IMAGE-BASED RECONSTRUCTION AND MODELING OF GEOMETRY WITH COMPLEX REFLECTANCE IN NON-DIFFUSIVE MEDIA

BY

SEBASTIAN MAGDA

M.S., University of Illinois at Urbana Champaign, 2003
B.S., University of Illinois at Urbana Champaign, 1999
B.S., University of Illinois at Urbana Champaign, 1999

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Computer Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2007

Urbana, Illinois

Abstract

Reconstruction and modeling of 3-D geometry are among the core issues in computer vision. While geometry composed of radiometrically simple materials is currently relatively easy to reconstruct and model, geometry with complex reflectance properties presents a substantial challenge. The research work presented here is an investigation of this problem. Two methods are proposed that are specifically designed to handle geometry with complex photometric properties.

The first proposed method is designed to reconstruct large-scale geometry with arbitrary and possibly anisotropic BRDFs. Existing reconstruction techniques typically make explicit or implicit assumptions about the reflectance properties of a surface. The proposed method uses the idea of photometric ranging, where no such assumptions are necessary. In a photometric stereo-like setup of multiple images obtained from a single viewpoint under controlled illumination, photometric ranging recovers the depth directly for each camera pixel, rather than through surface normal field integration. It exploits the basic concept of radiant energy density falloff with distance from a point light source. Double-covering the incident light field allows to find sets of coincidental pairs of light directions where this can be used to align the reflected light fields and directly reconstruct the depth of a scene. Unlike photometric stereo, in photometric ranging no assumptions are required about the surface smoothness, the presence or absence of shadowing, or the nature of the BRDF, which may vary over the surface. Once the depth of a scene is known, the reflected light field can also be resampled to relight the scene, that is to render the same scene from the camera view, but under novel lighting, including nearby and distant sources.

The second proposed method aims to model small-scale geometry of volumetric surface materials with complex reflectance. Instead of recovering the intricate geometry itself, it uses an appearance-based approach of volumetric texturing, which is a popular technique for rendering rich 3-D detail when a polygonal surface representation would be ineffective. Although efficient algorithms for rendering volumetric textures have been known for years, capturing the richness of real volumetric materials remains a challenging problem. The proposed technique generates a volumetric representation of a complex 3-D texture with unknown reflectance and structure. From the acquired reflectance data in the form of a 6-D Bidirectional Texture

Function (BTF), the proposed method creates an efficient volumetric representation in the form of a stack of semi-transparent layers each representing a slice through the texture's volume. In addition to negligible storage requirements, this representation is ideally suited for hardware-accelerated real-time rendering.

To my mom and dad, who always supported me in difficult moments of my life.

Table of Contents

Chapter 1

Introduction

Achieving visual realism in computer generated rendeirng is a daunting task. The natural world is a visually complex system, with light interactions taking place on a quantum level. Although modeling various light phenomena has been a central topic in computer graphics for years, it is unlikely that a comparable level of detail to the real world will ever be reached. Light interactions take place in an environment that is likewise extremely complex. Modeling the complexity of the natural environment is the goal of computer graphics. As computing power continues to grow rapidly, modeling is becoming more accurate and realistic. This is especially true when we look at the improvements of the capabilities of graphics hardware. The quality of offline image synthesis has already reached the level where it becomes very difficult for humans to distinguish synthetic images form real ones. Moreover, thanks to the rapid advancements in graphics hardware, we are quickly approaching the same level in real-time image synthesis (see an example in Figure 1.1). However, as rendered scenes become more complex, even more effort must be put into scene modeling. Although it is entirely possible to manually build realistically-looking synthetic environments, it is usually a lot more efficient to scan the real thing. For example, the natural environment can be scanned using a digital camera or a laser range scanner, and the data can then be used to create a computer model representation. This process of *reconstruction* and *modeling* is the topic of this research work.

1.1 Research Work Overview

The research work presented here is an investigation of reconstruction and modeling of 3-D geometry with complex reflectance. It focuses on media that do not scatter light or are very weakly scattering, such as the air. Two algorithms are presented which are intended for two different objectives in reconstruction. Their common feature is the ability to handle geometry with complicated reflectance. While there are many reconstruction algorithms for reconstruction of geometry with simple reflectance, there are still relatively few that can handle complex photometric properties.

Figure 1.1: An example of state-of-the-art real-time rendering engine: CryENGINE™2 [46]. In the top row are frames rendered in real-time and modeled after the real photographs in the bottom row.

1.1.1 Photometric Ranging of Large-Scale Geometry

The first proposed method is intended for reconstruction of large-scale geometry with complex unknown reflectance. It is a shape reconstruction method that is appropriate for opaque scenes in non-diffusive environments. While existing shape reconstruction techniques typically make explicit or implicit assumptions about the reflectance properties of a surface, the algorithm is based on the idea of *photometric ranging*, where no assumptions about the reflectance properties are necessary.

The reflectance of a surface is typically described using a bidirectional reflectance distribution function (see Appendix A.3). A BRDF is a 4-D function of the incident and outgoing light direction that describes the reflectance properties of a surface as a ratio of the outgoing radiance to the incident irradiance. In computer vision and graphics, simplifications are typically necessary to the BRDF. Over the years, many parametric models have been created to simulate the BRDF of common materials. Nearly all geometric reconstruction methods make use of BRDF models, yet each of these characterizes only a limited class of materials. Most existing reconstruction methods require some knowledge or make implicit assumptions about the BRDF. In contrast, the proposed photometric ranging method does not have such requirements.

In a photometric stereo-like setup of multiple images obtained from a single viewpoint under controlled illumination, the algorithm can be used to recover the depth directly, rather than through surface normal field integration typical of other photometric methods. It achieves that by double-covering the incident light field. That allows it to find sets of coincidental pairs of light directions where the principle of irradiance

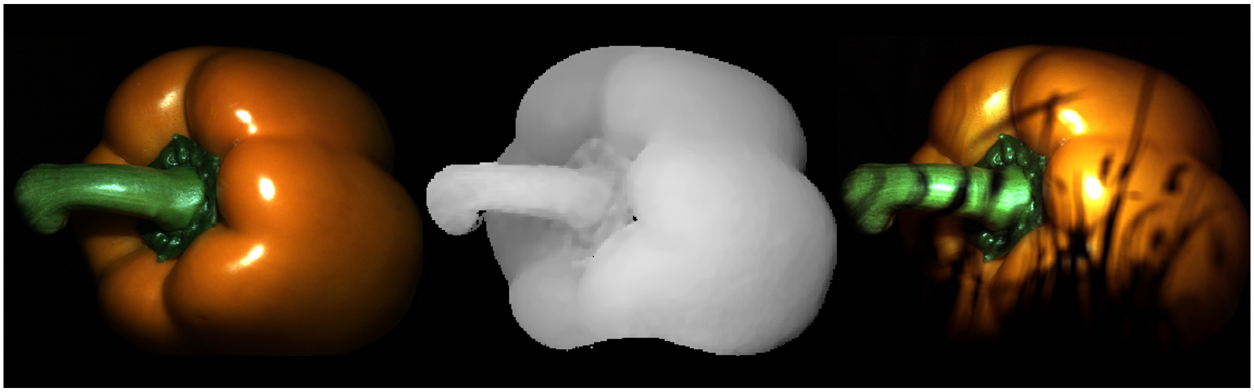

Figure 1.2: Shape reconstruction and relighting example using the proposed algorithm. On the left is one of the input images. In the center is the recovered depth map shown as a gray scale image. On the right is an example of geometry relighting using the original images and the recovered depth map.

falloff from a point light source can be used to align the reflected light fields and directly reconstruct the depth of a scene. Unlike photometric stereo, in photometric ranging no assumptions are required about the surface smoothness, the presence or absence of shadowing, or the nature of the BRDF, which may vary over the surface. Once the depth of a scene is known, the reflected light field can also be resampled to relight the scene, that is to render the same scene from the camera view, but under novel lighting, including nearby and distant sources (see Figure 1.2).

Key advantages of the proposed reconstruction algorithm can be summarized as follows:

- Shape reconstruction method that works for objects with arbitrarily complex BRDFs.
- The depth of a scene is recovered directly and independently at each pixel.
- Provides a 2-D slice of the apparent BRDF at each pixel.
- The original data and the recovered depth map can be used to render the object under novel illumination including nearby and distant light sources.

1.1.2 Modeling of Small-Scale Geometry

The second half of this research work contends with the problem of reconstruction and modeling of small-scale geometry of volumetric surface materials with complex reflectance properties. Unlike the first method, it focuses on efficient modeling, rather than on precise geometric reconstruction. Materials covering large elements of a scene, although they can have a very complex small-scale geometry, are rarely viewed up-close when rendered. Detailed rendering of small-scale geometry is therefore often superfluous. At the end, what

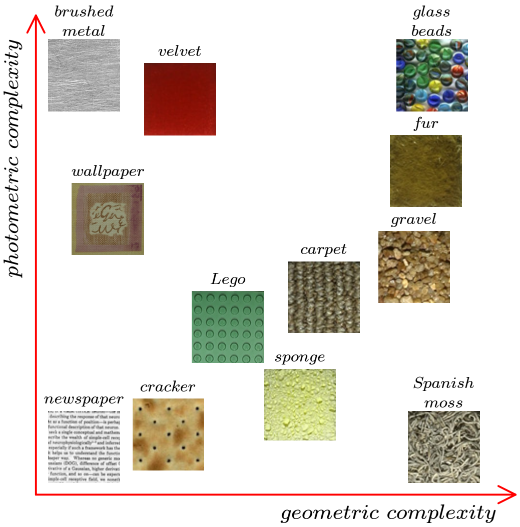

Figure 1.3: Appearance complexity space for textures. Horizontal and vertical axes indicate increasing geometric and photometric complexities. Textures in the lower right corner have relatively simple reflectance, yet their BTFs are complex due to geometry.

really matters in rendering is the visual resemblance of the rendered image to the real scene, as well as minimization of the rendering and the data storage costs.

Interactive applications are traditionally limited to a polygonal representation, which becomes very inefficient at surface macro scales. The standard solution to this problem is to approximate geometric detail with textures. Various texture representations offer a trade-off between visual realism and rendering/storage costs. The complexity of the modeled material is often the deciding factor. As such, relatively simple geometry of various smooth materials, such as concrete, wood, painted surfaces, can be visually reproduced with a flat albedo map represented with a color texture. To represent rougher material textures, the albedo map is usually supplemented with a normal map. Materials with more complex geometry and visible macro scale features require a more advanced representation for the same level of realism. Adding a displacement texture provides a way to represent materials for which geometry can be described with a depth map. More complicated materials typically require volumetric representations. While various methods have been proposed to render volumetric textures, there has been very little work on how to reconstruct them from image data. The proposed algorithm offers a possible solution.

The algorithm starts with a set of images of a volumetric surface material stored as bidirectional texture function samples (see Section 5.1.1). A BTF can be used directly for rendering, but high storage and synthesis costs make that unattractive in real-time applications. Real-time applications require efficient compression, resampling, and rendering algorithms. All of these are often very difficult to achieve. Yet for

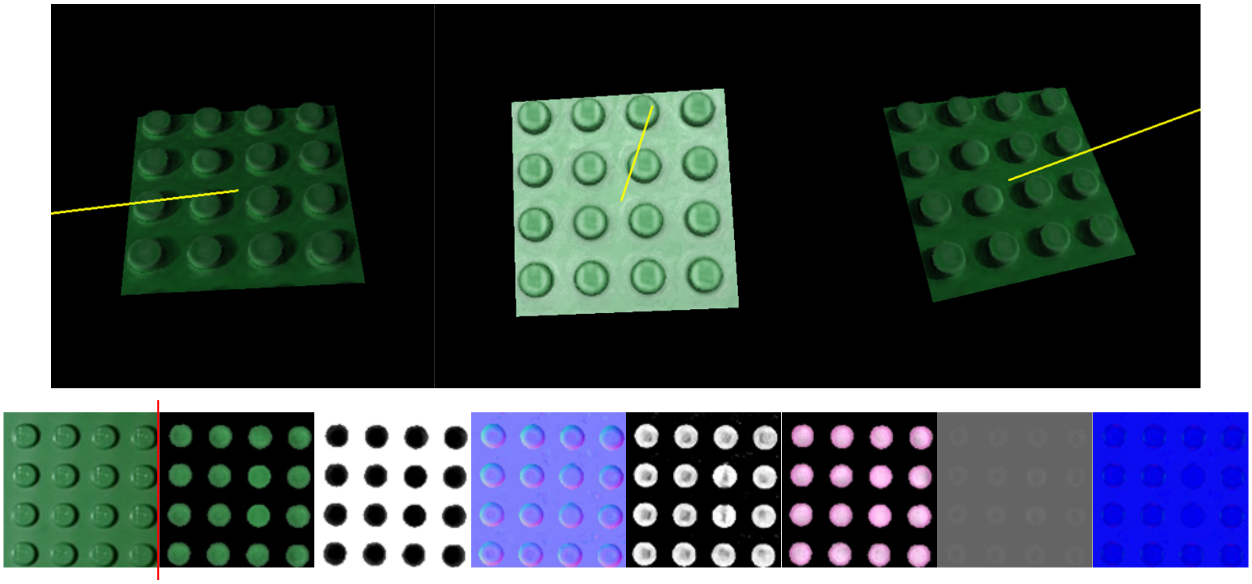

Figure 1.4: Three views of a volumetric texture synthesized from a BTF set. The yellow line indicates the light direction. Bottom row shows parameter textures for a single layer with a sample input texture to the left of the line.

many materials, this is the result of complex geometry, rather than the reflectance. Consider some examples in the lower right corner of Figure 1.3. Although these materials have relatively simple reflectance, their BTFs will be very complex due to geometry alone. The proposed algorithm works by separating this geometric information from the reflectance data making it possible to obtain a more compact representation. As a second benefit, extracting the geometry information also allows for more realistic rendering of volumetric materials at grazing angles, where direct BTF rendering methods cannot reproduce the occluding contour, as BTFs are parametrized in 2-D.

The proposed algorithm focuses on volumetric surface materials and takes an appearance-based approach to modeling. Instead of precisely modeling geometric details, the material is represented as a stack textures, each representing a volumetric slice of the original material sample. This volumetric representation is generated from a set of images taken with an ordinary camera and parametrized as a BTF. In addition to geometry, it also captures reflectance properties as a parametric BRDF model (see Figure 1.4). This results in a compact representation, effectively compressing a BTF. When viewed at a normal viewing distance, the appearance of a surface rendered using this approximation is hardly distinguishable from the original.

Key advantages of the proposed volumetric texture reconstruction and modeling method can be summarized as follows:

- Addresses an open problem of geometry extraction from BTFs.

- Works well for textures with complex geometry, reflectance, and shadowing.
- Separates the geometry from the reflectance and produces a very compact representation.
- Can be used to render the occluding contour in real-time applications.

1.2 Document Organization

This document is organized as follows. It starts with a review of the light transport model in Chapter 2. Emphasis is placed on common approximations used in 3-D scanning, computer vision, and computer graphics. Then Chapter 3 gives a broad overview of various 3-D scanning and reconstruction techniques. This includes true volumetric scanning typically used in medical imaging and also shape scanning techniques typically used for computer graphics applications. Chapter 4 introduces the algorithm for shape reconstruction of large-scale geometry with complex reflectance. The second algorithm intended for volumetric surface materials is described in Chapter 5. Finally the discussion of the two algorithms and future prospects follow in Chapter 6.

Chapter 2

The Light Transport Model

The core of computer image synthesis is simulation of light. In quantum physics “light” is viewed as a stream of elementary particles called *photons* that act as energy carriers for the electromagnetic force. Photons travel at the speed of light (299,792,458 m/s in vacuum) and they carry energy given by their oscillation frequency. Since frequency translates into wavelength, photons can be alternatively viewed as electromagnetic wave packets traveling at the speed of light. Interaction between these can lead to phenomena characteristic of waves, such as interference and diffraction. Visible light is a narrow band of electromagnetic radiation that is detectable by our eyes in the range of approximately 400 nm to 700 nm. This particle-wave duality is the manifestation of the dual nature of matter theorized by De Broglie [53]. It allows the light to be treated as a stream of particles, or as a wave, depending on the situation. In image synthesis applications both approaches often lead to complementary solutions, although the particle model of light is currently the dominant one. This is mostly a pragmatic choice, since the particle model is much easier to efficiently simulate on a computer, given the desired environment complexity and physical effects. However, with the particle model of light it is very difficult to model wave phenomena such as interference or diffraction.

This work also follows the particle model of light for the same efficiency reasons. While simultaneously modeling the wave nature of light would certainly allow for greater realism and flexibility, the particle model, as experiments will show, is typically sufficient. The next section describes in details the particle model of light used in this work, and, inadvertently, in computer graphics. It expands on the analysis of this topic presented by Andrew Glassner in his book “Principles of Digital Image Synthesis” [77].

2.1 The Particle Model of Light

In “Kinetic Theory of Gases” Ludwig Boltzmann formulated a statistical interpretation of thermal diffusion processes in dilute gases. The theory is a mathematical analysis of how the kinetic energy of molecules is distributed through collisions until a thermal equilibrium is reached. This heat diffusion process, or *heat transport*, is formalized in terms of statistical analysis of particle flow through a differential volume.

Interestingly, the same statistical treatment can be applied to many other particle systems reminiscent of gases. In fact, a few decades later, Boltzmann's theory has been adapted for simulating neutron interactions in atomic reactions. Since light can be thought of as a flow of photon particles through an environment, it too can be modeled in a similar fashion.

There are, of course, some major differences between photonic and molecular interactions. Photons are essentially infinitesimally small packets of energy and do not collide with each other like molecules. However, they do collide with scattering particles (dust, molecules, atoms, ions, etc.) present in a medium. Collisions can actually be a lot more complex. Most scattered photons traveling through a medium (air, for example) undergo elastic collisions in a process referred to as Rayleigh scattering (or equivalent Brillouin scattering in solids) [61]. In elastic collisions, the energy of an incoming photon does not transfer a collided particle to a higher energy level, but the photon is re-emitted in some other direction. This effect is very prominent every day when we look at the blue sky. The apparent blue color is the result of Rayleigh scattering varying with the wavelength of light. Inelastic scattering of visible light is usually a lot less frequent and can take a form of Raman scattering [184]. If an incoming photon has enough energy to transfer a molecule to a higher energy state, then any remaining energy is re-emitted as a lower energy photon. This can be viewed as a form of inelastic collision. High energy photons (more typical of x-rays) can also cause ionization by ejecting electrons. When the energy of an incoming photon is greater than the ionization energy of a particle and an electron is ejected, the remaining energy is re-emitted as a lower energy photon. This form of inelastic scattering is known as Compton scattering [41]. When the energy of a photon is completely absorbed, we are dealing with the process of *absorption*.

