler, and N. Holzschuch. Accurate fitting of measured reflectances using a shifted gamma micro-facet distribution. *Computer Graphics Forum*, 31(4):1509–1518, 2012. [page 101]
- [20] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers. *Advanced high dynamic range imaging: theory and practice*. CRC Press, 2011. [page 59]
- [21] F. Banterle, M. Callieri, M. Dellepiane, M. Corsini, F. Pellacini, and R. Scopigno. EnvyDepth: An Interface for Recovering Local Natural Illumination from Environment Maps. *Computer Graphics Forum (Proc. of Pacific Graphics 2013)*, 32(7):411–420, 2013. [page 63]
- [22] P. Barla, L. Belcour, and R. Pacanowski. In praise of an alternative brdf parametrization. In *Workshop on Material Appearance Modeling*, 2015. [page 97]

- [23] B. Bayer. Color imaging array. US Patent 3 971 065, 1976. [page 71]
- [24] M. A. Beaumont. Estimation of population growth or decline in genetically monitored populations. *Genetics*, 164(3):1139–1160, 2003. [page 49]
- [25] A. Bilgili, A. Öztürk, and M. Kurt. A general brdf representation based on tensor decomposition. *Computer Graphics Forum*, 30(8):2427–2439, 2011. [page 102]
- [26] J. Blinn and M. Newell. Texture and reflection in computer generated images. *Communications of the ACM*, 19(10):542–547, 1976. [page 58]
- [27] J. F. Blinn. Models of light reflection for computer synthesized pictures. In *SIGGRAPH '77*, pages 192–198. ACM, 1977. [pages 4 and 100]
- [28] C. Bloch. *The HDRI Handbook 2.0*. Rocky Nook, 2012. [pages 5, 12, 59, 61, and 64]
- [29] M. Born and E. Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Cambridge university press, 1999. [page 95]
- [30] W. J. Bouknight. A procedure for generation of three-dimensional half-toned computer graphics presentations. *Commun. ACM*, 13(9):527–536, 1970. [page 3]
- [31] A. Brady, J. Lawrence, P. Peers, and W. Weimer. genbrdf: discovering new analytic brdfs with genetic programming. *ACM Transactions on Graphics (TOG)*, 33(4):114, 2014. [page 100]
- [32] A. Buades, B. Coll, and J.-M. Morel. A non-local algorithm for image denoising. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 60–65. IEEE, 2005. [page 114]
- [33] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation (SIAM interdisciplinary journal)*, 4(2):490–530, 2005. [page 77]
- [34] B. Cabral, N. Max, and R. Springmeyer. Bidirectional reflection functions from surface bump maps. In *Siggraph*, 1987. [page 101]
- [35] D. A. Calian, K. Mitchell, D. Nowrouzezahrai, and J. Kautz. The shading probe: fast appearance acquisition for mobile AR. In *SIGGRAPH Asia Technical Briefs*, 2013. [page 59]
- [36] S. Chandrasekhar. *Radiative transfer*. Dover, 1960. [page 24]

- [37] S. Choi, Q.-Y. Zhou, and V. Koltun. Robust reconstruction of indoor scenes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5556–5565, 2015. [page 115]
- [38] E. L. Church, P. Z. Takacs, and T. A. Leonard. The prediction of brdfs from surface profile measurements. In *Proceedings of SPIE*, volume 1165, pages 136–150, 1989. [pages 101 and 105]
- [39] D. Cline, J. Talbot, and P. Egbert. Energy redistribution path tracing. *ACM TOG (SIGGRAPH)*, 24(3):1186–1195, 2005. [page 47]
- [40] R. L. Cook. Stochastic sampling in computer graphics. *ACM Trans. Graph.*, 5(1):51–72, 1986. [page 38]
- [41] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics*, 1(1):7–24, 1982. [pages 101 and 106]
- [42] M. Corsini, M. Callieri, and P. Cignoni. Stereo Light Probe. *Computer Graphics Forum (Proc. of Eurographics 2008)*, 27(2):291–300, 2008. [page 63]
- [43] C. Dachsbacher, J. Křivánek, M. Hašan, A. Arbree, B. Walter, and J. Novák. Scalable realistic rendering with many-light methods. *Computer Graphics Forum*, 33(1):88–104, 2014. [page 41]
- [44] J. Dahlin, F. Lindsten, and T. B. Schön. Particle metropolis–hastings using gradient and hessian information. *Statistics and computing*, 25(1):81–92, 2014. [page 48]
- [45] K. J. Dana and J. Wang. Device for convenient measurement of spatially varying bidirectional reflectance. *JOSA A*, 21(1):1–12, 2004. [page 98]
- [46] P. Debevec. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH*, pages 1–10, 1998. [pages 4, 5, 57, and 59]
- [47] P. Debevec. Making the parthenon. In *International Symposium on Virtual Reality, Archaeology, and Cultural Heritage*, 2005. [page 64]
- [48] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proc. of ACM SIGGRAPH*, pages 369–378. ACM, ACM, 1997. [pages 5, 71, 73, 77, and 78]
- [49] P. Debevec, P. Graham, J. Busch, and M. Bolas. A single-shot light probe. In *SIGGRAPH Talks*, pages 10:1–10:1, 2012. [page 59]