A photon-excited particle usually fairly quickly relaxes to its ground energy level. As it transfers to lower energy levels, it emits one or more photons. When the process produces photons with a different energy than the original source, it is called *fluorescence*. In unusual cases, the process can take a fairly long time, even minutes or hours. If photon emission is delayed more than about 10 ns, the process is called *phosphorescence*.

2.1.1 The Transport Equation

The particle model is a substantially simplified view of light transport. Photons are modeled as infinitesimally small massless particles traveling at a constant speed. The speed c depends only on the medium. Photons can be absorbed or scattered by the medium, but do not interact with each other. Each photon can be mathematically described by its position in space and direction (x, y, z, θ, ϕ). We can equivalently think of them as points in a 5-d Euclidean space $\mathcal{R}^5$, or a Cartesian product space of positions and direction vectors $\mathcal{R}^3 \otimes \mathcal{S}^2$ called the *phase space* [77]. Each photon carries an energy $\mathcal{E}$ proportional to its frequency ν (the

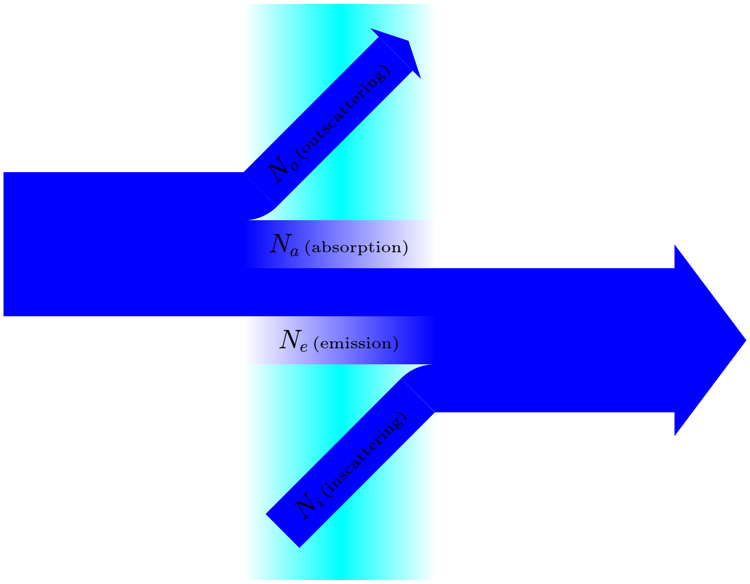

Figure 2.1: Schematic of a photon flow model in scattering media.

inverse of its wavelength λ) expressed as:

$$\mathcal{E} = h\,\nu = \frac{h\,c}{\lambda} \tag{2.1}$$

h is known as the *Planck's constant* equal to $6.62620 \times 10^{-34}\ \mathrm{J} \cdot \mathrm{s}$. Given some initial distribution of photon positions and their directions at time t, the particle model of light describes the photon phase space distribution after some time period T. In other words, it is a model of the photon flow through a volume V within some solid angle vector $\vec{\Omega}$. This 3-D particle transport model has been known since the 50s [34, 187, 76]. For an in-depth analysis of the transport model, refer to a book by Ishimaru [108].

The flow of photons through the medium can be expressed as a balance equation. Let N be the total number of photons within a range of wavelengths Λ traveling through a phase volume $(V, \vec{\Omega})$ at a time t. In scattering medium, some of these photons may flow through the medium with no interactions, but others will be absorbed or scattered, as indicated in the diagram in Figure 2.1. The scattering process can work either for or against the flow, meaning that photons could be "scattered into" the flow along $\vec{\Omega}$ or "scattered out" from the flow. Some additional photons can be also created inside the volume through emission. We can therefore express the change in the number of photons after traveling through the volume V as:

$$\Delta N = -N_a - N_o + N_i + N_e \tag{2.2}$$

N_a - loss due to absorption by the medium
N_o - loss due to outscattering from the direction $\vec{\Omega}$
N_i - gain due to inscattering into the direction $\vec{\Omega}$
N_e - gain due to emission from the medium

Let $\varrho(\mathbf{r}, \vec{\omega}, \lambda, t)$ $^{\text{particles}}/_{\text{m}^3\cdot\text{sr}}$ be the phase space time density of photons with a wavelength λ at a point $\mathbf{r}$ moving in a direction $\vec{\omega}$. The total number of photons N present in a flow through a volume V in a solid angle direction $\vec{\Omega}$ with a wavelength range Λ at a time t can be calculated by integrating the photon density function ϱ:

$$N = \int_{\vec{\Omega}} \int_V \int_\Lambda \varrho(\mathbf{r}, \vec{\omega}, \lambda, t)\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \tag{2.3}$$

If the density ϱ changes over a time period T, the number of photons will change accordingly:

$$\Delta N = \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T \frac{d\varrho(\mathbf{r}, \vec{\omega}, \lambda, t)}{dt}\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \tag{2.4}$$

Note, that the position of photons $\mathbf{r}$ also changes in time and therefore $\frac{d\varrho}{dt}$ must be expressed using the *chain rule* [52]:

$$\frac{d\varrho}{dt} = \frac{\partial\varrho}{\partial t} + \frac{\partial\varrho}{\partial\mathbf{r}} \cdot \frac{\partial\mathbf{r}}{\partial t} \tag{2.5}$$

$\frac{\partial\varrho}{\partial\mathbf{r}}$ is the directional derivative of ϱ, or the *gradient* $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$. $\frac{\partial\mathbf{r}}{\partial t}$ is the photon velocity. Since photons travel in the direction of $\vec{\omega}$, Equation 2.5 can be simplified to:

$$\frac{d\varrho}{dt} = \frac{\partial\varrho}{\partial t} + (\nabla\varrho) \cdot (c\,\vec{\omega}) \tag{2.6}$$

This partial derivative expansion can now be used to rewrite Equation 2.4 as a two term expression:

$$\Delta N = \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T \frac{\partial\varrho(\mathbf{r}, \vec{\omega}, \lambda, t)}{\partial t}\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} + \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\,\nabla\varrho(\mathbf{r}, \vec{\omega}, \lambda, t) \cdot \vec{\omega}\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \tag{2.7}$$

The first term measures the time rate of change in the photon density, while the second represents the total outflow (also known as *streaming*) of photons from the volume V. In the absence of photon interactions with the medium (light moving through an “empty” space), the number of photons remains constant ($\Delta N = 0$). The time rate of change of density depends only on the difference between the number of inflowing and outflowing photons:

$$\frac{\partial\varrho}{\partial t} = -c\,\nabla\varrho \cdot \vec{\omega} \tag{2.8}$$

Note that, if the photon density is constant in time, it implies that the light source is static, i.e. the number of photons emitted by the light source and traveling through the volume V does not vary over time.

In a scattering environment, all terms in Equation 2.2 play a role in the overall balance and ΔN is no longer zero. Each of them either contributes to the photon flow, or causes losses. In most situations media act as attenuators and the overall balance is negative ($\Delta N < 0$).

The absorption term N_a stands for the number of photons absorbed in the volume V over the time period T. Let $\mu_a(\mathbf{r}, \vec{\omega}, \lambda)$ be the *directional absorption coefficient* [91] that defines the percentage of photons with a wavelength λ absorbed when traveling through the volume at a point $\mathbf{r}$ along a solid angle vector $\vec{\omega}$ over a distance $ds = c\,dt$. The absorption term can then be expressed as:

$$N_a = \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\,\mu_a(\mathbf{r}, \vec{\omega}, \lambda) \varrho(\mathbf{r}, \vec{\omega}, \lambda, t)\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \tag{2.9}$$

Outscattering is the result of interactions between photons and the medium not ending in photon's absorption. Photons are instead deflected in directions other than $\vec{\Omega}$. Scattering can be described using a *directional scattering coefficient* $\mu_s(\mathbf{r}, \vec{\omega}, \lambda)$ [91], similar to the absorption coefficient μ_a. Unlike an absorption event, in an outscattering event only the direction of travel and possibly the wavelength of a photon change. The distribution of scattered photons can be described by a normalized *scattering phase function* $p(\mathbf{r}, \vec{\omega} \to \vec{\omega}', \lambda \to \lambda')$. The function specifies the directional and spectral distribution of scattered photons (see Appendix A.3.3)[1]. Outscattering can be mathematically described as an integral over the number of photons with a wavelength λ that were out-scattered from the direction of $\vec{\omega}$, possibly changing the wavelength:

$$\begin{aligned} N_o &= \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\,\mu_s(\mathbf{r}, \vec{\omega}, \lambda) \int_{4\pi} \int_0^\infty p(\mathbf{r}, \vec{\omega} \to \vec{\omega}', \lambda \to \lambda') \varrho(\mathbf{r}, \vec{\omega}, \lambda, t)\, d\lambda'\, d\vec{\omega}'\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \\ &= \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\,\mu_s(\mathbf{r}, \vec{\omega}, \lambda) \varrho(\mathbf{r}, \vec{\omega}, \lambda, t)\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \end{aligned} \tag{2.10}$$

Since the integral of the normalized scattering phase function over the entire sphere of scatter directions and all scatter wavelengths is by definition equal to one (see Equation A.16), the final equation is actually quite simple, involving just the directional scattering coefficient μ_s and the density ϱ.

Inscattering is just the opposite of outscattering. Photons traveling in directions other than $\vec{\omega}$ change their direction (and possibly the wavelength) and contribute to the flow. The process can be described as an integral over all photons in-scattered from all directions into $\vec{\omega}$ undergoing a change of wavelength to λ:

---

[1] Note that, unlike the definition from Section A.3.3, this form of the phase function also models wavelength shifts.

$$N_i = \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\, \mu_s(\mathbf{r}, \vec{\omega}, \lambda) \int_{4\pi} \int_0^\infty p(\mathbf{r}, \vec{\omega}' \to \vec{\omega}, \lambda' \to \lambda) \varrho(\mathbf{r}, \vec{\omega}', \lambda', t)\, d\lambda'\, d\vec{\omega}'\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \tag{2.11}$$

Finally the emission term can be specified using a directional emission density function $\varrho_e(\mathbf{r}, \vec{\omega}, \lambda, t)$. The function defines the density of photons with a wavelength λ emitted over a distance $c\,dt$ along the direction of $\vec{\omega}$:

$$N_e = \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\, \varrho_e(\mathbf{r}, \vec{\omega}, \lambda, t)\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \tag{2.12}$$

After substituting all photon streaming terms, Equation 2.2 expands to:

$$\begin{aligned}
&\int_{\vec{\Omega}} \int_V \int_\Lambda \int_T \frac{\partial \varrho(\mathbf{r}, \vec{\omega}, \lambda, t)}{\partial t}\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} + \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\, \nabla \varrho(\mathbf{r}, \vec{\omega}, \lambda, t) \cdot \vec{\omega}\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} = \\
&- \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\, \mu_a(\mathbf{r}, \vec{\omega}, \lambda) \varrho(\mathbf{r}, \vec{\omega}, \lambda, t)\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \\
&- \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\, \mu_s(\mathbf{r}, \vec{\omega}, \lambda) \varrho(\mathbf{r}, \vec{\omega}, \lambda, t)\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \\
&+ \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\, \mu_s(\mathbf{r}, \vec{\omega}, \lambda) \int_{4\pi} \int_0^\infty p(\mathbf{r}, \vec{\omega}' \to \vec{\omega}, \lambda' \to \lambda) \varrho(\mathbf{r}, \vec{\omega}', \lambda', t)\, d\lambda'\, d\vec{\omega}'\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega} \\
&+ \int_{\vec{\Omega}} \int_V \int_\Lambda \int_T c\, \varrho_e(\mathbf{r}, \vec{\omega}, \lambda, t)\, dt\, d\lambda\, d\mathbf{r}\, d\vec{\omega}
\end{aligned} \tag{2.13}$$

The outermost integrals present in every term can be dropped. By dividing both sides by the speed of light c we obtain the *integro-differential form* [77] of the photon transport equation (some notation has been dropped for clarity):

$$\frac{1}{c} \frac{\partial \varrho}{\partial t} + \nabla \varrho \cdot \vec{\omega} = -(\mu_a + \mu_s) \varrho + \mu_s \int_{4\pi} \int_0^\infty p(\mathbf{r}, \vec{\omega}' \to \vec{\omega}, \lambda' \to \lambda) \varrho(\mathbf{r}, \vec{\omega}', \lambda', t)\, d\lambda'\, d\vec{\omega}' + \varrho_e \tag{2.14}$$

The absorption and the scattering coefficients are usually combined into a single *extinction coefficient*, μ, measuring the probability of a photon loss from the flow:

$$\mu' = \mu_a + \mu_s \tag{2.15}$$

The average distance a photon can travel in a medium before a collision, or its *mean free path*, is equal to the inverse $\frac{1}{\mu}$. The ratio of the number of scattered photons to the total number of collisions is sometimes referred to as the *scattering albedo* $\varpi = \frac{\mu_s}{\mu}$ [52]. The final form of the integro-differential transport equation is:

$$\frac{1}{c}\frac{\partial \varrho}{\partial t} + \nabla \varrho \cdot \vec{\omega} = -\mu \varrho + \mu_s \int_{4\pi} \int_0^{\infty} p(\mathbf{r}, \vec{\omega}' \to \vec{\omega}, \lambda' \to \lambda) \varrho(\mathbf{r}, \vec{\omega}', \lambda', t)\, d\lambda'\, d\vec{\omega}' + \varrho_e \tag{2.16}$$

Equation 2.16 is quite similar the original particle transport equation originally developed by Boltzmann known as the *Boltzmann equation*. It gives a general description of photon balance in a scattering medium over time. The equation remains an approximation based on several important assumptions:

- **Photons are point particles.** Although they are actually wave packets, they are treated as points. This limits the minimum spatial resolution of the model.

- **Photon interactions with the media occur instantaneously.** In reality, scattering involves absorption and then subsequent emission that does not happen instantaneously. This limits the minimum time scale resolution.

- **Photons do not interact with each other.** This essentially implies a low photon density relative to their actual size.

- **Photons travel in straight lines.** The refractive index of a medium is the same everywhere, otherwise photon light waves will undergo refraction, changing the path of individual photons.

- **Photon polarization can be ignored.** Although Equation 2.16 does not account for the polarization of light, polarization actually may still be modeled by a set of transport equations, one for each polarization parameter.

- **Other photon wave phenomena can be ignored.** This implies that the distance between scattering events is much greater than the size of photon wave packets and therefore interference can be ignored.

Of these assumptions the last three are the most limiting in computer graphics. Note, that all three refer to various wave phenomena of light, which cannot be modeled using a particle model.

The complete specification of the environment for the light transport equation includes the initial photon density, scattering, attenuation, and emission parameters, as well as the boundary conditions. The boundary conditions refer to the directional density of incident photons, which may also vary over time. This is the same as defining the external lighting.

2.1.2 Monochromatic Transport Equation for Isotropic Media

Equation 2.16 can be substantially simplified in the case of isotropic media. A medium is *isotropic* when the phase function p does not depend on the direction of the incidence and scatter, but only on the relative angle between these. The phase function can be parametrized in terms of the dot product of the two directions. Similar directional isotropy applies to emission, scattering, and absorption coefficients. Isotropic media are very common and include air, fog, water - essentially all gases and liquids.

Another common simplification is often made to light scattering. In many media the majority of scattering is elastic and therefore wavelength shifts usually do not occur. The light transport model that assumes elastic scattering is called *monochromatic*, as interactions can be considered at each wavelength independently. In isotropic media Equation 2.16 is then a model of photon transport for a single wavelength:

$$\frac{1}{c}\frac{\partial \varrho(\mathbf{r},\vec{\omega},t)}{\partial t} + \nabla \varrho(\mathbf{r},\vec{\omega},t)\cdot\vec{\omega} + \mu(\mathbf{r})\varrho(\mathbf{r},\vec{\omega},t) = \mu_s(\mathbf{r})\int_{4\pi} p(\mathbf{r},\vec{\omega}'\cdot\vec{\omega})\varrho(\mathbf{r},\vec{\omega}',t)\,d\vec{\omega}' + \varrho_e(\mathbf{r},t) \tag{2.17}$$

Notice, that the emission density function ϱ_e is now isotropic.

Recall that the phase function (Equation A.16) specifies the directional probability distribution of a scattering event. In an isotropic medium, the distribution is the same for all incident directions and can be parametrized simply with the cosine of the angle formed by the incident and the scattered directions. The mean value of the cosine of the scattering angle is called an *anisotropy coefficient* [22]:

$$g = \int_{4\pi} p(\vec{\omega}'\cdot\vec{\omega})\,\vec{\omega}'\cdot\vec{\omega}\,d\vec{\omega}' \tag{2.18}$$

g can vary from -1 for totally backward scattering, to 1 for forward scattering. When g is equal to 0, the scattering is considered isotropic (equal scattering probability in all directions). In a diffusive medium ($\mu_a \ll \mu_s$) with strongly biased scattering in the forward direction, further approximation can be made by replacing the scattering coefficient with the *reduced scattering coefficient* μ_s' [91]. The reduced scattering coefficient is equal to the scattering coefficient modified by the anisotropy:

$$\mu_s' = \mu_s(1-g) \tag{2.19}$$

μ_s' can be thought of as the effective diffusion of photons over a distance of $1/\mu_s'$, the result of many smaller diffusion steps over $1/\mu_s$. The absorption and the reduced scattering coefficients are then often combined into a single *reduced extinction coefficient* [91]:

$$\mu' = \mu_a + \mu'_s \tag{2.20}$$

This simplification is quite commonly used to approximate isotropic media such as air, water, fog, or smoke. It is also used in medicine for *Optical Tomography* (see Section 3.1.2. Human tissues can be considered isotropic and typical values of the anisotropy coefficient g typically vary between 0.65 and 0.95 [37]. Light tends to be scattered forward along the direction of a beam.