- [50] P. E. Debevec, C. J. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 11–20. ACM, 1996. [page 4]
- [51] S. DiVerdi, J. Wither, and T. Hollerei. Envisor: Online environment map construction for mixed reality. *IEEE Virtual Reality*, pages 19–26, 2008. [page 59]
- [52] Y. Dong, S. Lin, and B. Guo. *Material Appearance Modeling: A Data-Coherent Approach*. Springer, 2013. [page 93]
- [53] J. Dongarra and F. Sullivan. Guest editors' introduction: The top 10 algorithms. *Computing in Science and Engg.*, 2(1):22–23, 2000. [page 29]
- [54] J. Dorsey, H. Rushmeier, and F. Sillion. *Digital modeling of material appearance*. Morgan Kaufmann, 2010. [page 93]
- [55] A. Doucet and A. M. Johansen. A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12:656–704, 2009. [page 115]
- [56] J. Dupuy, E. Heitz, J.-C. Iehl, P. Poulin, and V. Ostromoukhov. Extracting microfacet-based brdf parameters from arbitrary materials with power iterations. *Computer Graphics Forum*, 34(4):21–30, 2015. [page 101]
- [57] A. Dür. An improved normalization for the ward reflectance model. *Journal of graphics, gpu, and game tools*, 11(1):51–59, 2006. [page 100]
- [58] R. Durrett. *Probability: theory and examples*. Cambridge university press, 2010. [pages 30 and 35]
- [59] P. Dutre, P. Bekaert, and K. Bala. *Advanced global illumination*. CRC Press, 2006. [page 17]
- [60] R. Eckhardt. Stan ulam, john von neumann, and the monte carlo method. *Los Alamos Science*, 15:131–136, 1987. [page 30]
- [61] A. Einstein. English translation of the 1905 paper : "concerning an heuristic point of view toward the emission and transformation of light". *American Journal of Physics*, 33(5):367, 1965. [page 70]
- [62] K. Engel, M. Hadwiger, J. Kniss, C. Rezk-Salama, and D. Weiskopf. *Real-time volume graphics*. CRC Press, 2006. [page 25]

## 122 Bibliography

- [63] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg. A model of visual adaptation for realistic image synthesis. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '96*, pages 249–258, New York, NY, USA, 1996. ACM. [page 72]
- [64] A. Fournier, A. Gunawan, and C. Romanzin. Common Illumination between Real and Computer Generated Scenes. In *Graphics Interface*, pages 254–254, 1993. [page 57]
- [65] J. Froehlich, S. Grandinetti, B. Eberhardt, S. Walter, A. Schilling, and H. Brendel. Creating cinematic wide gamut hdr-video for the evaluation of tone mapping operators and hdr-displays. In *IS&T/SPIE Electronic Imaging*, pages 90230X–90230X. International Society for Optics and Photonics, 2014. [page 74]
- [66] M. Fuchs, V. Blanz, H. Lensch, and H.-P. Seidel. Adaptive sampling of reflectance fields. *ACM Transactions on Graphics (TOG)*, 26(2):10, 2007. [page 98]
- [67] Y. Furukawa, B. Curless, S. M. Seitz, and R. Szeliski. Reconstructing building interiors from images. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 80–87, 2009. [page 64]
- [68] R. Gal, Y. Wexler, E. Ofek, H. Hoppe, and D. Cohen-Or. Seamless montage for texturing models. *Computer Graphics Forum*, 29(2):479–486, 2010. [page 66]
- [69] O. Gallo, M. Tico, R. Manduchi, N. Gelfand, and K. Pulli. Metering for exposure stacks. *Computer Graphics Forum*, 31(2):479–488, 2012. [page 73]
- [70] A. Ghosh, S. Achutha, W. Heidrich, and M. O’Toole. Brdf acquisition with basis illumination. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007. [page 98]
- [71] A. Goldenshluger and A. Nemirovski. On spatially adaptive estimation of nonparametric regression. *Mathematical methods of Statistics*, 6(2):135–170, 1997. [page 86]
- [72] C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Battaile. Modeling the interaction of light between diffuse surfaces. *SIGGRAPH Comput. Graph.*, 18(3):213–222, 1984. [pages 3 and 28]
- [73] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H. Seidel, and H. Lensch. Optimal hdr reconstruction with linear digital cameras. In *CVPR*, 2010. [pages 5, 73, and 77]

- [74] N. Greene. Environment mapping and other applications of world projections. *Computer Graphics and Applications*, 6(11):21–29, 1986. [page 58]
- [75] M. D. Grossberg and S. K. Nayar. Modeling the space of camera response functions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(10):1272–1282, 2004. [page 71]
- [76] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau. Demosaicking: color filter array interpolation. *IEEE Signal Processing Magazine*, 22(1):44–54, 2005. [pages 71 and 77]
- [77] T. Hachisuka, S. Ogaki, and H. W. Jensen. Progressive photon mapping. *ACM Transactions on Graphics (TOG)*, 27(5):130, 2008. [page 41]
- [78] T. Hachisuka, J. Pantaleoni, and H. W. Jensen. A path space extension for robust light transport simulation. *ACM Trans. Graph.*, 31(6):191:1–191:10, 2012. [page 39]
- [79] S. Hajisharif, J. Kronander, E. Miandji, and J. Unger. Real-time image based lighting with streaming HDR-light probe sequences. In *SIGRAD*, 2012. [page 62]
- [80] S. Hajisharif, J. Kronander, and J. Unger. HDR reconstruction for alternating gain (ISO) sensor readout. In *Eurographics Short Papers*, 2014. [pages 14 and 77]
- [81] S. Hajisharif, J. Kronander, and J. Unger. Adaptive dualISO HDR-reconstruction. *Submitted to EURASIP Journal on Image and Video Processing*, 2015.
- [82] J. Hanika, A. Kaplanyan, and C. Dachsbacher. Improved half vector space light transport. *Computer Graphics Forum (Proceedings of Eurographics Symposium on Rendering)*, 34(4):to appear, 2015. [page 45]
- [83] J. E. Harvey. Light scattering characteristics of optical surfaces. In *Proceedings of SPIE*, volume 107, pages 41–47, 1977. [page 103]
- [84] S. Hasinoff, F. Durand, and W. Freeman. Noise-optimal capture for high dynamic range photography. In *CVPR*, 2010. [pages 73 and 79]
- [85] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. [page 37]
- [86] V. Havran, M. Smyk, G. Krawczyk, K. Myszkowski, and H.-P. Seidel. Interactive system for dynamic scene lighting using captured video environment maps. In *Rendering Techniques (Proc. of EGSR)*, pages 31–42, 2005. [pages 6 and 60]

## 124 Bibliography

- [87] E. Hayward and J. Hubbell. The albedo of various materials for 1-mev photons. *Physical Review*, 93(5):955, 1954. [page 29]
- [88] X. D. He, K. E. Torrance, F. X. Sillion, and D. P. Greenberg. A comprehensive physical model for light reflection. *SIGGRAPH '91*, 25(4):175–186, 1991. [page 101]
- [89] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. Pająk, D. Reddy, O. Gallo, J. L. abd Wolfgang Heidrich, K. Egiazarian, J. Kautz, and K. Pulli. Flexisp: A flexible camera image processing framework. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia)*, 33(6), 2014. [page 78]
- [90] E. Heitz. Understanding the masking-shadowing function in microfacet-based brdfs. *Journal of Computer Graphics Techniques (JCGT)*, 3(2):48–107, 2014. [page 101]
- [91] J. Hoberock and J. C. Hart. Arbitrary importance functions for metropolis light transport. *Computer Graphics Forum*, 29(6):1993–2003, 2010. [page 43]
- [92] I. Ihrke, T. Stich, H. Gottschlich, M. Magnor, and H.-P. Seidel. Fast incident light field acquisition and rendering. *Journal of WSCG*, 16(1-3):25–32, 2008. [page 63]
- [93] S. Jackman. *Bayesian analysis for the social sciences*, volume 846. John Wiley & Sons, 2009. [page 106]
- [94] W. Jakob. Mitsuba renderer, 2010, <http://www.mitsuba-renderer.org>. [page 46]
- [95] W. Jakob. *Light transport on path-space manifolds*. PhD thesis, Cornell University, 2013. [pages 25, 26, and 28]
- [96] W. Jakob and S. Marschner. Manifold exploration: A markov chain monte carlo technique for rendering scenes with difficult specular transport. *ACM Trans. Graph.*, 31(4):58:1–58:13, 2012. [page 45]
- [97] W. Jarosz, C. Donner, M. Zwicker, and H. W. Jensen. Radiance caching for participating media. *ACM Transactions on Graphics (TOG)*, 27(1):7, 2008. [page 41]
- [98] H. W. Jensen. *Realistic image synthesis using photon mapping*. AK Peters, Ltd., 2001. [page 41]
- [99] N. Johansson. *Measuring and modelling light scattering in paper*. PhD thesis, Mid Sweden University, 2015. [page 98]