2.1.3 Steady-State Transport Equation

In many situations the properties of a medium and boundary conditions (the external illumination) do not vary in time. The density distribution of photons is then going to reach a steady-state i.e. $\frac{\partial \varrho}{\partial t} = 0$ and Equation 2.17 will no longer be time-dependent:

$$\nabla \varrho(\mathbf{r}, \vec{\omega}) \cdot \vec{\omega} + \mu(\mathbf{r}) \varrho(\mathbf{r}, \vec{\omega}) = \mu_s(\mathbf{r}) \int_{4\pi} p(\mathbf{r}, \vec{\omega}' \cdot \vec{\omega}) \varrho(\mathbf{r}, \vec{\omega}') \, d\vec{\omega}' + \varrho_e(\mathbf{r}) \tag{2.21}$$

Time variations occurring to a medium or the external illumination are really only relevant if they occur on a time scale comparable to the average time a photon spends in the medium, which is typically on a micro-, or just a nanosecond scale. This is, for example, well below the typical time frame rates used in computer graphics and therefore virtually all rendering algorithms use the steady-state form of the transport equation. On the other hand, the time-dependent form is explicitly used in many optical tomography techniques (see Section 3.1.2).

2.1.4 Integral Form of the Steady-State Transport Equation

The integro-differential form of the transport equation can be difficult to solve directly. Usually it is a lot more convenient in an integral form, as described by Duderstadt and Martin [62]. The first step in the process is to realize that, since photons travel in straight lines between collisions, their position $\mathbf{r}$ is really a function of the direction $\vec{\omega}$. Protons' positions can be traced back to some point s along a line and the position could be expressed as a vector line equation:

$$\mathbf{r}(s) = \mathbf{r} - s \, \vec{\omega} \tag{2.22}$$

This observation allows for some significant simplifications. The directional derivative of the density ϱ at $\mathbf{r}$ in the direction $\vec{\omega}$ can be expressed now as a derivative of ϱ along a straight line:

$$\nabla\varrho(\mathbf{r},\vec{\omega})\cdot\vec{\omega} = -\frac{\partial}{\partial s}\varrho\big(\mathbf{r}(s),\vec{\omega}\big)\Big|_{s=0} \tag{2.23}$$

Symbolically, Equation 2.21 can be rewritten as:

$$\begin{aligned}
-\frac{\partial}{\partial s}\varrho\big(\mathbf{r}(s),\vec{\omega}\big) + \mu\big(\mathbf{r}(s)\big)\varrho\big(\mathbf{r}(s),\vec{\omega}\big) &= \mu_s\big(\mathbf{r}(s)\big)\int_{4\pi} p\big(\mathbf{r}(s),\vec{\omega}'\cdot\vec{\omega}\big)\varrho\big(\mathbf{r}(s),\vec{\omega}'\big)\,d\vec{\omega}' + \varrho_e\big(\mathbf{r}(s)\big) && (2.24)\\
\varrho'\big(\mathbf{r}(s),\vec{\omega}\big) - \mu\big(\mathbf{r}(s)\big)\varrho\big(\mathbf{r}(s),\vec{\omega}\big) &= -G\big(\mathbf{r}(s),\vec{\omega}\big) && (2.25)
\end{aligned}$$

$G\big(\mathbf{r}(s),\vec{\omega}\big)$ represents the gain due to inscattering and emission. Note, that Equation 2.25 is a standard first-order differential equation. It can be solved by using an exponential integrating factor (see Appendix B for details) to give the photon density at a phase space location s_1. The result is expressed as an integral of all photon transport events occurring over the interval $[s_0, s_1]$ along the direction of $\vec{\omega}$. Solving for ϱ using Equation B.5 yields:

$$\varrho\big(\mathbf{r}(s_1),\vec{\omega}\big) = a(s_0)\varrho\big(\mathbf{r}(s_0),\vec{\omega}\big) + \int_{s_0}^{s_1} a(s)G\big(\mathbf{r}(s),\vec{\omega}\big)\,ds \tag{2.26}$$

where $a(s)$ is the integrating factor similar to B.3 representing the *ray attenuation function.* The attenuation function is defined as:

$$a(s) = \mathrm{e}^{-\int_0^{\|s_1-s\|}\mu\left(\mathbf{r}(u)\right)\,du} \tag{2.27}$$

Alternatively, it is more common to express the ray attenuation function in terms of the *optical thickness* τ [52]:

$$\begin{aligned}
&a(s) = \mathrm{e}^{-\tau\left(\mathbf{r}(s),\mathbf{r}(s_1)\right)} && (2.28)\\
&\text{where :}\\
&\tau(\mathbf{r}_0,\mathbf{r}_1) = \int_0^{\|\mathbf{r}_1-\mathbf{r}_0\|}\mu(\mathbf{r}_0 + \overrightarrow{\mathbf{r}_0\mathbf{r}_1}u)\,du && (2.29)
\end{aligned}$$

The optical thickness represents the total ray attenuation along the line between points $\mathbf{r_0}$ and $\mathbf{r_1}$. This quantity is indirectly measured during *computer tomography* (see Section 3.1.1) and later used to reconstruct the spatially-varying extinction coefficient function of the medium. Expanding Equation 2.26 using the attenuation function expressed in terms of the optical thickness τ yields:

$$\begin{aligned} &\varrho\big(\mathbf{r}(s_1), \vec{\omega}\big) = \\ &\quad \mathrm{e}^{-\tau\left(\mathbf{r}(s_0), \mathbf{r}(s_1)\right)} \varrho\big(\mathbf{r}(s_0), \vec{\omega}\big) \\ &\quad + \int_{s_0}^{s_1} \mathrm{e}^{-\tau\left(\mathbf{r}(s), \mathbf{r}(s_1)\right)} \left[\mu_s\big(\mathbf{r}(s)\big) \int_{4\pi} p\big(\mathbf{r}(s), \vec{\omega}' \cdot \vec{\omega}\big) \varrho\big(\mathbf{r}(s), \vec{\omega}'\big)\, d\vec{\omega}' + \varrho_e\big(\mathbf{r}(s)\big)\right] ds \end{aligned} \tag{2.30}$$

Equation 2.30 has a very intuitive physical interpretation. The photon phase space density at a point $\mathbf{r}(s_1)$ is the sum of the incident photons at a point $\mathbf{r}(s_0)$ plus the in-scattered and the volume-emitted photons along $\vec{\omega}$ minus the photons lost due to attenuation.

2.2 The Radiance Model

An alternative to the modeling of light as a flow of particles is to view it as a flow of energy. The energy of light and photon density are directly related quantities. Each photon is a wave-particle packet that carries energy inversely proportional to its wavelength (see Equation 2.1). The flow of light energy is expressed using *radiance* (see Appendix A). Radiance is a measure of energy flow per unit area per unit solid angle (see Equation A.8). Conversion of the transport equation from the particle form to a radiance form is simply a matter of multiplying the phase-space photon density by the photon energy:

$$L = \varrho \frac{h\,c}{\lambda} \tag{2.31}$$

In expressions involving radiance it is also common to express scattering using *bidirectional distribution functions* (BDFs) (see Appendix A). A BDF is similar to the scattering phase function, but it directly relates the outgoing radiance to the incoming irradiance. The radiance form of Equation 2.30 can be written as:

$$\begin{aligned} &L\big(\mathbf{r}(s_1), \vec{\omega}\big) = \\ &\quad \mathrm{e}^{-\tau\left(\mathbf{r}(s_0), \mathbf{r}(s_1)\right)} L\big(\mathbf{r}(s_0), \vec{\omega}\big) \\ &\quad + \int_{s_0}^{s_1} \mathrm{e}^{-\tau\left(\mathbf{r}(s), \mathbf{r}(s_1)\right)} \left[\int_{4\pi} f\big(\mathbf{r}(s), \vec{\omega}' \cdot \vec{\omega}\big) L\big(\mathbf{r}(s), \vec{\omega}'\big) \cos(\theta')\, d\vec{\omega}' + L_e\big(\mathbf{r}(s)\big)\right] ds \end{aligned} \tag{2.32}$$

Note that the BDF f incorporates the scattering coefficient μ_s. Conversion from irradiance to radiance in the integral introduced an extra cosine term of the angle θ between the in-scattered direction $\vec{\omega}'$ and the

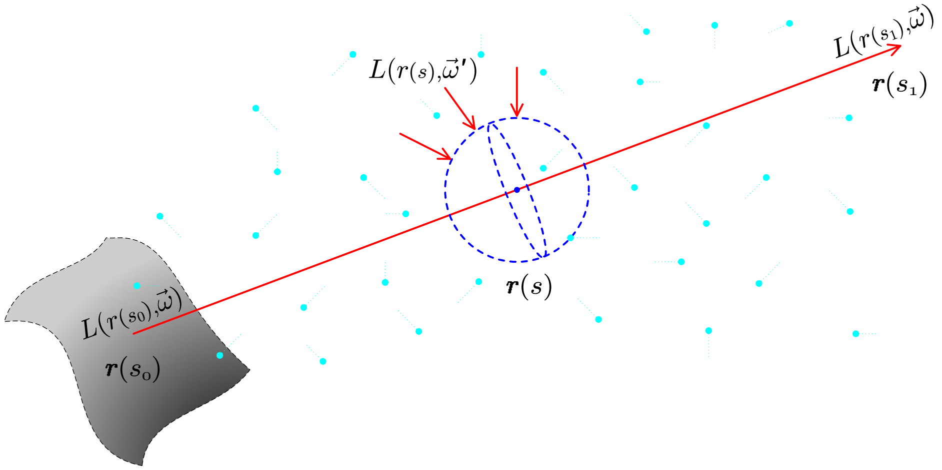

Figure 2.2: Photon transport in scattering media. The final radiance at the point $\mathbf{r}(s_1)$ is the sum of the incident radiance at a point $\mathbf{r}(s_0)$ plus the in-scattered radiance along the ray. In addition, the final radiance also includes the volume-emitted radiance gain and the radiance lost due to attenuation, which are not shown in the figure.

direction of the flow $\vec{\omega}$ (this assumes BDF isotropy). Similar to Equation 2.30, the radiance at the point $\mathbf{r}(s_1)$ is the sum of the incident radiance at a point $\mathbf{r}(s_0)$ plus the in-scattered and the volume-emitted radiance along $\vec{\omega}$ minus the radiance lost due to attenuation (see Figure 2.2). In typical situations most media do not emit light i.e. all light energy comes from external light sources and Equation 2.32 further reduces to:

$$L\big(\mathbf{r}(s_1),\vec{\omega}\big) = \int_{s_0}^{s_1} \mathrm{e}^{-\tau\left(\mathbf{r}(s),\mathbf{r}(s_1)\right)} \int_{4\pi} f\big(\mathbf{r}(s),\vec{\omega}_v \cdot \vec{\omega}\big) L\big(\mathbf{r}(s),\vec{\omega}'\big) \cos\theta' \, d\vec{\omega}' \, ds \tag{2.33}$$

Note, that boundary and initial condition terms were omitted from derivations for readability reasons.

2.2.1 Full Radiance Equation

The radiance model from Equation 2.32 is not a complete description of light in every situation, although it is sufficient in most cases. Still it can often be expanded to account for some phenomena. These may include polarization [2], phosphorescence, and fluorescence. Without going into derivation details (for full details, please refer to the book by Glassner [77]), this more complete *full radiance equation* can be expanded to include incandescent, phosphorescent, and fluorescent terms:

---

[2] Polarization plays a role in reflections, as the light reflected off a surface at the Brewster's angle is linearly polarized.

$$
\begin{aligned}
L\big(\mathbf{r}(s_1),\vec{\omega},\lambda,\vec{e},t\big) = \int_{s_0}^{s_1} \mathrm{e}^{-\tau\left(\mathbf{r}(s),\mathbf{r}(s_1)\right)} \Bigg[& \\
& L_e\big(\mathbf{r}(s),\vec{\omega},t,\lambda\big) \\
& + m_p(\vec{\omega}) \int_{-\infty}^{t} \gamma_p(t-t_\gamma)\mathcal{E}_p\big(\mathbf{r}(s),\lambda\big) \int_{4\pi} L\big(\mathbf{r}(s),\vec{\omega}',\lambda,\vec{e},t_\gamma\big)\cos(\theta')\,d\vec{\omega}'\,dt_\gamma \\
& + \int_{4\pi} f\big(\mathbf{r}(s),\vec{\omega}_v \rightarrow \vec{\omega},\lambda\big) \int_{\Lambda} \mathcal{E}_f\big(\mathbf{r}(s),\lambda_i \rightarrow \lambda\big) L\big(\mathbf{r}(s),\vec{\omega}',\lambda_i,\vec{e},t\big)\cos(\theta')\,d\lambda_i\,d\vec{\omega}' \\
& \Bigg] ds
\end{aligned}
\tag{2.34}
$$

$m_p(\vec{\omega})$ - phosphorescence modulation function
$\gamma_p(t-t_\gamma)$ - phosphorescence time decay function
$\mathcal{E}_p\big(\mathbf{r}(s),\lambda\big)$ - phosphorescence wavelength efficiency
$\mathcal{E}_f\big(\mathbf{r}(s),\lambda_i \rightarrow \lambda\big)$ - fluorescence efficiency for transfer of energy from λ_i to λ
$\vec{e}$ - ellipsometric light polarization vector

Note that boundary and initial condition terms were again omitted for clarity. The full radiance equation is rarely used directly in computer graphics due to the computational cost involved.

Chapter 3

Three-Dimensional Scanning Techniques

Scanning in computer science is a process by which some physical properties of the real world are measured and eventually digitized into a form that can be processed by a computer. The number of physical dimensions over which the scanning takes place determines the complexity of the process. Acquisition in one-dimension, for example, detection of a changing sound wave at one point (time would be the only dimension) , requires a relatively simple microphone. Scanning in two-dimensions typically requires a more complex scanning methodology. A 2-D slice of radiance arriving from a particular range of directions or an *image* can be captured using a digital camera. A digital camera is a relatively complex device containing a 2-D[1] light detector, typically either a CCD or a CMOS array [138]. Another 2-D scanning device, a flatbed scanner, uses a 1-D light detector moved mechanically across the page to scan in the second dimension.

Scanning physical properties in three dimensions is more complex yet. This chapter is an overview of various volumetric scanning techniques, especially relevant in computer-aided medical imaging, typically comprising scanning of volumetric density or attenuation. In addition, this chapter is an overview of *shape*, or *geometric reconstruction* methods. Shape reconstruction is an active research area in engineering and computer vision dealing with the problem of recovery of surfaces of solid objects in transparent media (most often the clear air). Reconstruction methods proposed in Chapter 4 and Chapter 5 also fall in the same category as they deal with solid surfaces in clear media. Yet while the method from Chapter 4 is designed to recover surfaces of solid, opaque objects, the method from Chapter 5 is closer to volumetric methods used in medical imaging and recovers a volumetric density representation.

3.1 Volumetric Reconstruction Methods

Medicine has always been on a lookout for a better way to peer inside the human body in order to study, diagnose, and heal. The biggest challenge has always been how to do so with minimal harm to the patient. The field of non-invasive medical imaging has its beginnings in 1895, when Wilhelm Roentgen took first

---

[1] Technically the wavelength of light could be considered an another dimension.

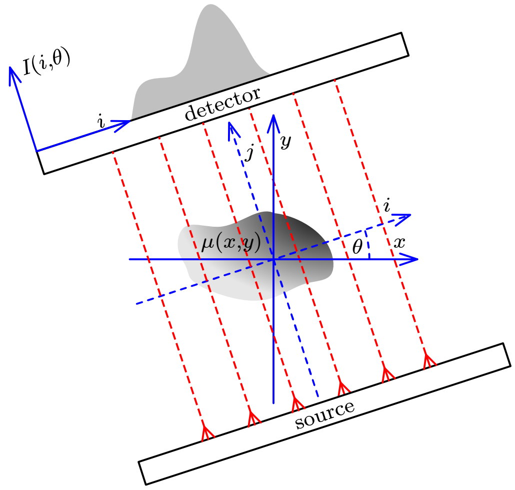

Figure 3.1: 2-D transmission computed tomography (CT). A single measurement $I(i, \theta)$ represents the total attenuation along the direction j for a particular rotation angle θ of the source and the detector.

X-ray images. X-rays are high energy photons that can penetrate soft tissues, but are strongly absorbed by denser tissues (such as bone and cartilage), thus providing an image of the interior of the body. The appearance of computers opened the possibility of interactive viewing of data in all three dimensions. Using X-rays to create 3-D images was the next logical step.

3.1.1 Computed Tomography

In the mid-70s appeared a medical scanning technique named *computed tomography* (CT) or, alternatively, *computed axial tomography* (CAT). CT was first developed by Hounsfield [100] and built by EMI Central Research Laboratories. In CT, a linear beam of X-rays is passed through a body and the beam attenuation is measured by a detector. The source and the detector are rotated around the body and a set of measurements are taken. Each measurement provides a one-dimensional slice showing the attenuation along a particular direction (see Figure 3.1). The volumetric reconstruction process is relatively simple because scattering and bending of X-rays as they travel through the body is marginal and can be ignored. The photon transport equation in Equation 2.32 reduces to the first term, as the only remaining process is the absorption. While tomography can be done in three dimensions, it is often simplified to measuring 2-D slices by rotating a parallel X-ray source and a detector on the opposite sides of a patient's body. In this simpler version, the attenuation for some source/detector rotation angle θ for a parallel beam of X-rays with power area density W_0 (see Equation A.2) is written as:

$$
\begin{aligned}
W(i,\theta) &= W_0 e^{-\int_{source}^{detector} \mu(x,y)dj} \\
&= W_0 e^{-\int \mu\left(i\cos(\theta)-j\sin(\theta),i\sin(\theta)+j\cos(\theta)\right)\,dj}
\end{aligned}
\tag{3.1}
$$

Here $\mu(x,y)$ denotes the 2-D X-ray extinction coefficient (see Equation 2.15). Since scattering is ignored, the extinction coefficient is essentially equivalent to the absorption coefficient μ_a. The detected 1-D projection along the i axis corresponds to the attenuation integral of μ along the j axis. Since the X-ray detector is actually providing a logarithmic measure of the beam power area density, the measured X-ray brightness I simplifies to:

$$
\begin{aligned}
I(i,\theta) &\propto \ln\left(W(i,\theta)\right) \\
&= I_0 \int \mu\left(i\cos(\theta)-j\sin(\theta), i\sin(\theta)+j\cos(\theta)\right)\,dj \\
&= I_0 \int\int \mu(x,y)\delta\left(i-x\cos(\theta)-y\sin(\theta)\right)\,dx\,dy
\end{aligned}
\tag{3.2}
$$

δ is the Dirac delta function (or the impulse function). Equation 3.2 is called the *Radon transform* [92] of the attenuation $\mu(x,y)$. Finding μ, requires solving the inverse problem i.e. finding the *inverse Radon transform*:

$$
\mu(x,y) = \int_0^\pi I\left(x\cos(\theta)+y\sin(\theta),\theta\right)d\theta \tag{3.3}
$$

In real-world applications only a discrete set of measurements is collected and θ is a discrete variable. The image reconstruction is performed using a discrete version of Equation 3.3.

The core requirement of CT is for the radiated waves to pass through the examined object along straight lines. Although X-rays work quite well, there are other lower energy methods less harmful to people. Yet lower energy waves are also more susceptible to diffraction and refraction and therefore are more difficult to use in tomography.

There are currently many variations of the basic axial tomographic technique. In general, tomography can be divided into three sub-genres: transmission, emission, and reflection tomography. All rely on solving for the inverse Radon transform therefore require that photons travel along straight lines.

Transmission Tomography

Transmission tomography measures the attenuation of radiation traveling completely through an object. X-ray transmission tomography described in the previous section was the first tomographic technique and remains highly popular today. Yet using X-rays have negative side-effects on the human body and finding an alternative has always been of great interest.

Even before invention of computers there have been attempts to employ visible light as an alternative to X-rays in examination of tissues [48]. Although the early experiments have shown that the light is inferior to X-rays due to scattering, its harmless nature still made it desirable as an examination tool. Since the early 90's visible light have been used in full-scale CT scans [110, 236, 230]. The newest transmission techniques typically use coherent light sources in the near infra-red part of the spectrum and belong to the growing area of *Optical Tomography* (OT) (see Section 3.1.2). In addition to visible light, researchers continue to investigate other forms of illumination, including ultrasound [81, 111, 182] and microwave radiation [116].

Emission Tomography

Rather than using an external radiation sources, emission tomography works by detecting an internally introduced radiation source. Usually these come in a form of compounds readily absorbed by the body containing radioactive isotopes with short half-lifes such as Carbon-11, Nitrogen-13, Oxygen-15, or Fluorine-18. The radioactive decay process produces positrons, which after traveling a short distance (on average less than 1mm) annihilate with randomly encountered electrons producing gamma radiation, which are very high energy photons. During annihilation process there are two photons produced that radiate out in the opposite directions. The scanning technique incorporates two detectors placed on the opposite sides of the body that are used to detect simultaneous occurrence of gamma photons on the two sides (see Figure 3.2). Such detection event implies that the two photons have likely originated from the same annihilation. Since they travel in opposite directions and essentially straight lines, the annihilation point must be located along the same line. Non-simultaneous photon occurrences are rejected as noise. This method has been first used by Brownell and Sweet to localize brain tumors [28]. It is commonly known as Positron Emission Tomography (PET).

A related method known as Single Photon Emission Computed Tomography (SPECT) was developed later in the 60s by Kuhl and Edwards [131]. The method uses isotopes that produce only a single gamma photon. To measure the direction of the incoming photon SPECT uses collimator detector array. Collimators essentially function as a form of optics for radiation that is not refracted by glass (cannot be bent). Depending on the configuration, collimators can act as parallel, converging, diverging, or even as a pinhole optics.

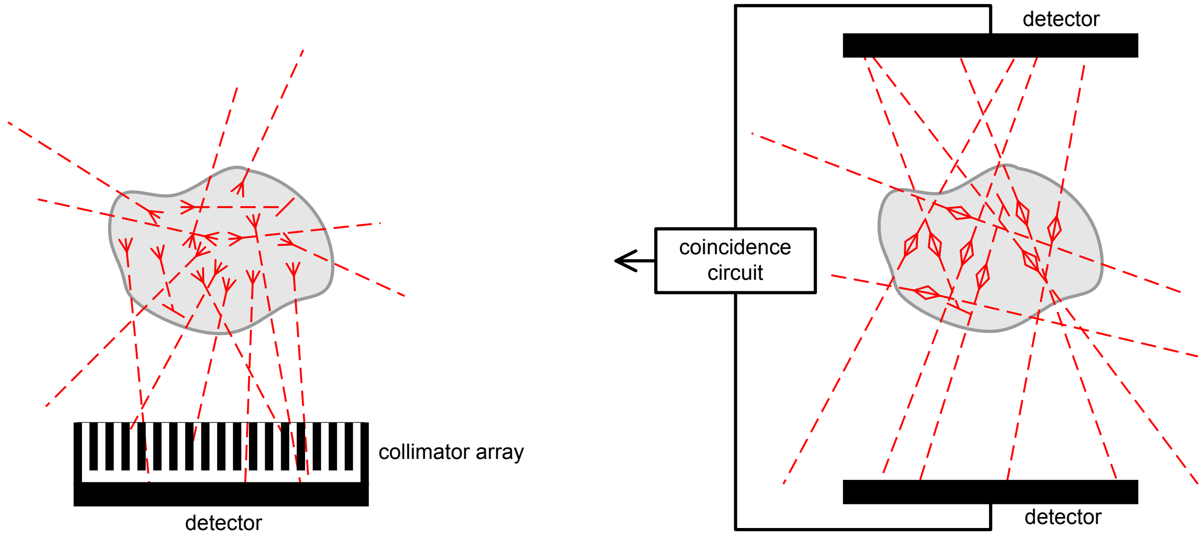

Figure 3.2: Emission tomography. On the left, Single Photon Emission Computed Tomography (SPECT) uses a detector with a collimator array. On the right, Positron Emission Tomography (PET) uses a pair of detectors with coincidence circuitry to detect pairs of photons coming from a single event.

SPECT usually uses parallel collimators, giving a 1:1 magnification. Each detector cell is only activated by photons coming almost perpendicular to it (see Figure 3.2). Given a set of radiation measurements around the body, the inverse Radon transform problem is then solved with iterative or analytical methods. Although the images are usually of lower quality than PET, this scan method is much cheaper.

Although emission tomography is invasive (introduction of a tracer material into the blood stream), it is an indispensable tool in imaging body processes, rather than tissues. Specific tracer compounds are used to observe absorption process by a particular organ. It is widely used in oncology for localizing tumors, cardiology for diagnosing blood flow problems, and in neurology to observe brain activity.

Reflection Tomography

Reflection tomography actually has its roots in ultrasound imaging. It turns out that sound waves can be used in tomography in a similar way as light waves. For years ultrasound tomographic techniques have been used not only in medical fields, but also in geological, oceanographical, and industrial applications. The use of wave echoes to measure distances started with the development of *SONAR* (SOund Navigation And Ranging) and later *RADAR* (RAdio wave Detection And Ranging). The principle of wave ranging relies on detecting and timing the wave reflected off a distant object. If the speed at which the wave travels is known, the distance to the object can be accurately gaged.

The first attempt to apply RADAR as a tool for analysis of interior of objects has been done by Stern in

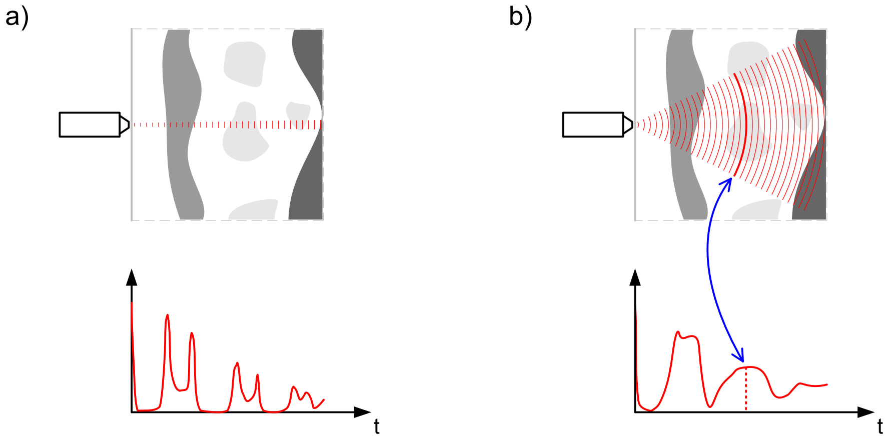

Figure 3.3: Ultrasound imaging versus reflection tomography. In ultrasound (a), reflections are measured along a narrow beam of radiation. The signal measured at a particular time corresponds to the reflection strength at the corresponding depth sampled along a line. In reflection tomography (b), reflections are measured along a spherical wavefront formed by a fan-shaped beam. The signal measured at a particular time is therefore the total sum of all reflections occurring at the corresponding spherical distance from the emitter. Reflection tomography can also employ a parallel wavefront emitter in a similar fashion.

late 20's when he used the technology to measure the thickness of a glacier in Austria [177]. Unfortunately it was largely forgotten until an accident in the late 50's, when several U.S. Air Force planes crashed in Greenland due to their radar penetrating the ice and causing the pilots to overestimate the altitude [177]. This event sparked a renewed interest in subsurface imaging using radio and sound waves, including ice, water, and ground [42, 11]. A system very similar to Stern's was eventually sent as one of the scientific experiments on Apollo 17 to the Moon [215]. First applications of ultrasound also appeared in industry, where it was used to detect flaws in airplane components and laid the foundation of non-destructive testing of materials [57, 71]. Medical uses were investigated in the early 50's, when Howry and Bliss obtained first cross-sectional images of a subject submerged in water [101]. Wild and Reid developed the first ultrasound breast cancer scanner [240]. First ultrasonic recording of the movement of the hart was demonstrated by Edler and Hertz [64].

In the simplest configuration an ultrasonograph works by emitting a pulse along a narrow beam and displays the arrival time and the strength of the detected echo on a temporal axis. The time lag of the echo is proportional to the depth of the reflection. This one-dimensional imaging technique is referred to as the amplitude, or the *A-mode* ultrasonography (see Figure 3.3a). A 2-D image can be obtained with the brightness, or the *B-mode*, where the brightness of each image pixel is proportional to the strength of the reflection and is positioned along the depth (or temporal) axis. The beam is swept in a region of interest along a series of closely-spaced scan lines. The resulting 2-D image is a volume slice along the sweep plane.

The main design feature characteristic of ultrasonographic imaging is volume sampling along a narrow beam of radiation. The reflected echo essentially corresponds to depth measurements along a straight line. It is a major limitation when the aim is to create a 2-D image, or eventually reconstruct a 3-D volume. Not only is the sampling region very small, but also the reflections traveling in all directions interfere with any attempt to make more than one measurement at a time. On top of that, the speed of sound in water (human body is mostly water) is a relatively slow $1.5\,{}^{\mathrm{mm}}/_{\mu\mathrm{s}}$. High-resolution reconstruction using ultrasonography can be a very time-consuming process. Reflection tomography has a potential to accelerate the process of reconstruction by sampling over the entire surface of the wavefront [120]. Instead of a narrow beam, reflection tomography uses a very wide fan-shaped beam. A narrow pulse emitted expands spherically and reflects off any inhomogeneities. The detected echo at time t is therefore the sum of all reflections at a distance of ct, where c is the speed of the wave. The locus of all points at that distance is a hemispherical surface and the detected reflection is therefore an integral over that surface (see Figure 3.3b).

Similarities to transmission tomography are easy to spot. In transmission tomography a measurement is an integral of the attenuation along the direction of the wave, in reflection tomography it is an integral of reflection on the wavefront surface perpendicular to the wave travel direction. By moving the emitter/detector on the surface enough measurements can be made to reconstruct the entire volume is a similar fashion. Alternative designs for reflection tomography can also use other wavefront shapes, including plane and cylindrical wave emitters, but the analysis is essentially the same. This tomographic reconstruction method from reflections was first proposed by Norton and Linzer [172]. Later Dines has shown the feasibility of this technique by reconstructing a volume consisting of needles suspended in water [59]. An in-depth study of ultrasound reflection tomography was performed by Jago and Whittingham [111]. Their clinical evaluation showed noticeably improved resolution over conventional B-mode ultrasonography, although the advantage was not conclusive.

Quite recently Huang et al. [102] demonstrated a new interferometric NIR reflection tomography method called *Optical Coherence Tomography* (OCT). The technique has its roots in *Optical Coherence-Domain Reflectometry* (OCDR), a one-dimensional interferometric range-finding method that was originally used to find faults in fiber-optic network components. It is also very related to optical transmission tomography discussed earlier. Unlike the former, where the transmitted part of the coherent beam is used, OCT makes the use of the portion reflected back toward the source. Initially OCT made its way into medical applications requiring precise anatomical eye measurements [69, 95]. Huang et al. developed the method further to the use for 3-D imaging of semi-transparent tissues. The method is similar to ultrasound imaging, as it also probes the delay on the reflected wave returning from the object. However, an OCT system uses light waves generated by a laser diode and measures the delay with an interferometer.

OCT offers superb resolution (on the order of $10\,\mu$m) when compared to other techniques, however the scanning depth is currently very limited. Scattering and absorption typically put the limits around 2 cm for transparent tissues (for example, an eye) and can rarely exceed a few millimeters in absorbent tissues [202]. Its applications include micro examination of the eye in ophthalmology, blood vessel examination in cardiology, and skin cancer diagnosis in dermatology. It is seen as the tool of the future in microsurgery.

Reflection tomography uses the reflected wave to reconstruct the volumetric structure of an object. Unfortunately reflections take place on all density interfaces, on the way in and out, leading to multiple inter-reflections or *reverberations* within the object. In addition to the undesired reverberations, the wave also refracts when passing from one density region to another - a phenomena known as *phase front aberrations*. Both of these result in noise and reduced resolution [5].

3.1.2 Optical Transmission Tomographic Techniques

The use of non-ionizing form of radiation, such as the near infra-red light (NIR) to examine patients has its obvious health advantages over the more traditional X-rays. Yet, unlike X-rays, the body is only weakly transparent to NIR. Below 600 nm the light is strongly absorbed by the hemoglobin present in the blood, while water absorption limits the penetration depth for wavelengths above $1\,\mu$m [85]. Finally, NIR is a subject to significant scattering causing the internal features to be blurred.

There are currently two different approaches used to overcome the problem of NIR scattering. The first one is based on the observation that photons that take the shortest time to travel through an object take the most direct or straight path. The method therefore aims to capture photons that arrive first at the detector and and filter out the scattered rest. The second approach is based on the notion that an observation of the scattering of light corresponds to a unique volumetric distribution of absorption and scattering. The process of reconstructing the volume is therefore put as an inverse problem of recovering that distribution using an appropriate light transport model.

Direct Imaging Methods

One way to improve the sharpness of a tomographic image is by filtering-out scattered photons. This can be accomplished with a collimator array in front of the detector [113, 110, 241]. This is very similar to the SPECT emission tomography method described earlier (see Figure 3.2). As with SPECT, collimator filtersout any photons not traveling perpendicular to the detector. The assumption is that unscattered photons took the straightest path through the sample. An equivalent method proposed by Wang et al. [235] works by filtering high frequencies in the Fourier spectrum of the spatial distribution of the transmitted light. This leaves only the portion of the transmitted beam arriving at the detector close to the perpendicular direction. Unfortunately collimated filtering only really works for relatively shallow depths (a few millimeters for soft body tissues). As the sample thickness increases the ratio of unscattered to scattered photons rapidly decreases [85].

Another way to distinguish between transmitted and scattered light is to look for light polarization changes. As light is scattered, it looses its initial polarization in collisions eventually leading to a total randomization. This approach has been evaluated by Schmitt et al. [201]. Unfortunately most of the light traveling through a scattering medium eventually looses polarization due to scattering. The experimental limit for biological tissues is about 1 cm, approximately equivalent to a theoretical depth limit of ten transport scattering mean free paths [29].

When light gets scattered, not only does it loose polarization, but also coherence. This can also be exploited to filter-out the undesired scattered portion by intersecting the emerging light with a reference beam [218, 139, 35]. The coherent unscattered portion will interfere with the beam producing a holographic image. This is exactly the same principle used in holographic imaging. The depth resolution in such systems depends on the coherence length of the light source. Since all coherence-based methods still require a portion of the transmitted light to remain coherent with the reference beam it puts a limit on the maximum sample thickness. In addition, the scattered light adds up to the background noise lowering the overall signal-to-noise ratio (SNR). Several improvements to increase the SNR have been proposed and tested, especially based on heterodyne detection [107, 90, 32]. That particular technique measures the amplitude of the beat frequency produced when the coherent portion of the transmitted light interferes with a frequency-modulated reference beam. The amplitude is proportional to the intensity of the transmitted coherent portion. This approach has been quite successful and thickness well over 1 cm have been measured. It is also the basis for the reflective *Optical Coherence Tomography* (OCT), discussed earlier (see Section 3.1.1).

Scattered light can also be directly blocked by a “lightning fast” shutter. Since the non-scattered photons arrive first, a fast shutter could let through the front of a pulse and block the tail. This happens extremely fast (usually on the order of picoseconds) and therefore the use of any mechanical shutter is out of the question. Shutters used in such cases rely instead on certain optical phenomena that can occur with a similar temporal resolution. One such method relies on the *optical Kerr effect* - a property of some isotropic transparent substances (usually liquids or liquid crystals) to become birefringent[2] under a strong coherent light. The shutter that utilizes the Kerr effect is called a Kerr gate and consists of a Kerr cell placed between two blocking polarizers. When a laser is directed at an angle at the cell the induced anisotropy rotates the plane of polarization allowing the blocked signal from the first polarizer to pass through the second one. Essentially the Kerr cell acts as a half-wave plate. With the disappearance of the laser stimulus the gate "closes" within a few picoseconds blocking the signal one again. The Kerr gate shutter method allows to block the trailing part of the scattered pulse letting through only photons traveling the shortest path. The subject is scanned with a series of short laser pulses. The light from each pulse is split, so that part of it is used to "open" the Kerr gate allowing the non-scattered portion of the other part of the pulse to pass through the subject to the detector. This method has been shown to produce quality images for samples up to 1 cm thick [236, 121, 143]. The opacity of a Kerr gate limits the minimum detectable fraction of the transmitted photons, which must be above a few percent of the pulse [85]. An improved design uses a light amplification process based upon the phenomena of *stimulated Raman scattering* (SRS) [153, 186, 13]. Ultimately this approach still relies on the strength of the non-scattered part of the laser beam. There are other alternative filter/amplifier designs exploiting various optical phenomena, such as combining the signal and reference pulses in a non-linear crystal to generate an amplified signal at twice the frequency (known as *second-harmonic generation*, or SHG) [247] and a related method of combining the signal with a reference pulse with twice the frequency in a non-linear crystal to produce an amplified signal at a frequency equal to the difference between the two [67, 239]. The former method is known as *optical parametric amplification* (OPA).