- [100] D. Jönsson, J. Kronander, T. Ropinski, and A. Ynnerman. Historygrams: Enabling interactive global illumination in direct volume rendering using photon mapping. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2364–2371, 2012.
- [101] J. T. Kajiya. The rendering equation. *Computer Graphics (Proc. of SIGGRAPH)*, 20(4):143–150, 1986. [pages 3, 21, and 38]
- [102] M. H. Kalos and P. A. Whitlock. *Monte carlo methods*. John Wiley & Sons, 2008. [page 40]
- [103] S. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High dynamic range video. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2003)*, 22(3):319–325, 2003. [page 73]
- [104] A. S. Kaplanyan, J. Hanika, and C. Dachsbacher. The natural-constraint representation of the path space for efficient light transport simulation. *ACM Transactions on Graphics (Proc. SIGGRAPH)*, 33(4), 2014. [page 45]
- [105] K. Karsch and D. Forsyth. Rendering Synthetic Objects into Legacy Photographs. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia 2011)*, 30(6):157:1–157:12, 2011. [page 58]
- [106] K. Karsch, K. Sunkavalli, S. Hadap, N. Carr, H. Jin, R. Fonte, M. Sittig, and D. Forsyth. Automatic scene inference for 3d object compositing. *ACM Transactions on Graphics*, 33(3), 2014. [page 58]
- [107] V. Katkovnik. A new method for varying adaptive bandwidth selection. *IEEE Transactions on signal processing*, 47(9):2567–2571, 1999. [page 86]
- [108] J. Kautz and M. D. McCool. Interactive rendering with arbitrary brdfs using separable approximations. In *Rendering Techniques' 99*, pages 247–260. Springer, 1999. [page 102]
- [109] E. Kee, J. F. O'brien, and H. Farid. Exposing photo manipulation from shading and shadows. *ACM Trans. Graph.*, 33(5):165:1–165:21, 2014. [page 115]
- [110] C. Kelemen, L. Szirmay-Kalos, G. Antal, and F. Csonka. A simple and robust mutation strategy for the metropolis light transport algorithm. *Computer Graphics Forum*, 21(3):531–540, 2002. [pages 45 and 47]
- [111] A. Keller. Instant radiosity. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '97*, pages 49–56, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co. [page 41]

- [112] E. A. Khan, E. Reinhard, R. W. Fleming, and H. H. Bühlhoff. Image-based material editing. *ACM Transactions on Graphics (Proc. of Siggraph 2006)*, 25(3), 2006. [page 58]
- [113] H. Knutsson and C.-F. Westin. Normalized and differential convolution. In *CVPR*, 1993. [pages 13 and 82]
- [114] J. Kronander and T. B. Schön. Robust auxiliary particle filters using multiple importance sampling. In *IEEE Statistical Signal Processing Workshop (SSP)*, 2014.
- [115] J. Kronander, J. Unger, T. Möller, and A. Ynnerman. Estimation and modeling of actual numerical errors in volume rendering. *Computer Graphics Forum (Proc. of Eurovis)*, 29(3):893–902, 2010.
- [116] J. Kronander, D. Jönsson, J. Low, P. Ljung, A. Ynnerman, and J. Unger. Efficient visibility encoding for dynamic illumination in direct volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 18(3):447–462, 2012.
- [117] J. Kronander, S. Gustavson, G. Bonnet, and J. Unger. Unified HDR reconstruction from raw CFA data. In *IEEE International Conference on Computational Photography (ICCP)*, 2013.
- [118] J. Kronander, J. Dahlin, D. Jönsson, M. Kok, T. B. Schön, and J. Unger. Real-time video based lighting using HDR video and Sequential Monte Carlo samplers. In *Proceedings of EUSIPCO'14: Special Session on HDR-video*, 2014. [page 62]
- [119] J. Kronander, S. Gustavson, G. Bonnet, A. Ynnerman, and J. Unger. A unified framework for multi-sensor HDR video reconstruction. *Signal Processing: Image Communication*, 29(2), 2014.
- [120] J. Kronander, T. B. Schön, and J. Dahlin. Backward Sequential Monte Carlo for marginal smoothing. In *IEEE Statistical Signal Processing Workshop (SSP)*, 2014.
- [121] J. Kronander, F. Banterle, A. Gardner, E. Miandji, and J. Unger. Photorealistic rendering of mixed reality scenes. *Computer Graphics Forum (Proc. of Eurographics STARs)*, 34(2):643–665, 2015.
- [122] J. Kronander, T. B. Schön, and J. Unger. Pseudo-Marginal Metropolis Light Transport. *SIGGRAPH Asia Technical Briefs*, Preprint version, 2015.
- [123] J. Kronander, T. B. Schön, and J. Unger. Pseudo-Marginal Metropolis Light Transport. In *SIGGRAPH Asia Technical Briefs*, 2015.