Capturing the whole temporal profile of the transmitted pulse is yet another possibility. This can be done using a *streak camera*. A streak camera essentially is an inverted version of a cathode ray tube: the incoming light is used to generate electrons which are swept in an electric field and impact a phosphor screen producing a temporal image that is recorded in a conventional way with a CCD or a film. Streak cameras have been used to generate sharp 2-D time-gated images [87, 159, 86], and also to generate true tomographic images used to reconstruct the 3-D volume [88, 204]. An alternative to a streak camera is a *time-to-amplitude converter* (TAC). TAC is used to convert a photon arrival delay to a voltage forming a histogram of photon arrival times, equivalent to the temporal dispersion curve. This approach has been investigated by Berg et al. [15] and Benaron et al. [14]. Other promising detection devices include an ultra-fast photodiode [54] and an avalanche photodiode detector [126, 200].

An indirect way to recover the pulse profile is to approach the problem in the frequency domain rather

---

[2] Birefringence is also known as double refraction, or the property of some transparent materials (calcite is a good example), to refract entering light along two distinctive directions, depending on the polarization of the light.

than time domain. Frequency methods use an amplitude-modulated (usually in the radio frequency range) coherent light beam and measure its attenuation and the modulation phase shift. Most of the time the heterodyne detection method is used. A measurement at a single modulation frequency is enough to recover the first moment of the temporal point spread function (TPSF), or the average photon arrival time. Although additional measurements at different frequencies could, in principle, allow to recover the full TPSF, it is usually not necessary. By adapting the Boltzmann transport equation to the amplitude-modulated beam (or the photon density wave) inside the scattering media, Fishkin and Gratton [72] showed how these measurements can be used to recover the approximate absorption and diffusion coefficients essentially providing information about the internal structure. They concluded that achieving desirable resolution would require frequencies on the order of several gigahertz. Unfortunately with the existing technology it is still very difficult to achieve high-frequency modulation at sufficient power levels and current devices are only capable of a few hundreds of megahertz [85]. Working at these lower frequencies still provides enough resolving power to be a viable medical diagnostic tool [176, 114, 249] leading even to first commercial implementations by Carl Zeiss [123], Siemens [203], and Philips [179]. Many recent designs also include several modifications that aim to improve the resolution. One of the more promising relies on amplitude-phase compensation [33]. The method uses a combination of two sources of photon waves with the same amplitude but with a phase shift of 180° and a detector equidistant to the two sources. If the two photon waves pass through a homogeneous medium the interference at the detector would be destructive. Inhomogeneities result in partial cancellation detectable using the standard heterodyne method. Another common modification is to use light emitters operating at different NIR frequencies allowing for faster acquisition time and better tissue discretization. This permits temporal, or functional, tissue analysis [33].

Model-Based Inversion Methods

Inversion methods aim to solve the inverse problem of the photon transform under the assumption that a particular distribution function of absorption and scattering produces a unique distribution of photons exiting the sample. Inversion methods are known under the name of *diffuse optical tomography* (DOT), as the diffusion is explicatively modeled. In a sense the term “tomography” here is quite misleading, as DOT does not use the principle of radiation travel along straight lines. The term is being used mostly for historical reasons. In DOT light scattering is explicitly modeled. A number of calibrated light sources are used to illuminate the sample from different directions and for each direction the distribution of the diffused light is measured using detectors placed on the volume’s surface. These measurements are then used to estimate the likely distribution of absorption and scattering coefficients inside the volume. Since diffuse optical tomography is based on diffusion it will not be discussed here. For a review of DOT refer to the review by Gibson et al. [75].

3.1.3 Magnetic Resonance Imaging

An alternative non-invasive volumetric scanning method called *Magnetic Resonance Imaging* does not use high-energy radiation, but relies on the principle of *nuclear magnetic resonance*. Nuclear magnetic resonance, discovered independently by Bloch et al. [21] and Purcel et al. [183], describes the behavior of protons in a magnetic field. In a strong magnetic field protons align their magnetic moments (or spin) with the field lines (the object becomes magnetized). By the law of momentum conservation, this alignment will undergo *precession* (analogous to a spinning gyro). Although this precession is random, in MRI it is synchronized by applying a perpendicular magnetic field at the precession or *resonance frequency* (usually in the range of 10-100 MHz). After the resonance field is removed, protons will gradually loose their spin synchronization. These changes can be detected and used to infer various spatially-varying properties of the examined volume.

In MRI, a subject is placed in a strong magnetic field that initially magnetically aligns protons. Then a second oscillating magnetic field is applied to induce precession resonance (an RF pulse). The resonating field is then switched off and a sensitive antenna is used to pick up variations in the magnetic field as the protons relax their spins. Varying the length and frequency of the RF pulses allows for measuring of various characteristics.

Typically three quantities are measured: the proton density (PD), the spin-lattice (T1), and the spin-spin (T2) relaxation. PD is a measure of the spatial proton density. It is essentially equivalent to measuring hydrogen density. T1 measures the magnetization relaxation time in the longitudinal direction, while T2 measures the time for the transverse magnetization to decay due to dephasing. T1 time is usually much longer than T2. Unlike CT, in MRI volume is typically described using three parameters.

In order to reconstruct the 3-D volume, the spacial distribution of each parameter must be determined. The methodology used to measure volumetric variation in the MRI signal was proposed by Lauterbur [137]. The crux of the method is the dependence of the proton's precess frequency on the magnetic field strength: creating a smooth gradient in the magnetic field induces a corresponding gradient in precess frequencies. The location dependent read-out takes place by sending a resonance inducing RF-pulse at a particular frequency. Since the precession frequency of the protons was made location dependent by the gradient field, only a narrow "slice" of the volume will match with the RF-pulse frequency and resonate. To determine the 2-D location within the slice, additional two gradient fields are used in the other two perpendicular directions. One encodes the signal in terms of frequency and the other uses phase. The whole 3-D volume can therefore be reconstructed by sending a series of RF-pulses and varying the three gradient fields.

MRI has excellent soft tissue discrimination when compared to other scanning methods. Today it is the primary tool used as a non-invasive method for cancer diagnosis [49]. The main drawback of MRI is the high cost and long acquisition times[3]. Multiple RF-pulses and gradient fields adjustments must be performed in order to reconstruct the entire volume. Strong magnetic fields also prohibit most metallic objects in the scanner.

3.2 Shape Scanning Methods

Although shape scanning could be achieved by most volumetric methods discussed in the previous section, there are many others that were developed specifically for that purpose. Shape scanning methodologies often take a form reminiscent to reflection tomography. In fact, many reflection tomography methods are also used in a simplified form for shape scanning. They are among active scanning methods, where a scanner first emits energy into the environment and then measures the reflection. For example, laser time-of-flight scanners use the principle of timing the reflected signal to determine the depth of the reflection surface. The same principle is used in an A-mode ultrasonography. Shape scanning is also often done passively by detecting the electromagnetic radiation naturally emitted from the environment. This is similar to emission tomography, although emission tomography typically requires a tracer material. Most passive shape scanning methods detect reflected visible light and use digital cameras as detectors[4].

Shape scanning methods are based around the concept that the radiance along a ray in an empty space remains constant. This is called the *ray law* of radiance (see Appendix A.2). Technically it refers to scanning of solid shapes in a vacuum environment. In practice, the meaning of "empty" is not very strict. Although in a non-vacuum environment photons eventually collide with other particles resulting in attenuation, the ray law is often a good approximation over short distances. It is frequently applied to light traveling in clear air. The use of the empty space assumption leads to a significant simplification: all light interactions are limited solely to surfaces of objects. The radiance transport equation in Equation 2.33 reduces to an expression of boundary conditions, which essentially has the form of Equation A.10.

3.2.1 Active Shape Scanning Methods

Active shape scanners emit an energy wave into the environment and infer the geometry from the detected reflection. This methodology is very similar to reflection tomography. Some methods already mentioned

---

[3] There is a derived method called *functional MRI* (fMRI), that can measure relatively short-time changes in the blood oxygen level and is often used to monitor brain activity, but the resolution is much lower.

[4] Infra-red cameras could be used to detect emitted, rather than reflected light.

in Section 3.1.1 can be used directly. These include SONAR and RADAR, although they offer limited maximum resolution due to relatively long wavelengths.

Laser Range Scanning

The development of laser scanners began in the 70s and early 80s, with a lot of pioneering work done by the National Research Council of Canada [188]. For a historical review of the laser scanner technology refer to a review by Blais [20]. Lasers are probably the best tool currently available for measuring distances with great accuracy. Their main advantage over conventional light sources is coherence, both temporal and spacial. These can be used in different ways to construct range scanners, which can be subdivided into three main categories: triangulation, time of flight, and modulation scanners [23].

The main principle behind the triangulation method for laser range scanning is to take advantage of the spatial coherence of the laser beam over long distances. The scanner consists of a laser emitter and a camera placed at an angle [195]. The camera is used to observe the reflection of the laser beam. The position of the reflection in the camera's field of view is used to compute the distance from the reflection surface to the emitter. Triangulation laser scanners can be very accurate, although their range is limited due to geometry. On the other hand, triangulation methods can be extended to also scan the reflectance of a surface in color [9]. By using a multi-wavelength laser source the relative brightness of the reflected beam at each wavelength can be used to estimate the color reflectance of the surface.

Time of flight laser range scanners work by measuring the amount of time it takes for light to travel as it reflects from a surface and returns to the detector [142]. This is essentially the same principle used in SONAR and RADAR. Accordingly, a LIDAR (*Laser Imaging Detection and Ranging*) is a laser version of a RADAR. A time of flight scanner emits a pulse and then measures the time delay t of the reflected pulse. If the speed of light c in the medium is known, the distance d to the reflection surface will be just the half of the distance traveled by the light during that time [165]:

$$d = \frac{c\,t}{2} \tag{3.4}$$

Time-of-filght laser range scanners are typically used for scanning over large distances, as their accuracy is usually lower than triangulation laser scanners over shorter ranges [4]. The accuracy is limited by the ability of the electronic to precisely time the reflection time delay.

Amplitude modulation or phase shift methods are based on the same principle used in optical coherence reflection tomography (see Section 3.1.1). The basic layout of a phase shift scanner is the same as that of a time of flight scanner, as the detector is placed on the same path with the emitter. The laser light signal is modulated with a low frequency signal f_{mod} and then the phase shift $\varDelta f$ of the reflected signal is measured by the detector [223, 117]. The distance to the surface can then be calculated as:

$$d = \frac{1}{2}c\frac{\varDelta f}{2\pi}f_{mod} \tag{3.5}$$

The resolution of an amplitude-modulated scanner is usually much better than that of a typical time of flight scanner [4].

Recently frequency modulated laser scanners have been gaining popularity [219, 30, 58, 166]. They use a source with frequency periodically swept over a small range $\varDelta f_s$. As with amplitude modulation, the reflected signal is then combined with the emitted signal, but this time to measure the frequency shift difference $\varDelta f$. The distance from the detector to the surface is then proportional to the detected frequency shift:

$$d = \frac{1}{2}c\frac{\varDelta f}{f_s}T_s \tag{3.6}$$

T_s is the time period over which the frequency sweep $\varDelta f_s$ occurs. The amplitude of the frequency shift difference signal $\varDelta f$ is proportional to the amplitudes (not the power) of the two combined signals and therefore the dynamic range is about twice as large as for the other time of flight methods [4]. This is the main advantage of frequency modulated scanners which are typically more accurate and can work over longer ranges.

Although a point laser range finder can only detect the distance to one point at a time, it can be modified to scan a range of directions by using a set of computer controllable mirrors that change the direction of the beam. This results in a 2-D depth field measured from the sensor's location. Since the light travels extremely fast, the scanning speed depends on the mechanical components of the mirror array. Typical point scanning speeds are on the order of 10-100 kHz [4].

Structured Light Scanners

While laser range finders can be very accurate, they only allow point sampling of depth and the speed of acquisition is ultimately limited by mechanical components responsible for directing the beam. Overcoming this limitation has led to active scanning methods with consecutive sampling. For example, a laser beam can be split along a line allowing for simultaneous depth sampling along its length. Such solution is used in a slit laser scanner [3]. Early slit scanner designs were using incandescent light [210]. Spatial coherence of lasers over large distances has been the decisive factor in favor of using them in later designs. A slit scanner uses an extension of the triangulation laser scanning technique: a camera placed at an angle measures the reflection of the beam. The displacement of the reflection along a line in the camera view then corresponds to the depth of the reflection point. Depth of the visible scene can be measured by performing a one-dimensional sweep of the laser line across the view.

A step above slit scanners are stripe scanners [112, 6]. By projecting a series of parallel stripes, stripe scanners can estimate depth of an entire scene in a fraction of time required by slit scanners. Going even further, a a dense stripe pattern could be used to capture geometry from just a single image by tracing individual lines simultaneously. A measurement can then be performed instantaneously and no mechanical components are necessary for scanning [20]. Unfortunately additional lines make the line identification algorithm more difficult. While a slit scanner could use a simple threshold, or edge filter to detect the reflection of the line in the image frame, identifying dense pattern features usually requires feature coding and stereographic reconstruction algorithms [162].

In stereo reconstruction scanning is traditionally done passively, as discussed in Section 3.2.2. Yet passive stereo can be difficult to perform under some conditions, such as when a surface lacks features. To aid the stereo algorithm, *active stereo* methods use a projector to project a coded pattern onto a scene. Features of the pattern can then be identified for correspondence matching and in turn used for depth reconstruction. Various active stereo methods were proposed utilizing different projection patterns, including color stripes [25], square grids [238], or other more complex forms [105, 161, 82, 104, 144]. Pattern methods have become quite popular due to their robustness, speed, and relatively low cost. A basic setup requires only a camera and a projector. They are currently one of the favorite methods for geometry scanning in computer graphics. Active stereo methods are mostly limited to Lambertian scenes.

3.2.2 Passive Shape Scanning Methods

Passive shape scanning does not involve emission of a radiation by the scanner, but rather relies on the detection of the radiation already emitted by the scene. Visible light is the most obvious choice. In computer vision, shape scanning uses images taken with a camera. An image is a 2-D slice of the scene's radiance field that captures the interaction between the lighting and geometry. Even though a single image encapsulates the actual geometry of the scene, this problem of shape reconstruction is ill-posed. People are naturally good at solving it (to roughly recreate the geometry using their prior experience), but they can also be easily fooled. For example, we cannot differentiate an image of a scene from an image of a photograph of the same scene.

In computer vision the geometry also cannot be recovered directly from a single image without prior knowledge and limiting assumptions about reflectance. The problem has been extensively explored and the majority of existing techniques use some form of visual cues. Not surprisingly, often these are also the same cues that are used by people. Probably the most popular one is shading and the problem is referred to as *Shape from Shading* [106, 97]. Another possibility is to use texture in a problem known as *Shape from Texture* [242, 221]. Object contours can also serve a a source of geometric information [122]. Feature sharpness due to a lens with a low depth of field can also be used. The problem is generally known as *Depth form Defocus* [65]. Alternatively some of the methods rely on direct human input [99, 229, 212, 45, 256]. Applications of single image techniques are rather limited and will not be discussed here in detail.

Robust geometry recovery usually requires multiple images. In general, there are two possible passive approaches to shape reconstruction from multiple images: stereographic and photometric. The choice dictates how the additional images are captured. In a stereographic approach additional images are captured while changing the camera view, while photometric approach uses variations in the lighting.

Traditional Stereography

The stereographic approach to shape scanning makes use of the view parallax of the scene as the location of the camera changes. Early stereo reconstruction work was conducted in 70's and early 80's (refer to the survey paper by Barnard [10]). Initial work focused on the problem of geometry reconstruction from a pair of stereo images [83] and later was generalized to multi-view stereo reconstruction. The next section gives an overview of different multi-view stereo algorithms used for object reconstruction in object-centric stereo.

In the most basic two-frame stereography, the geometry of a scene is recovered from two images taken by two cameras or the same camera after a translation. The apparent movement of scene features between the two camera views will depend on the distance z of these features from the camera. As the feature point is moved along the x axis of the image plane by Δx, the parallax Δx_p in the image plane for an ideal pinhole camera can be expressed as:

$$\Delta x_p = -f\frac{\Delta x}{z} \tag{3.7}$$

where f is the focal length of the camera. This can also be inverted by moving the camera as shown in Figure 3.4. A feature point P visible in a stereo pair of images created by moving a camera along the x axis of the image plane by b will appear at two different offsets x_1 and x_2 along the *epipolar lines* in the two images. These offsets can be expressed as:

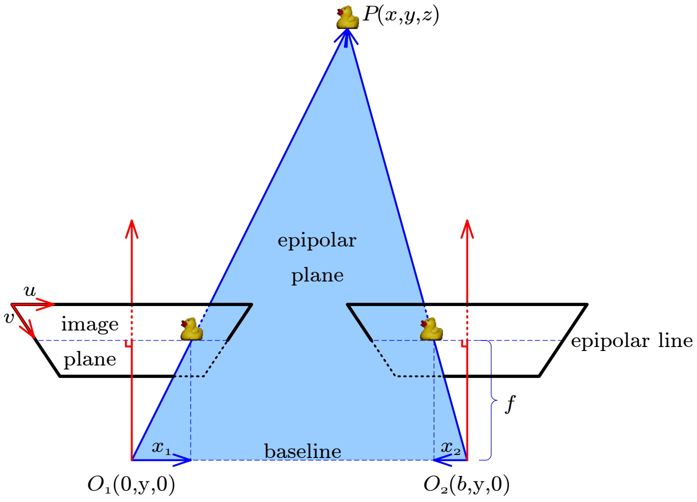

Figure 3.4: Stereo reconstruction. A feature point P visible in a stereo pair of images created by moving a camera along the x axis by b will appear at two different offsets x_1 and x_2 along the *epipolar lines* in the two images. This can be used to recover the depth z.

$$x_1 = f\frac{x}{z} \qquad\qquad x_2 = f\frac{x-b}{z} \tag{3.8}$$

The point correspondence between the two images can be represented with a *disparity map*. Disparity d of a point P is defined to be equal to the difference in the offsets x_1 and x_2:

$$d = x_1 - x_2 = f\frac{b}{z} \tag{3.9}$$

If the baseline distance b and the focal length f is known, the depth z of the feature point in the scene can now be easily computed:

$$z = f\frac{b}{x_1 - x_2} \tag{3.10}$$

The key to all stereographic methods is the ability to correlate points in the scene as the camera view changes. When a camera is moved, this *correspondence problem* is often ambiguous. This is due to several possible situations, such as the lack of texture, occlusions, or view-dependent reflectance changes of the scene. Although a number of possible constraints are often used (for example epipolar geometry, surface smoothness, image brightness constancy), there is no straightforward solution to this problem. A number of possible approaches have been proposed, which can be organized as either local or global correspondence matching [27].

Local matching methods attempt to match image features over small regions. The most popular approach is block-based matching, where the goal is to find the best match for a small region of pixels or a pixel *template* over some search region [7, 68, 253, 17]. Although template matching is relatively cheap to compute and is one of the favorite choices for real-time stereo implementations [27], it does not do very well in the presence of stereo ambiguities. Local matching can also be done on specific geometric features present in images, such as edges [232, 18] or curves [199]. These are generally more robust, but do not produce dense depth maps. When the disparity between images is small, correspondence is often framed as an *optical flow* problem. In a typical situation images come form a slowly moving video camera and the relative movement of feature points from frame to frame can be formulated in terms of differential point motion [98]. The search region is limited to a small window around the original point and the disparity can be estimated by finding a least-squares solution to a system of differential correspondence equations over the window [149].

Global matching methods aim to solve local stereo ambiguities by enforcing global constraints on the visible scene. In general, they are more robust then local methods, although at a higher computational cost. The most common approach is to use *dynamic programming* [43]. For example, a global constraint such as the epipolar ordering and smoothness can be enforced on a local level by finding the minimum cost path through the *disparity space* [173, 44, 19]. Disparity space is a 2-D representation of correspondence cost along two matching epipolar lines. Smoothness constraint is used to find the minimum cost path.

An alternative dynamic programming approach proposed by Tomasi and Manduchi [226] uses *intrinsic curves*, or local feature descriptor vectors at each pixel. These descriptors are created by applying local operators, such as edge, or corner detectors, to each pixel in an image. The idea is that these remain invariant as the camera moves and the correspondence problem turns into a nearest neighbor search in the intrinsic curve space.

Dynamic programming methods still make it difficult to introduce smoothness constraints between epipolar lines. A more recent approach to the stereo correspondence problem is to utilize *graph cuts* [43]. The method has been popularized by Roy and Cox [190]. The idea is to represent image pixels (u, v) and disparities d as a three-dimensional graph along (u, v, d) axes. In the graph, edges have an associated *flow capacity* corresponding to the disparity and smoothness costs of adjacent graph vertices. The *minimum cut* solution produced by the graph cut algorithm is then the minimum cost stereo disparity solution. A recent alternative graph cut representation proposed by Kolmogorov and Zabih [128] uses pixel pair correspondences (rather than pixels themselves) as graph nodes. This way the graph cut cost function can also include a disparity smoothness term and an occlusion penalty term when one of the pixels is occluded. This approach has been shown to be among the best in comparative tests [197]. Other less popularized global matching methods include disparity diffusion [208, 196, 155] and Markov belief network propagation [220], the latter producing results matching in quality the best graph cuts methods [197].

Binocular stereo uses images taken from two overlapping views to recover the depth. However, having just two images is very often insufficient to unabigiously identify disparities. Including additional images taken from different views can help to resolve them. Many stereo systems now often use three cameras instead of just two. In *trinocular stereo* [246, 174] point matching is performed over three overlapping image frames. A single point maps to three pairs of epipolar lines, two for each of the three frames. This results in an additional epipolar constraint, as the disparity point in each image frame must be the intersection of the epipolar line pair. Disparity search is simplified to a simple verification at a precise location in the third image [8]. Of course using more images can result in even higher quality depth maps. Epipolar stereo reconstruction from multiple images is generally known as *multi-baseline* stereo [250, 175, 26]. A video sequence captured with a moving camera can also be framed as a stereographic method. Geometric reconstruction form video is known as *structure from motion* [169, 228] and is usually solved using the optical-flow approach.

Multi-View Stereographic Reconstruction

In a typical stereography camera views are assumed to be close, so that disparities between images can be easily identified. A separate class of stereographic algorithms focuses on the problem of geometric reconstruction of entire objects from many radically different view points. As the camera can be positioned at radically different locations, depth maps are not the best geometric representations. Multi-view stereo object reconstruction methods usually utilize point cloud, polygonal, level set, or voxel representations [205]. As a consequence, disparity cannot be easily used to reconstruct geometry. Instead, most of these methods minimize some form of a global *photo-consistency* cost function [133]. The error is generally computed by comparing existing photographs of an object with identical views of the rendered model.

To reconstruct the shape of an object, a camera is used to capture the object's appearance from many different directions. This actually reassembles computed tomography (see Section 3.1.1). Images obtained by moving the camera are synonymous with attenuation measurements obtained by moving the detector in CT. Unlike tomography, a camera captures the reflected radiation from the object's surface. Some multi-view stereo object reconstruction methods are even simpler and look only at object occlusions of a known background. As only the silhouette of an object is used for its reconstruction, these methods are referred to as *shape from silhouettes* or *shape from contours* [63]. Shape from silhouettes methods use a known background to identify the silhouettes of an opaque foreground object. Silhouettes are then used to carve the volume by back-projection. Although this methodology can be applied to opaque objects with complex reflectance properties, the reconstructed geometry is only an approximation of its convex hull, or the *visual hull* [136]. This limitation can be overcome by examining the light reflected from an object. Such methods explicitly exploit photo-consistency, either by starting with a fully transparent volume and then filling it in a photo-consistent way or by starting with an opaque volume and carving photo-inconsistent parts. The voxel coloring algorithm [206] uses the first approach. Since the appearance of a Lambertian surface observed under fixed illumination is view independent, opacity and color can be iteratively assigned to each voxel so that it is most consistent over all input images. Space carving [192, 133, 132, 216, 254] takes the opposite approach to reconstruction. Voxels with inconsistent appearance are progressively eliminated slowly carving the shape of the object. The process typically requires several passes over the volume, usually as an iterative plane sweep from different directions [133]. There are also other methods that start with an opaque volume yet do not use voxel representations [189, 255, 66]. Since the Lambertian assumption can be quite restrictive, most recent work focused on extending these algorithms to handle more radiometrically complex surfaces [248, 115, 251]. An evaluation of various multi-view stereo reconstruction algorithms was recently done by Seitz et al. [205].

Photometric Methods

Taking the photometric approach requires an analysis of the reflectance properties of the scene. Observed changes in image brightness due to changing lighting are then used to deduce the underlying geometry of the scene. The camera location and parameters are kept fixed, while the scene is observed under a changing illumination.

Photometric stereo can be traced back to the work of Woodham [244], who proposed a method for simultaneous normal and albedo recovery of an observed surface. The radiance reflected by an opaque surface due to the received irradiance can be related using the reflectance equation (see Equation A.10). The most basic setup analyzed by Woodham assumes a known directional light source in a clear medium and ignores reflections and shadowing. The reflectance equation reduces to the product of the BRDF and the irradiance of the observed surface (refer to Equation A.10 for parameter definitions):

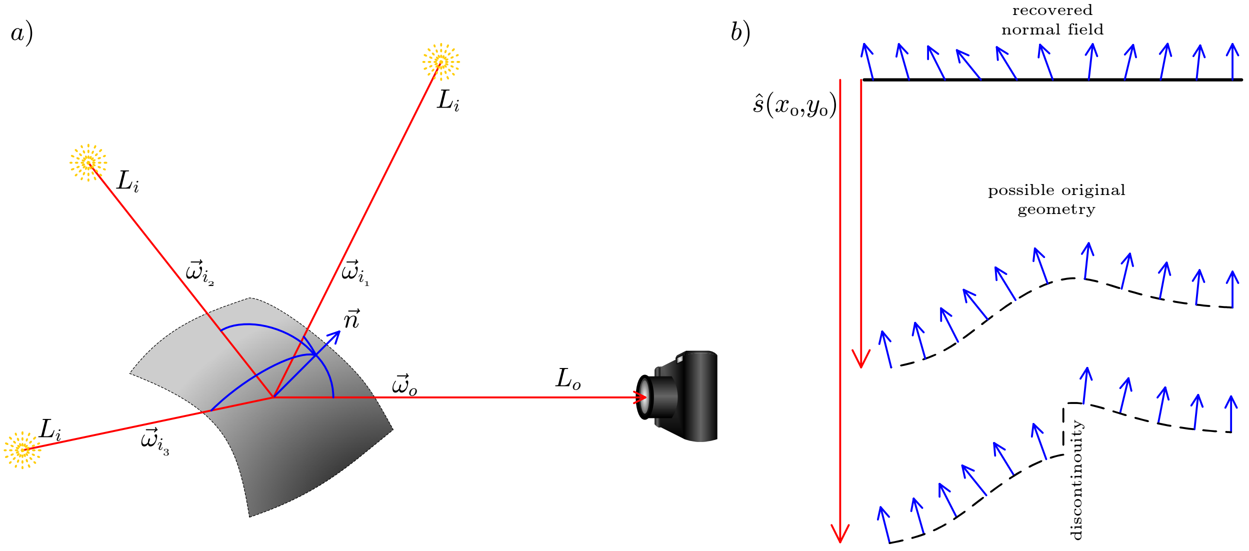

Figure 3.5: Photometric stereo. (a) Reflected radiance L_o measured for a minimum of three distinct light directions is used to recover the orientation of the surface normal vector $\vec{n}$. (b) The shape of the surface is reconstructed by integrating the recovered surface normal vectors over the image plane. The absolute depth of the surface z_0 and shape discontinuities are ambiguous and cannot be reconstructed.

$$\begin{aligned} L(\vec{\omega}_o) =& f(\vec{\omega}_i \rightarrow \vec{\omega}_o) E(\vec{\omega}_i) \\ L(\vec{\omega}_o) =& f(\vec{\omega}_i \rightarrow \vec{\omega}_o) L(\vec{\omega}_i)\, \vec{\omega}_i \cdot \vec{n} \\ L_o =& f(\vec{\omega}_i \rightarrow \vec{\omega}_o) L_i\, \vec{\omega}_i \cdot \vec{n} \end{aligned} \tag{3.11}$$

If the BRDF of the surface is known, the surface normal $\vec{n}$ can be recovered using just two measurements of the reflected radiance for two distinct light directions [96]. This is because $\vec{n}$ has only two degrees of freedom. Unfortunately the BRDF of the observed surface is usually unknown and therefore more images are needed. For example, if the surface is assumed to be Lambertian (see Equation A.13) then the minimum is three images. Silver [214] has shown how to reconstruct the shape of objects with Lambertian reflectance, but unknown albedo. This is the standard setup for photometric stereo methods (see Figure 3.5a). The additional measurement is required to disambiguate the surface albedo (or the diffuse coefficient K_d if we use a radiometrically calibrated setup). The problem can be solved using a simple set of equations:

$$K_d\vec{n} = \frac{\pi}{L_i}\begin{bmatrix} \vec{\omega}_{i_1} \\ \vec{\omega}_{i_2} \\ \vec{\omega}_{i_3} \end{bmatrix}^{-1} \begin{bmatrix} L_{o_1} \\ L_{o_2} \\ L_{o_3} \end{bmatrix} \tag{3.12}$$

K_d is solved for by using the unit vector constraint on the surface normal $\vec{n}$. To reduce errors more than three measurements can be used, so that Equation 3.12 becomes over-constrained. The normal $\vec{n}$ is then found by using a least-squares approach.

Although traditional photometric stereo can only recover the orientation of the surface $\vec{n}$, scene geometry can often be recovered by integrating the normal vector field over the image plane. The depth of a scene $\hat{s}(x, y)$ seen with a camera from a fixed view point is a projection of the geometric shape $s(x, y, z) \mapsto [x, y, \hat{s}(x, y)]$. Under this mapping, the surface tangential and normal vectors can be expressed as:

$$\vec{t}_x = \frac{\frac{\partial s}{\partial x}}{\left\|\frac{\partial s}{\partial x}\right\|} = \frac{\left[1, 0, \frac{\partial \hat{s}}{\partial x}\right]}{\sqrt{1 + \frac{\partial \hat{s}}{\partial x}^2}} \tag{3.13}$$

$$\vec{t}_y = \frac{\frac{\partial s}{\partial y}}{\left\|\frac{\partial s}{\partial y}\right\|} = \frac{\left[0, 1, \frac{\partial \hat{s}}{\partial y}\right]}{\sqrt{1 + \frac{\partial \hat{s}}{\partial y}^2}} \tag{3.14}$$

$$\vec{n} = \vec{t}_x \times \vec{t}_y = \frac{\left[-\frac{\partial \hat{s}}{\partial x}, -\frac{\partial \hat{s}}{\partial y}, 1\right]}{\sqrt{1 + \frac{\partial \hat{s}}{\partial x}^2 + \frac{\partial \hat{s}}{\partial y}^2}} \tag{3.15}$$

Note the relation between the normal vector $\vec{n}$ components and the derivatives of the depth $\hat{s}$. Given a recovered normal vector $\vec{n}$, the depth gradient vector can therefore be written as:

$$\nabla \hat{s} = \left[\frac{\partial \hat{s}}{\partial x}, \frac{\partial \hat{s}}{\partial y}\right] = \left[-\frac{\vec{n}_x}{\vec{n}_z}, -\frac{\vec{n}_y}{\vec{n}_z}\right] \tag{3.16}$$

To recover the geometry, the depth gradient field can be integrated over the image plane. This requires making an additional assumption of surface continuity (see Figure 3.5b). Even with this assumption the gradient field integration can only be used to recover the relative depth, as the absolute depth $\hat{s}(x_0, y_0)$ is unknown. The relative depth can be estimated by integrating the depth gradient field along some path starting at a chosen surface depth point $\hat{s}(x_0, y_0)$. Under the surface continuity assumption the result should be independent of the path, as the gradient field should satisfy the second partial derivatives equality:

$$\frac{\partial}{\partial y}\frac{\partial \hat{s}}{\partial x} = \frac{\partial}{\partial x}\frac{\partial \hat{s}}{\partial y} \tag{3.17}$$

The simplest path that could be used to estimate an unknown depth $\hat{s}(x_1, y_1)$ could start at $\hat{s}(x_0, y_0)$ and integrate the gradient field along the x direction to get $\hat{s}(x_1, y_0)$ and then along the y direction to arrive at $\hat{s}(x_1, y_1)$:

$$\hat{s}(x_1, y_1) = \hat{s}(x_0, y_0) + \int_{x_0}^{x_1} \frac{\partial \hat{s}}{\partial x}(x, y_0)dx + \int_{y_0}^{y_1} \frac{\partial \hat{s}}{\partial y}(x_1, y)dy \tag{3.18}$$

Unfortunately in practice the estimated values of $\nabla \hat{s}$ are noisy and the recovered value for $\hat{s}(x_1, y_1)$ will likely depend on the selected path. A solution to this is to estimate the depth by minimizing a surface fit cost function [96] such as:

$$\int\int_{image} (\frac{\partial \hat{s}_{fit}}{\partial x} - \frac{\partial \hat{s}}{\partial x})^2 + (\frac{\partial \hat{s}_{fit}}{\partial y} - \frac{\partial \hat{s}}{\partial y})^2 \, dxdy \tag{3.19}$$

where $\hat{s}_{fit}(x, y)$ is the surface being fitted. $\hat{s}_{fit}(x, y)$ could be discrete or be represented using basis functions (splines, for example). The optimization is done iteratively until a minimum is reached.

Although the basic photometric stereo was developed for purely Lambertian surfaces, it may often be applied to surfaces with more complex reflectance properties. This may by done if specular highlights and shadows can be treated as outliers. Shafer [207] first observed that for many objects the reflected light can be treated as a linear combination of diffuse Lambertian and specular components. Often the specular component of a surface is nearly zero except for a few sharp highlight areas. As long as there are a minimum of three measurements for every pixel with no highlight or shadow, the surface could be treated as Lambertian, and Lambertian photometric stereo could still be applied. Coleman and Jain [118] first extended photometric stereo to surfaces with simple specular highlights by rejecting measurements with detected specularities. Provided that sampled light directions are far apart with respect to the size of the specular lobe, their method could work using just four measurements. For color images, the separation could be done by performing a chromatic analysis of the reflected light [145, 12]. When the color of the light source is different than the ambient color of the surface, a specular reflection will have a different chromatic spectrum which can be detected. Four-image photometric stereo works well as long as specular highlights are sharp and do not overlap each other in images. The overlap problem is especially difficult for scenes with regions of high curvature.

Rather than rejecting measurements with specular highlights as noise, it is generally better to remove the specular part from an image leaving only the Lambertian component. Shlüns and Witting [198] used sparse set of images to achieve the separation, but required the surface to have homogeneous reflectance parameters. Sato and Ikeuchi [193] also used dense sampling of light directions, but with color histogram analysis at each pixel. Recently Shen et al. [209] used a sparse technique with no homogeneity restriction on the Lambertian component, while Wu and Tang [245] used energy minimization via a graph cut in a very dense photometric stereo setting. Finally Mallick et al. [154] proposed a color space transformation that effectively removes the specular component and can work with as few as three images. For all *dichromatic* color separation methods the major challenge is saturation of specular highlights in a camera leading to a complete loss of color information.