- [124] M. Kurt and D. Edwards. A survey of brdf models for computer graphics. *SIGGRAPH Comput. Graph.*, 43(2):4:1–4:7, 2009. [page 93]
- [125] E. P. Lafortune and Y. D. Willems. Bi-directional path tracing. In *Proceedings of CompuGraphics*, volume 93, pages 145–153, 1993. [page 41]
- [126] E. P. Lafortune and Y. D. Willems. Using the modified phong reflectance model for physically based rendering. Technical report, K.U. Leuven, 1994. [page 100]
- [127] J.-F. Lalonde, A. a. Efros, and S. G. Narasimhan. Estimating the Natural Illumination Conditions from a Single Outdoor Image. *International Journal of Computer Vision*, 98(2):123–145, 2012. [page 58]
- [128] P. Lancaster and K. Salkauskas. Surfaces generated by moving least squares methods. *Mathematics of Computation*, 87:141–158, 1981. [page 82]
- [129] D. Landau and K. Binder. *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, New York, NY, USA, 2005. [page 23]
- [130] J. Lawrence, S. Rusinkiewicz, and R. Ramamoorthi. Efficient brdf importance sampling using a factored representation. In *ACM Transactions on Graphics (Proceedings of SIGGRAPH '04)*, pages 496–505. ACM, 2004. [page 102]
- [131] V. Lempitsky and D. Ivanov. Seamless mosaicing of image-based texture maps. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1 –6, 2007. [page 66]
- [132] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, J. Ginsberg, J. Shade, and D. Fulk. The digital michelangelo project: 3d scanning of large statues. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00*, pages 131–144, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co. [page 4]
- [133] R. R. Lewis. Making shaders more physically plausible. *Computer Graphics Forum*, 13(2):109–120, 1993. [pages 95 and 100]
- [134] S. Lindholm and J. Kronander. Accounting for uncertainty in medical data: A cuda implementation of normalized convolution. In *SIGRAD*, 2011.
- [135] F. Lindsten, M. I. Jordan, and T. B. Schön. Particle gibbs with ancestor sampling. *Journal of Machine Learning Research*, 15:2145–2184, 2014. [page 48]

- [136] J. S. Liu. *Monte Carlo strategies in scientific computing*. Springer Science & Business Media, 2008. [pages 37 and 43]
- [137] C. Loader. *Local regression and likelihood*. New York: Springer-Verlag, 1999. [page 82]
- [138] J. Löw, A. Ynnerman, P. Larsson, and J. Unger. HDR Light Probe Sequence Resampling for Realtime Incident Light Field Rendering. In *Spring Conference on Computer Graphics*, pages 43–50, 2009. [page 63]
- [139] J. Löw, J. Kronander, A. Ynnerman, and J. Unger. BRDF models for accurate and efficient rendering of glossy surfaces. *ACM Transactions on Graphics (TOG)*, 31(1):9, 2012.
- [140] R. Lukac. *Single-sensor imaging: methods and applications for digital cameras*. CRC Press, 2008. [page 72]
- [141] S. N. MacEachern, M. Clyde, and J. S. Liu. Sequential importance sampling for nonparametric bayes models: The next generation. *Canadian Journal of Statistics*, 27(2):251–267, 1999. [page 47]
- [142] A. Manakov, J. F. Restrepo, O. Klehm, R. Hegedüs, E. Eisemann, H.-P. Seidel, and I. Ihrke. A reconfigurable camera add-on for high dynamic range, multi-spectral, polarization, and light-field imaging. *ACM Trans. Graph. (Proc. SIGGRAPH 2013)*, 32(4):47:1–47:14, 2013. [pages 76 and 91]
- [143] S. Mann and R. W. Picard. On being ‘undigital’ with digital cameras: Extending dynamic range by combining differently exposed pictures. In *IS&T*, 1995. [pages 5 and 73]
- [144] R. K. Mantiuk, K. Myskowski, and H.-P. Seidel. High dynamic range imaging. In *Wiley Encyclopedia of Electrical and Electronics Engineering*. Wiley, 2016. [page 72]
- [145] S. R. Marschner, S. H. Westin, E. P. F. Lafortune, and K. E. Torrance. Image-based bidirectional reflectance distribution function measurement. *Appl. Opt.*, 39(16):2592–2600, 2000. [page 98]
- [146] V. Masselus, P. Peers, P. Dutre, and Y. D. Willems. Relighting with 4d incident light fields. *ACM Transactions on Graphics (Proc. of SIGGRAPH 2003)*, 22(3):613–620, 2003. [page 63]
- [147] W. Matusik, H. Pfister, M. Brand, and L. Mcmillan. A data-driven reflectance model. *ACM Transactions on Graphics*, 22:759–769, 2003. [pages 99, 101, 102, 103, 107, and 110]