Photometric stereo is not limited in any way to the Lambertian reflectance model. More complex reflectance models allow for reconstruction of a greater variety of surface materials but reconstruction also require many more images to recover the additional parameters. Reflectance is described using a bidirectional reflectance distribution function (BRDF) (see Appendix A.3). Tagare and deFigueriredo [225] developed a theory for photometric stereo for a large class of reflectance models. They also investigated the problem of the minumin number of measurements required to reconstruct the normal field. The theory was later extended by Kay and Caelly [125] who specifically looked at the problem of non-linear parameter recovery. Early experimental work on non-Lambertian reflectance models was done by Nayar et al. [167] who used extended light sources over many directions to explicitly recover the surface and reflectance parameters of a hybrid Torrance-Sparrow and Beckmann-Spizzichino model. In a similar way Solomon and Ikeuchi [217] used a simplified Torrance-Sparrow BRDF model to obtain a measure of surface roughness but required just four images. The Torrance-Sparrow BRDF model with fixed global specular parameters has been used by Georghiades [74] in an uncalibrated photometric stereo setting to reconstruct the surface normal field and unknown light directions. An alternative photometric approach for isotropic BRDFs introduced by Hertzman et al.[94, 93] uses orientation-constancy of surface appearance under the same illumination. Using this principle, a reference object with known geometry and the same surface BRDF can be used to map surface normals to the object of interest, when observed in the same illumination conditions. Several reference objects with different BRDFs can be used to compute a material surface map of objects composed of several materials with different BRDFs. Under a constraint that each object's surface point is a pairwise combination of parametric BRDFs of a small set of materials, Goldman et al.[78] later extended this technique by eliminating the need for reference objects. The number of materials must still be provided by the user. The algorithm recovers BRDF model parameters and material weights for each pixel. The algorithm was demonstrated for the Ward BRDF model.

A completely generic approach that makes no assumptions about the BRDF model has been proposed by Zickler et al.[152, 257, 258]. The algorithm is based on the principle of BRDF reciprocity first described by Helmholtz [233]. This principle of reflectance symmetry about the incoming and the outgoing light directions allows for geometric reconstruction without specific knowledge of the BRDF. The algorithm requires reciprocal positioning of light source and camera pairs and therefore combines some aspects of stereographic and photometric approaches.

Chapter 4

Photometric Ranging of Large-Scale Geometry

Geometric reconstruction of objects with complex surface reflectance properties remains a challenging problem in computer vision. This chapter explores the problem of geometric reconstruction of the shape of objects with arbitrary and possibly anisotropic BRDFs (see Appendix A.3) and introduces a method that can be used for that purpose.

4.1 Introduction

In computer vision, geometric reconstruction uses images as the source of data. With limited knowledge of the scene, recovering of scene's geometry usually requires multiple images. In general, there are two possible approaches to shape reconstruction from images: geometric and photometric. These approaches were discussed in detail in Section 3.2.2. The proposed method can be viewed as photometric, since it uses controlled illumination to recover the geometry of a scene from a fixed view. However, unlike photometric stereo, it recovers the depth of the scene directly. The basic principle is to exploit the law of radiant energy density falloff with distance from a point light source. When a point light source is moved away in free space from a surface, the change in the surface's irradiance is inversely proportional to the square of the distance to the source. This idea was first explored by Iwahori et al. [109], who presented an algorithm to recover the depth and the normal field of a Lambertian scene. They have shown that for a diffusive surface with uniform reflectance this could be accomplished with just four images. Using the same concept, Clark [38] later proposed an active photometric ranging algorithm that is not restricted to Lambertian scenes. His differential algorithm utilizes the fact, that the BRDF gradient component along the light direction vector is always zero, as a BRDF does not depend on the distance to the light source. As this method requires differentiation of the irradiance function, it is highly sensitive to noise. Later Clark and Pekau [39] proposed an alternate integral formulation of Clark's algorithm using a planar light source that is more robust to noise, yet it was only demonstrated to work on Lambertian objects with uniform reflectance. Along with the algorithm of Zickler et al., the method discussed was introduced in [152, 129]. Very recently the principle of using point light sources to recover depth has also been theoretically explored by Mulligan and Brolly [164], who coined the term *photometric ranging*.

The proposed algorithm exploits the same concept of the radiant density falloff from a point light source to recover the scene's depth, but, uses a different approach specifically suited for lightfield rendering with nearby light sources. Using reflectance measurements of coincidental point light source pairs made over a range of light positions, the reflected light fields and be algned to recover the distance to the surface at each pixel location. Moreover, this method is likewise independent of the reflective properties of the surface, and produces high-quality results using real objects.

4.2 Radiance Ratio for Coincidental Illumination

We start by analyzing the basic principle used in photometric ranging. In traditional photometric stereo light sources are intentionally positioned at large distances from an observed surface in order to approximate directional light sources at infinity. On the other hand, photometric ranging uses point light sources at close distances in order to detect irradiance changes caused by varying the separation between the source and the surface. Consider a convex surface illuminated by a point light source and observed with a stationary camera, as shown in Figure 4.1. The illuminator is an ideal point light source with intensity [1] $\mathcal{I}_0$ [W/sr]. Consider the problem of computing the irradiance E_i received by a small surface patch dA at a distance l from the illuminator. If the patch subtends a solid angle $d\boldsymbol{\omega}_i$ with respect to the illuminator, then E_i can be expressed as:

$$E_i = \frac{\mathcal{I}_0 \, d\boldsymbol{\omega}_i}{dA} \tag{4.1}$$

Since dA is a small, differential patch, we can apply a solid angle approximation for $d\boldsymbol{\omega}_i$:

$$\begin{aligned} E_i &= \frac{\mathcal{I}_0 \left(\frac{dA \cos\theta_i}{l^2} \right)}{dA} \\ &= \frac{\mathcal{I}_0}{l^2} \cos\theta_i \end{aligned} \tag{4.2}$$

where θ_i is the angle between the surface normal $\vec{n}$ and the direction to the light source, and l is the distance to the source. Note that the $\cos\theta_i$ factor reflects the foreshortening of the patch with respect to the light

---

[1] Intensity is defined as the measure of radiant power exiting from a point per unit solid angle [77], although it is often casually overused to mean the camera-recorded pixel value related to pixel's irradiance. To differentiate, we will refer to the former as pixel *brightness*.

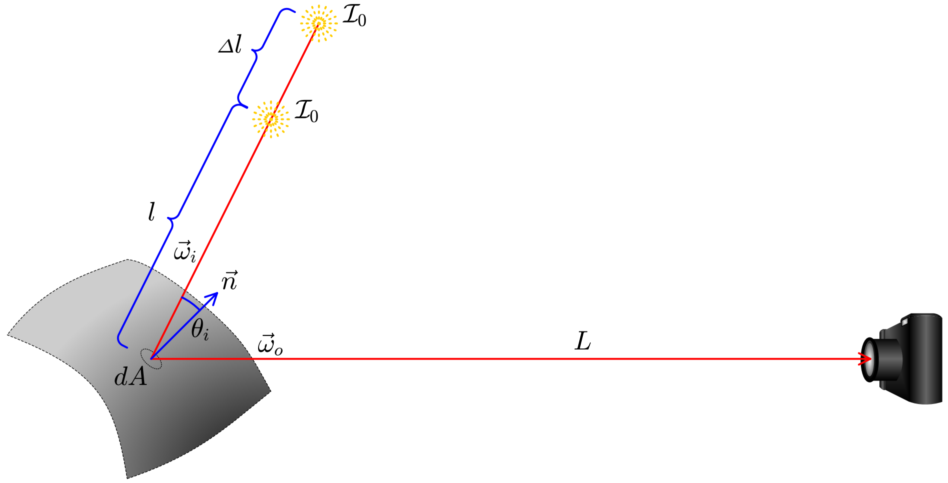

Figure 4.1: Coincidental illumination observed by a camera. A small surface patch dA is illuminated by a point light source with intensity $\mathcal{I}_0$. As the light source is moved along the surface incident direction $\vec{\omega}_i$, the change in observed radiance L depends only on the change in the separation between the surface and the light.

source. The phenomenon described by Equation 4.2 is often referred to as the *inverse-square law*.

Some of the light reaching the patch is subsequently reflected toward the camera. The outgoing radiance along the solid angle vector $\vec{\omega}_o$ is related to the incoming irradiance E_i through the BRDF of the surface. In general, the BRDF can depend on several other parameters, such as wavelength, polarization, etc. For simplicity, we will only consider the BRDF's directional dependence for now. The outgoing radiance is $f_r(\vec{\omega}_i \rightarrow \vec{\omega}_o)E_i$, where f_r is the BRDF of the surface for the patch dA. Combining these, we can express the outgoing radiance in terms of the original light source intensity $\mathcal{I}_0$. For a light source at a distance l from dA, the outgoing radiance L_1 is:

$$L_1 = f_r(\vec{\omega}_i \rightarrow \vec{\omega}_o)\frac{\mathcal{I}_0}{l^2}\cos(\theta_i) \tag{4.3}$$

Now imagine moving the point light source back by a distance Δl along $\vec{\omega}_i$ to a new position (see Figure 4.1). The irradiance received by the surface patch will decrease due to the increased separation:

$$L_2 = f_r(\vec{\omega}_i \rightarrow \vec{\omega}_o)\frac{\mathcal{I}_0}{(l+\Delta l)^2}\cos(\theta_i) \tag{4.4}$$

Since the camera remains stationary, the camera vector $\vec{\omega}_o$ is the same in both cases. We will combine the BRDF and the foreshortening term into a two-dimensional camera-oriented reflectance function of the incoming light direction:

$$f_{r\to c}(\vec{\omega}_i) = f_r(\vec{\omega}_i \to \vec{\omega}_o)\cos(\theta_i) \tag{4.5}$$

When a point light source is moved along the incident direction $\vec{\omega}_i$, this reflectance function remains constant. This is because the BRDF is a function of the incoming and outgoing light directions, which do not change in this case. We can therefore relate Equation 4.3 and Equation 4.4 through the reflectance function:

$$f_{r\to c}(\vec{\omega}_i)\mathcal{I}_0 = L_1 l^2 = L_2(l + \vartriangle l)^2 \tag{4.6}$$

Note, that we have effectively eliminated the reflectance term and in turn the BRDF. The result is a simple formula that relates the unknown distance l to the measured irradiance values L_1, L_2 and the known offset $\vartriangle l$:

$$L_1 l^2 = L_2(l + \vartriangle l)^2 \tag{4.7}$$

It may appear, that Equation 4.7 could be directly used to recover the distance to the surface patch, by solving it for l. Unfortunately it is not possible in this basic setup to correlate the location of the surface patch dA with the corresponding point on the camera's image plane. The location of the surface patch in space can only be determined given measurements of reflected radiance L_1 and L_2, yet we cannot determine the direction $\vec{\omega}_i$ to obtain the second radiance measurement L_2 without knowing the location of dA. In the next section we will present an approach that solves this problem by double-covering the set of incident directions and then uses a global optimization strategy to estimate the scene's depth.

4.2.1 Depth Estimation using Dense Sampling

Even though it is not possible to determine the depth directly from a single pair of measurements, the problem can be solved using a global approach. As shown in Figure 4.2, let us define a global coordinate system $\{\mathcal{O}\}$, whose origin O will be conveniently located near the center of the object. Imagine, that a point light source is moved over a convex surface around the object while we gather images capturing the reflected radiance. Let the convex surface be defined by a radial distance function $l(\phi, \theta)$ relative to $\{\mathcal{O}\}$. $L(\phi, \theta; u, v)$ will be the radiance function measured at an image point (u, v), as the light source is moved over $l(\phi, \theta)$. Note, that L is measured for a single, fixed viewing direction, and (ϕ, θ) are parameters of the direction of the light source. The measured radiance will be a function of the surface's BRDF, foreshortening, source's

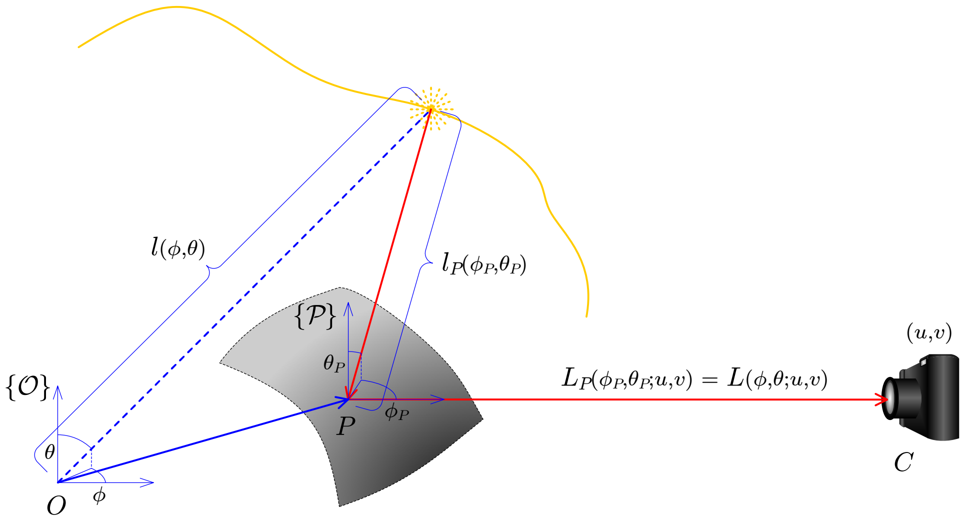

Figure 4.2: Mapping between the radiance field $L(\phi, \theta; u, v)$ in global coordinates $\{\mathcal{O}\}$ and the radiance field $L_P(\phi_P, \theta_P; u, v)$ in local coordinates $\{\mathcal{P}\}$ at a point P.

intensity, and the distance to the source, as defined earlier in Equation 4.3. An example of a radiance function $L(\phi, \theta; u, v)$ for a particular pixel (u, v) is shown in Figure 4.3. Measurements were done by moving a point light source on a spherical surface and then repeated for a different sphere radius.

For some point P along $\vec{\omega}_o$, also mapping to the point (u, v) on the camera plane, we can define a similar local radiance and radial distance functions $L_P(\phi_P, \theta_P; u, v)$ and $l_P(\phi_P, \theta_P)$. For that we define a local coordinate system $\{\mathcal{P}\}$ with P as the new origin whose axis are aligned with those of $\{\mathcal{O}\}$. Note, that we can remap the original functions into the $\{\mathcal{P}\}$ frame of reference. Let $\mathcal{W}_P$ define a 2-D change of coordinates that transforms the (ϕ, θ) into local coordinates (ϕ_P, θ_P):

$$(\phi_P, \theta_P) = \mathcal{W}_P(\phi, \theta) \tag{4.8}$$

While the radiance function is remapped, the distance function must additionally be offset by a vector from the origin O to the point P. This can be written as:

$$L_P(\phi_P, \theta_P; u, v) = L(\phi, \theta; u, v) \tag{4.9}$$

$$l_P(\phi_P, \theta_P) = \left\| l(\phi, \theta) \begin{bmatrix} \sin\phi \sin\theta \\ \cos\phi \sin\theta \\ \cos\theta \end{bmatrix} - \vec{P} \right\| \tag{4.10}$$

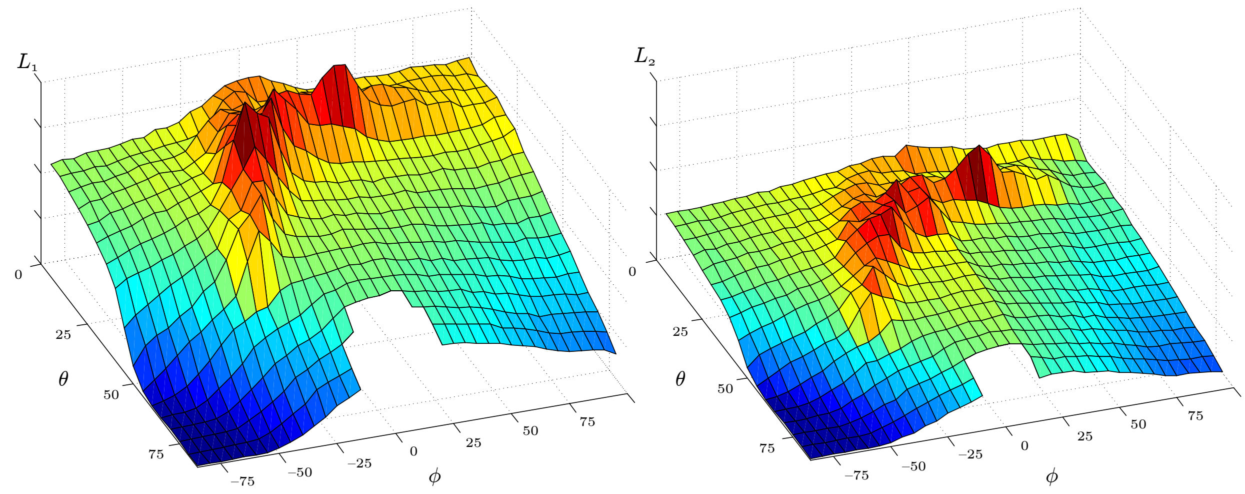

Figure 4.3: An example of a complex radiance function measured at a single pixel location as a point light source is moved on a spherical surface in the (ϕ, θ) space. The figure shows the green channel measurements for motion over two spheres of different radii. Missing data at the bottom of each graph is due to the robot arm occlusions.

Suppose, that we have multiple measurements obtained by moving a point light source over two non-intersecting surfaces described by two distance functions l_1 and l_2, resulting in two radiance functions L_1 and L_2. Projecting both into the local coordinates of the surface point P, we obtain a double-covered local radiance function (Figure 4.4). For every direction (ϕ_P, θ_P) we have two radiance measurements corresponding to point sources at two distances from P. Note that this is very similar to the situation described by Equation 4.6. For any light source direction (ϕ_P, θ_P) we can identify a pair of collinear light source locations on l_1 and l_2. Analogous to Equation 4.7, we can therefore form an equation relating the local radiance and light source distance functions independent of the local BRDF:

$$L_1(\phi_P, \theta_P; u, v) l_1(\phi_P, \theta_P)^2 = L_2(\phi_P, \theta_P; u, v) l_2(\phi_P, \theta_P)^2 \tag{4.11}$$

Effectively we have eliminated the unknown surface BRDF $f_r(\vec{\omega}_i \to \vec{\omega}_o)$, the foreshortening term $\cos(\theta_i)$, and the unknown source intensity $\mathcal{I}_0$. The location of the surface point P is still unknown, even though this location is necessary to define the mapping $\mathcal{W}_P$. However, now radiance is defined as a function of (ϕ_P, θ_P) and, using the $\mathcal{W}_P$ operator, it can be re-mapped for any point P.