- [148] M. McGuire and J. F. Hughes. Optical Splitting Trees for High-Precision Monocular Imaging. *IEEE Computer Graphics And Applications*, 27:32–42, 2007. [page 74]
- [149] N. Metropolis. The beginning of the monte carlo method. *Los Alamos Science*, 15(584):125–130, 1987. [page 30]
- [150] N. Metropolis and S. Ulam. The monte carlo method. *Journal of the American statistical association*, 44(247):335–341, 1949. [page 29]
- [151] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953. [page 37]
- [152] E. Miandji, J. Kronander, and J. Unger. Geometry independent surface light fields for real time rendering of precomputed global illumination. In *SIGRAD*, 2011.
- [153] E. Miandji, J. Kronander, and J. Unger. Learning based compression of surface light fields for real-time rendering of global illumination scenes. In *SIGGRAPH Asia Technical Briefs*, 2013.
- [154] E. Miandji, J. Kronander, and J. Unger. Compressive image reconstruction in reduced union of subspaces. *Computer Graphics Forum (Proc. of Eurographics)*, 34(2), 2015.
- [155] P. Milanfar. A tour of modern image filtering: New insights and methods, both practical and theoretical. *IEEE Signal Processing Magazine*, 30(1):106 – 128, 2013. [page 82]
- [156] G. Miller and C. R. Hoffman. Illumination and reflection maps: Simulated objects in simulated and real environments. In *Proc. of SIGGRAPH 86*, 1986. [page 58]
- [157] T. Mitsunaga and S. K. Nayar. Radiometric self calibration. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 1. IEEE, 1999. [page 73]
- [158] J. Murray-Coleman and A. Smith. The automated measurement of brdfs and their application to luminaire modeling. *Journal of the Illuminating Engineering Society*, 19(1):87–99, 1990. [page 98]
- [159] A. a. Mury, S. C. Pont, and J. J. Koenderink. Representing the light field in finite three-dimensional spaces from sparse discrete samples. *Applied optics*, 48(3):450–7, 2009. [page 63]

- [160] P. Musialski, P. Wonka, D. G. Aliaga, M. Wimmer, L. Gool, and W. Purgathofer. A survey of urban reconstruction. *Computer graphics forum*, 32(6):146–177, 2013. [page 64]
- [161] S. G. Narasimhan and S. K. Nayar. Enhancing Resolution Along Multiple Imaging Dimensions Using Assorted Pixels. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4):518–530, 2005. [page 77]
- [162] S. Nayar and T. Mitsunaga. High dynamic range imaging: Spatially varying pixel exposures. In *CVPR*, 2000. [page 76]
- [163] L. Neumann, A. Neumann, and L. Szirmay-Kalos. Reflectance models with fast importance sampling. *Computer Graphics Forum*, 18:249–265, 1999. [pages 97 and 104]
- [164] A. Ngan, F. Durand, and W. Matusik. Experimental analysis of brdf models. In *Proceedings of the Eurographics Symposium on Rendering*, pages 117–226. Eurographics Association, 2005. [page 101]
- [165] F. E. Nicodemus, J. C. Richmond, J. J. Hsia, I. W. Ginsber, and T. Limperis. *Geometrical Considerations and Nomenclature for Reflectance*. U.S. Dept. of Commerce, National Bureau of Standards, NBS MN-160, 1977. [page 94]
- [166] J. B. Nielsen, H. W. Jensen, and R. Ramamoorthi. On optimal, minimal brdf sampling for reflectance acquisition. *ACM Transactions on Graphics (Proc. of SIGGRAPH Asia)*, 2015. [page 98]
- [167] J. Novák, A. Selle, and W. Jarosz. Residual ratio tracking for estimating attenuation in participating media. *ACM TOG (SIGGRAPH Asia)*, 33(6):179, 2014. [pages 51 and 52]
- [168] J. Ohta. *Smart CMOS Image Sensors*. CRC press, 2008. [pages 71 and 114]
- [169] M. Oren and S. K. Nayar. Generalization of lambert’s reflectance model. In *SIGGRAPH ’94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 239–246, New York, NY, USA, 1994. ACM. [page 101]
- [170] A. Owen and Y. Zhou. Safe and effective importance sampling. *Journal of the American Statistical Association*, 95(449):135–143, 2000. [page 33]
- [171] A. B. Owen. *Monte Carlo theory, methods and examples*. 2013. [pages 30 and 33]
- [172] M. Pauly, T. Kollig, and A. Keller. Metropolis light transport for participating media. In *Rendering Techniques (Proceedings of the Eurographics Workshop on Rendering)*. Springer, 2000. [pages 26, 27, and 51]

- [173] K. Perlin and E. M. Hoffert. Hypertexture. In *Comp. Graph. (Proceedings of SIGGRAPH)*, pages 253–262. ACM, 1989. [page 50]
- [174] M. Pharr and G. Humphreys. *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2 edition edition, 2010. [pages 17, 18, 21, 36, 38, and 41]
- [175] B. T. Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, 1975. [pages 99 and 104]
- [176] A. D. Polyanin and A. V. Manzhirov. *Handbook of integral equations*. CRC press, 2008. [page 28]
- [177] J. E. Proctor and P. Y. Barnes. Nist high accuracy reference reflectometer-spectrophotometer. *Journal of Research of the National Institute of Standards and Technology*, 101(5):619–627, 1996. [page 98]
- [178] Y. Reibel, M. Jung, M. Bouhifd, B. Cunin, and C. Draman. Ccd or cmos camera noise characterisation. *The European Physical Journal - Applied Physics*, 21:75–80, 2003. [page 72]
- [179] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski. *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann, second edition, 2010. [pages 59 and 77]
- [180] R. Richtmeyer and J. von Neumann. Statistical methods in neutron diffusion. Technical report, Los Alamos National Laboratory, 1947. [page 29]
- [181] C. Robert and G. Casella. *Monte Carlo statistical methods*. Springer Science & Business Media, second edition, 2004. [pages 37, 38, and 47]
- [182] F. Romeiro and T. Zickler. Blind reflectometry. In *ECCV 2010*, 2010. [page 99]
- [183] F. Romeiro, Y. Vasilyev, and T. Zickler. Passive reflectometry. In *Computer Vision—ECCV 2008*, pages 859–872. Springer, 2008. [pages 97 and 102]
- [184] S. M. Rusinkiewicz. A new change of variables for efficient brdf representation. In *Rendering techniques' 98*, pages 11–22. Springer, 1998. [page 96]
- [185] F. Salvat, J. M. Fernández-Varea, and J. Sempau. Penelope-2006: A code system for monte carlo simulation of electron and photon transport. In *Workshop Proceedings*, volume 4, page 7, 2006. [page 24]

- [186] I. Sato, Y. Sato, and K. Ikeuchi. Acquiring a Radiance Distribution to Superimpose Virtual Objects onto a Real Scene. *IEEE Transactions on Visualization and Computer Graphics*, 5(1):1–12, 1999. [page 63]
- [187] M. Schoberl, A. Belz, J. Seiler, S. Foessel, and A. Kaup. High dynamic range video by spatially non-regular optical filtering. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 2757–2760. IEEE, 2012. [page 77]
- [188] T. B. Schön and F. Lindsten. *Learning of dynamical systems - Particle filters and Markov chain methods*. 2015. [page 35]
- [189] N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring image collections in 3d. *ACM Transactions on Graphics (SIGGRAPH 2006)*, 2006. [page 64]
- [190] J. Spanier and E. M. Gelbard. Monte carlo principles and neutron transport problems. *Reading, Mass., USA*, 1969. [page 23]
- [191] A. Srikantha and D. Sidibé. Ghost detection and removal for high dynamic range images: Recent advances. *Image Commun.*, 2012. [pages 5 and 73]
- [192] M. M. Stark, J. Arvo, and B. Smits. Barycentric parameterizations for isotropic brdfs. *IEEE Transactions on Visualization and Computer Graphics*, 11(2):126–138, 2005. [page 97]
- [193] J. C. Stover. *Optical Scattering: Measurement and Analysis, 2nd edition*. SPIE Publications, 1995. [pages 15, 100, 101, 103, and 109]
- [194] R. Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010. [page 77]
- [195] H. Takeda, S. Farsiu, and P. Milanfar. Kernel regression for image processing and reconstruction. *IEEE Transactions On Image Processing*, 16(2): 349–366, 2007. [pages 82 and 87]
- [196] H. Takeda, P. Milanfar, M. Protter, and M. Elad. Super-resolution without explicit subpixel motion estimation. *Image Processing, IEEE Transactions on*, 18(9):1958–1975, 2009. [page 91]
- [197] N. I. Technologies. Magic technology - scene contrast indexed image sensing with wdr. White paper, [www.new-imaging-technologies.com/download-wpan.html](http://www.new-imaging-technologies.com/download-wpan.html), 2008. [page 74]
- [198] R. Tibshirani and T. Hastie. Local likelihood estimation. *Journal of the American Statistical Association*, 82(398):559–567, 1987. [pages 79, 82, and 84]

- [199] L. Tierney. Markov chains for exploring posterior distributions. *the Annals of Statistics*, pages 1701–1728, 1994. [page 38]
- [200] M. D. Tocci, C. Kiser, N. Tocci, and P. Sen. A Versatile HDR Video Production System. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2011)*, 30(4):41:1–41:10, 2011. [pages 75, 77, 78, 88, and 89]
- [201] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV*, pages 839–846, 1998. [page 82]
- [202] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9): 1105–1112, 1967. [page 4]
- [203] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America*, 57(9): 1105–1112, 1967. [page 101]
- [204] A. Tsirikoglouy, S. Ekeberg, J. Vikström, J. Kronander, and J. Unger. S(wi)ss: A flexible and robust sub-surface scattering shader. In *SIGRAD*, 2014.
- [205] J. Unger and S. Gustavson. High-dynamic-range video for photometric measurement of illumination. In *SPIE Electronic Imaging*, 2007. [pages 73 and 77]
- [206] J. Unger, A. Wenger, T. Hawkins, A. Gardner, and P. Debevec. Capturing and rendering with incident light fields. In *Rendering Techniques 2003 (Proc. of EGSR)*, pages 1–10, 2003. [page 63]
- [207] J. Unger, S. Gustavson, M. Ollila, and M. Johannesson. A real time light probe. In *In Proceedings of the 25th Eurographics Annual Conference*, volume Short Papers and Interactive Demos, pages 17–21, 2004. [page 77]
- [208] J. Unger, S. Gustavson, and A. Ynnerman. Densely sampled light probe sequences for spatially variant image based lighting. In *GRAPHITE*, pages 341–347, 2006. [page 63]
- [209] J. Unger, S. Gustavson, P. Larsson, and A. Ynnerman. Free form incident light fields. *Computer Graphics Forum (Proc. of EGSR)*, 27(4):1293–1301, 2008. [page 6]
- [210] J. Unger, J. Kronander, P. Larsson, S. Gustavson, J. Löw, and A. Ynnerman. Spatially varying image based lighting using hdr-video. *Computers & graphics*, 37(7):923–934, 2013.

- [211] J. Unger, J. Kronander, P. Larsson, S. Gustavson, and A. Ynnerman. Temporally and Spatially Varying Image Based Lighting using HDR-video. In *Proceedings of EUSIPCO'13: Special Session on HDR-video*, 2013.
- [212] V. Vaish, M. Levoy, R. Szeliski, C. L. Zitnick, and S. B. Kang. Reconstructing occluded surfaces using synthetic apertures: Stereo, focus and robust measures. In *Proc. CVPR '06*, pages 2331–2338, 2006. [page 64]
- [213] E. Veach. *Robust monte carlo methods for light transport simulation*. PhD thesis, Stanford University, 1997. [pages 26, 27, 39, 45, and 46]
- [214] E. Veach and L. Guibas. Bidirectional estimators for light transport. In *Photorealistic Rendering Techniques*, pages 145–167. Springer, 1995. [page 41]
- [215] E. Veach and L. J. Guibas. Optimally combining sampling techniques for monte carlo rendering. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 419–428. ACM, 1995. [page 33]
- [216] E. Veach and L. J. Guibas. Metropolis light transport. In *SIGGRAPH*, pages 65–76, 1997. [pages 41, 44, 45, and 47]
- [217] S. VR. Spherocam hdr. [www.spheron.com](http://www.spheron.com), 2008. [page 59]
- [218] J. Waese. A real time high dynamic range light probe. In *SIGGRAPH Sketches*, 2003. [page 61]
- [219] B. Walter, S. R. Marschner, H. Li, and K. E. Torrance. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, pages 195–206. Eurographics Association, 2007. [page 101]
- [220] G. J. Ward. Measuring and modeling anisotropic reflection. *Computer Graphics (Proc. of SIGGRAPH)*, 26(2):265–272, 1992. [pages 4, 98, and 100]
- [221] G. J. Ward, F. M. Rubinstein, and R. D. Clear. A ray tracing solution for diffuse interreflection. *Computer Graphics (Proc. of SIGGRAPH)*, 22(4): 85–92, 1988. [page 41]
- [222] S. H. Westin, J. R. Arvo, and K. E. Torrance. Predicting reflectance functions from complex surfaces. *Computer Graphics (Proc. of SIGGRAPH)*, 26(2):255–264, 1992. [pages 101 and 102]
- [223] T. Weyrich, J. Lawrence, H. Lensch, S. Rusinkiewicz, and T. Zickler. Principles of appearance acquisition and representation. *Foundations and Trends in Computer Graphics and Vision*, 4(2):75–191, 2009. [page 93]

- [224] T. Whitted. An improved illumination model for shaded display. *Communications of the ACM*, 23(6):343–349, 1980. [page 3]
- [225] B. Widrow and I. Kollár. Quantization noise. *Cambridge University Press*, 2:5, 2008. [page 72]
- [226] D. N. Wood, D. I. Azuma, K. Aldinger, B. Curless, T. Duchamp, D. H. Salesin, and W. Stuetzle. Surface light fields for 3d photography. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques, SIGGRAPH '00*, pages 287–296, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co. [page 65]
- [227] G. Xing, X. Zhou, and Q. Peng. Lighting Simulation of Augmented Outdoor Scene Based on a Legacy Photograph. *Computer Graphics Forum (Proc. of Pacific Graphics 2013)*, 32(7), 2013. [page 58]
- [228] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar. Generalized Assorted Pixel Camera : Postcapture Control of Resolution , Dynamic Range , and Spectrum. *IEEE Transactions on Image Processing*, 19(9):2241–2253, 2010. [page 77]
- [229] C. Yu, Y. Seo, and S. W. Lee. Global optimization for estimating a brdf with multiple specular lobes. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 319–326. IEEE, 2010. [page 101]
- [230] K. Zhang, X. Gao, D. Tao, and X. Li. Single image super-resolution with non-local means and steering kernel regression. *Image Processing, IEEE Transactions on*, 21(11):4544–4556, 2012. [page 114]





**Part II**  
**Publications**



# Publications

The articles associated with this thesis have been removed for copyright reasons. For more details about these see:

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-122588>