Let $s(u, v)$ be the depth of the scene for a camera pixel (u, v). That is, $s(u, v)$ is the distance to a point P from the camera center C along the camera vector $\vec{\omega}_o$ intersecting the camera plane at a point (u, v). We can express the point P in terms of s as:

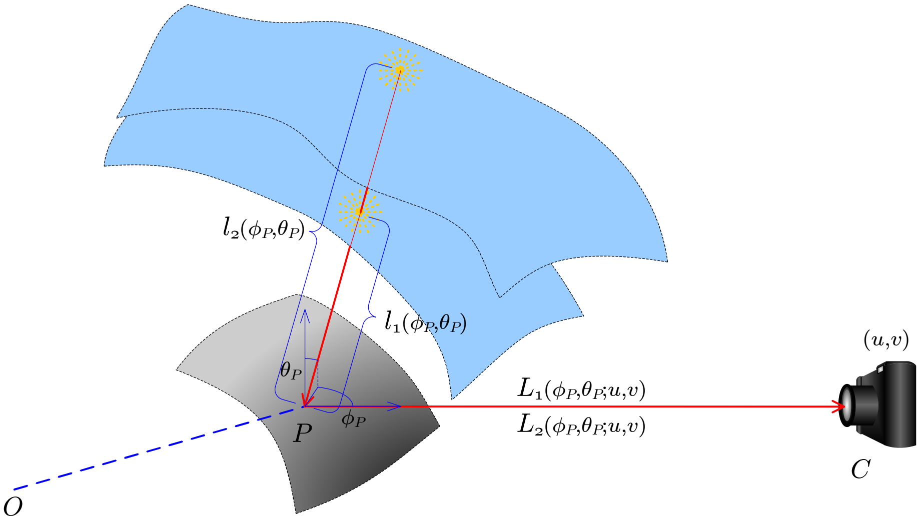

Figure 4.4: Radiance functions L_1 and L_2 in the local reference frame of the point P. For some direction (ϕ_P, θ_P) the distances to the two point source locations are given by l_1 and l_2.

$$P = C - s\,\vec{\omega}_o \tag{4.12}$$

If we knew the true surface depth $\hat{s}$, then under ideal conditions we would expect Equation 4.11 to be satisfied for every light direction (ϕ_P, θ_P). For the true depth the two re-mapped radiance functions L_1 and L_2 would be in alignment, differing only be the ratio of the squares of the light source distances l_1 and l_2. However, in general, Equation 4.11 will not be satisfied for an arbitrarily chosen value of s. We can define an error function over the depth s along the camera vector $\vec{\omega}_o$ as:

$$\mathcal{E}(s; u, v) = \int_\Phi \int_\Theta \left\| L_1(\phi_s, \theta_s; u, v){l_1}^2(\phi_s, \theta_s) - L_2(\phi_s, \theta_s; u, v){l_2}^2(\phi_s, \theta_s) \right\| d\theta_s d\phi_s \tag{4.13}$$

Appropriately we have replaced location subscripts P with depth s. Our aim is to find the minimum of Equation 4.13 over some likely range of depth values. This minimum corresponds to the best depth estimate $\hat{s}(u, v)$:

$$\hat{s}(u, v) = \underset{s}{\operatorname{argmin}} \int_\Phi \int_\Theta \left\| L_1(\phi_s, \theta_s; u, v){l_1}^2(\phi_s, \theta_s) - L_2(\phi_s, \theta_s; u, v){l_2}^2(\phi_s, \theta_s) \right\| d\theta_s d\phi_s \tag{4.14}$$

It should be noted that for each image point (u, v), the depth $\hat{s}$ is estimated independently.

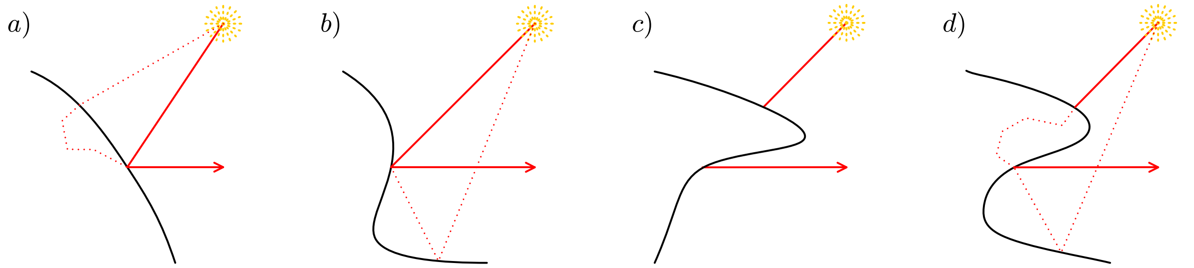

Figure 4.5: Physical violations of the direct illumination assumption: (a) subsurface backscattering, (b) surface interreflection, (c) shadowing, (d) combination of all three.

4.3 An Implementation of Photometric Ranging

The objective function in Equation 4.14 is non-linear and multimodal, and finding its minimum requires a global iterative approach. Ideally, the minimum value corresponding to the true depth would be equal to zero, but in practice, there may be several local minima, and the global minimum may likely be greater than zero due to data noise. In addition, the direct illumination assumption by a point light source can sometimes be violated due to shadows, interreflections, or sub-surface scattering (see Figure 4.5). Shadows violate the direct illumination assumption as the direct light path is blocked by geometry. However, measurements affected by shadows can be easily determined and discarded through brightness thresholding. This may be more difficult if the other two phenomena are present. Interreflections and sub-surface scattering are more difficult to detect and account for. One possible solution is to use the direct illumination separation method proposed by Nayar et al. [168]. Still, for many opaque objects, sub-surface scattering is minor and interreflections are not significant if the surface is mostly convex or the surface albedo is low.

4.3.1 Finding the Minimum

In order to find the global minimum of Equation 4.14, it is necessary to resample radiance functions L_1 and L_2 into the local coordinates of $\{\mathcal{P}\}$. Since radiance functions were discretely sampled, evaluating them at arbitrary values of (ϕ, θ) requires interpolation. Mapping between the global (ϕ, θ) and the local (ϕ_P, θ_P) space, which requires resampling, may therefore result in sampling artifacts if the sampling rate of the global radiance is inadequate. This is a potential problem if the radiance function has high-frequency components, for example, high-frequency BRDF components or shadow boundaries. Although a denser set of measurements would be preferred, we can alternatively apply a low-pass filter. In experiments low-pass filtering had few negative side effects on the reconstruction. As a second benefit, filtering also removes high-frequency noise. We have achieved the best filtering results by first applying a Savitzky-Goley polynomial smoothing filter [194] and subsequently performing bicubic re-sampling during the $\mathcal{W}_P$ mapping operation. We then approximate integrals from Equation 4.14 over the (Φ, Θ) range by summation at discrete points.

To find the global minimum corresponding to the true depth, we can use a number of recursive non-linear minimization algorithms. Experiments discussed in this chapter use a variation of the convex global underestimator method [180]. The method recursively uses a quadratic underestimate of the overall shape of the objective function to find the global minima through direct sampling. In the final iteration uses the minimum from the quadratic estimator, as it turned out to be better in the presence of noise.

4.3.2 Light Geometry Cofiguration

To estimate the depth of a scene point, it must be directly illuminated from a 2-D subset for two surfaces of light positions. The simplest configuration of the source surfaces are two spheres, one enclosing the other. The sampling should be sufficiently dense to guarantee that all visible parts of a scene are under direct illumination in some image. This is the configuration that we have used in our experiments (see Figure 4.9).

Note, that in Equation 4.14 the error is minimized over the depth s, which is measured along the camera direction $\vec{\omega}_o$ rather than the light direction $\vec{\omega}_i$. When the light ray is perpendicular to the camera ray, the gradient of the error function in Equation 4.13 with respect to s is zero. In fact, the gradient is always zero for the component of the light source vector orthogonal to the camera vector. This may cause difficulties in localizing the minimum when the data is noisy. To account for that, more weight can be placed on samples where the light direction is closer to the camera direction by weighting each sample with the dot product: $\|\vec{\omega}_i \cdot \vec{\omega}_o\|$. This weighting is applied to the objective function from Equation 4.14.

4.3.3 Radiometric Calibration of the Light Source

Unlike other photometric techniques that assume that the light source is at infinity, photometric ranging uses a point light source at close range. This poses some difficulties, as effects of a real, non-ideal light source become more pronounced at finite distances. These include the source's finite size and possibly its anisotropic angular radiation pattern.

Radiant energy density due to a finite size source decreases with distance at a lower rate than for a point

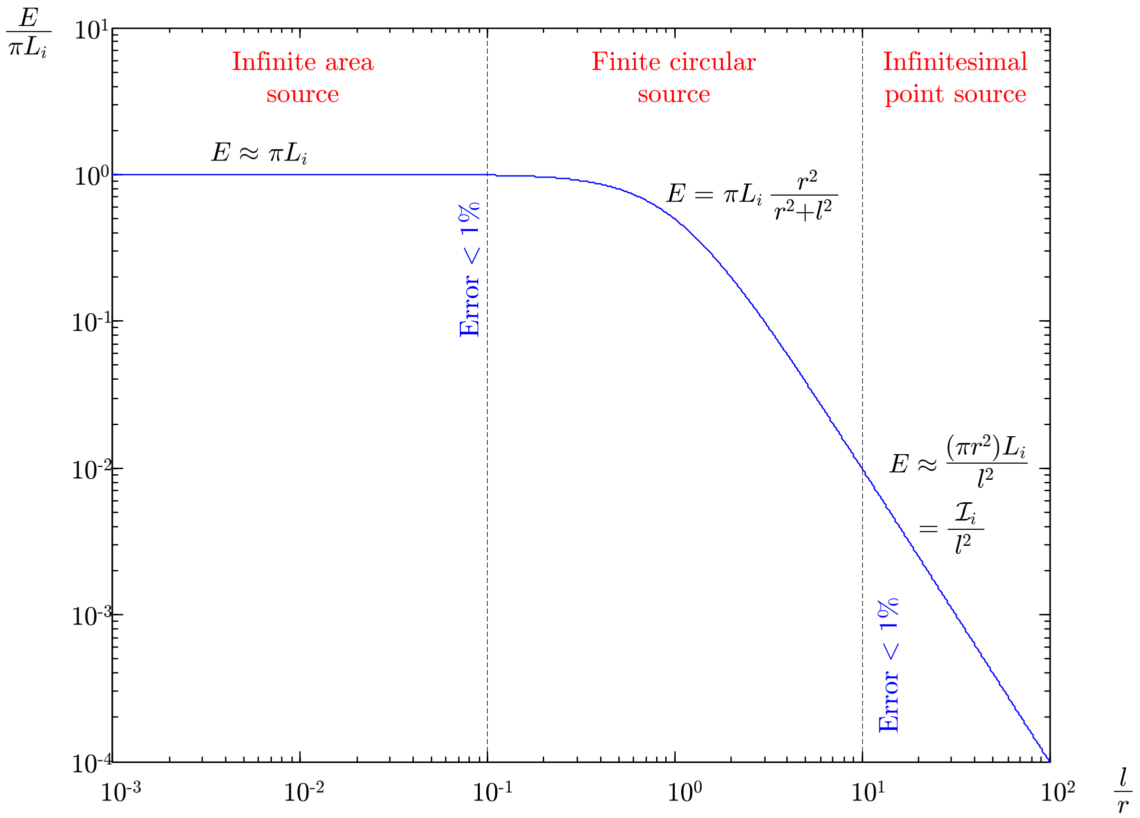

Figure 4.6: Theoretical surface irradiance falloff behavior when lit from a distance l by a uniform circular light source with a radius r [191].

source. Fortunately, we can minimize the impact of the finite size by moving the source "far enough" from the surface. With increasing distance, the energy density falloff approaches that of an ideal point source. In Figure 4.6 a graph shows the theoretical power density falloff from a circular light source of radius r versus the distance l from the surface. At distances less than 0.1 of the radius, a circular light source can be treated as an infinite plane source, and the observed irradiance does not change much with distance. At greater distances the physical dimensions of the source become significant. Yet at distances above 10 times the radius, a circular light source can essentially be approximated as a point light source. For example, to achieve the goal of approximating the radiant density falloff to within 1%, the light must be kept at a distance over 10 times the light's diameter. For necessary derivations, refer to the book by Ryer [191]. At the same time it is important to realize, that the signal-to-noise ratio of the reconstuction is dictated by the proximity of the light source, and therefore the size of the light source should still be as small as possible, so that the minimum acceptable distance is small.

In experiments we have been using a compact super-bright white light emitting diode (LED) OTLWHTA-

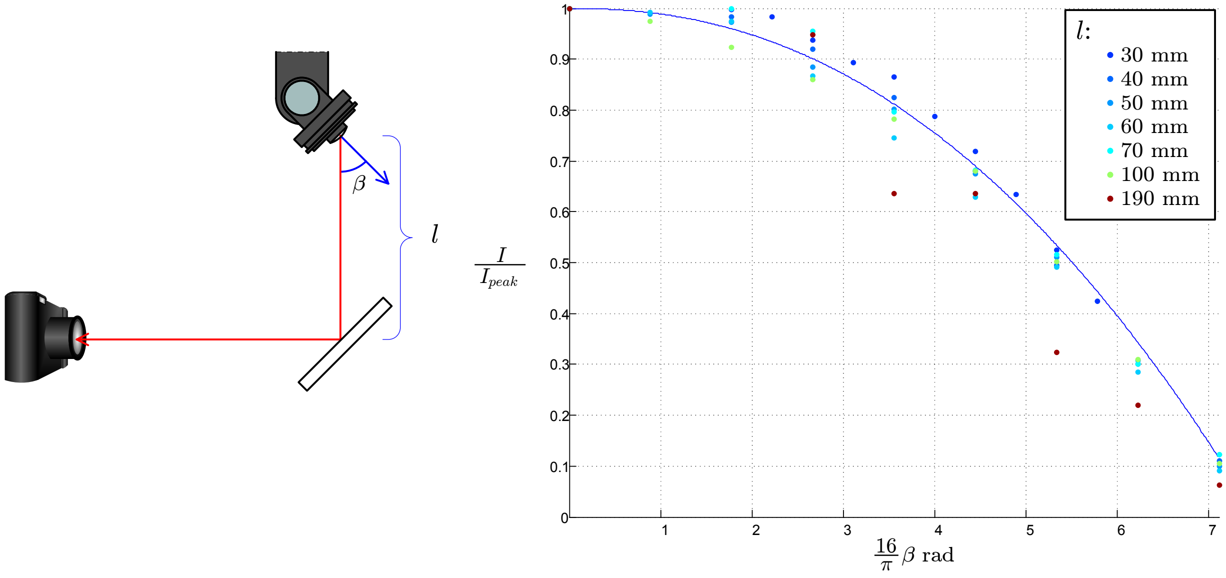

Figure 4.7: Radiation pattern measurement for our LED. A camera is used to measure changes in reflected radiance from the center of a diffuse surface as a light source is rotated. Measurements are repeated for a range of distances from the surface.

5-10-66-E [178]. With a radius of 0.5 cm, the minimum distance for the 1% point source approximation was 5 cm. In the experimental setup (Figure 4.9) the diode was always kept at a radius of more than 14 cm (the maximum radius was limited by the mechanical constraints of the robot to 19 cm). In accordance with the 1% approximation rule, it provided a working space with a radius of about 9 cm.

All real light sources typically also have a non-uniform angular power radiation pattern caused by the physical shape of the light emitting area. For example, a typical LED has a peak power output in a forward-facing direction and quickly decreases away from it. To keep the source power constant, the light source should ideally always be pointed toward the examined surface. However, since the the surface depth is unknown a-priori, the best we can do is to aim the light toward the origin O. The resulting power difference can be compensated for by measuring the angular power distribution pattern of the light source. We have measured it by taking a series of observations of a lit tilted diffuse plane. While the light source was rotated about its center point, a camera was used to measure the brightness at the center the diffuse plane (see Figure 4.7). Measurements were repeated for a range of distances to the surface and a least-squares polynomial fit was then used on the measured data. Given the surface light vector $\vec{\omega}_i$ and the actual forward direction of the light source aimed at the origin O, the resulting function can then used to correct the expected measured irradiance.

The last issue to consider is the temporal stability of the illumination source. This includes both the light's temperature and the power output. These can be easily assured through selection of a stable power supply and a source with a short warm-up time.

4.3.4 Calibration of the Camera and Noise Control

For the method to be accurate, we must be able to detect small changes in irradiance due to variations in light source distance. In essence, the algorithm uses ratios of irradiance measured for two different source distances to the surface. Equation 4.7 can be re-written as:

$$R_L = \frac{L_1}{L_2} = \left(\frac{l + \Delta l}{l}\right)^2 \tag{4.15}$$

A question one might ask is how much can we expect the ratio R_L to change, as the distance l changes. This indicates how sensitive the camera needs to be in order to detect changes in l. Let us assume, that our distance l is large enough, so that the inaccuracy in light source positions can be ignored as a source of error. Differentiating Equation 4.15 with respect to l, we get:

$$\frac{dR_L}{dl} = -2\frac{\Delta l\,(l + \Delta l)}{l^3} \tag{4.16}$$

To obtain a relative measure, we divide both sides by R_L and rearrange the equation a bit to obtain:

$$\frac{dR_L}{R_L} = -2\frac{\Delta l}{l + \Delta l}\frac{dl}{l} \tag{4.17}$$

Recall that Δl is the separation between the two light sources (see Figure 4.1). The observed relative change in the radiance ratio R_L will therefore be the greatest when the separation between light sources is maximized and will be equal to at most twice the relative change in the distance l. This is the fundamental limitation of this method for a single measurement and is equivalent to a theoretical observation when the light source is placed as close as possible to the surface and a second observation with the light source placed at infinity. In practice, it is difficult to achieve a value greater than one times the relative change in the distance l.

Unfortunately measuring small percentage changes in irradiance with a typical digital camera can be difficult. The biggest challenge in using digital cameras is limited sensitivity. This is mostly due to the sensor noise inevitably present in any measurement of the light reflected off the scene. The response of the camera can also be non-linear. Since the algorithm requires linear radiance measurements, the radiance response curve must be recovered in a calibration procedure. Refer to Appendix C for more details on the procedure.

In the algorithm, the location of the camera and its optical characteristics (focal length, exposure, etc.)

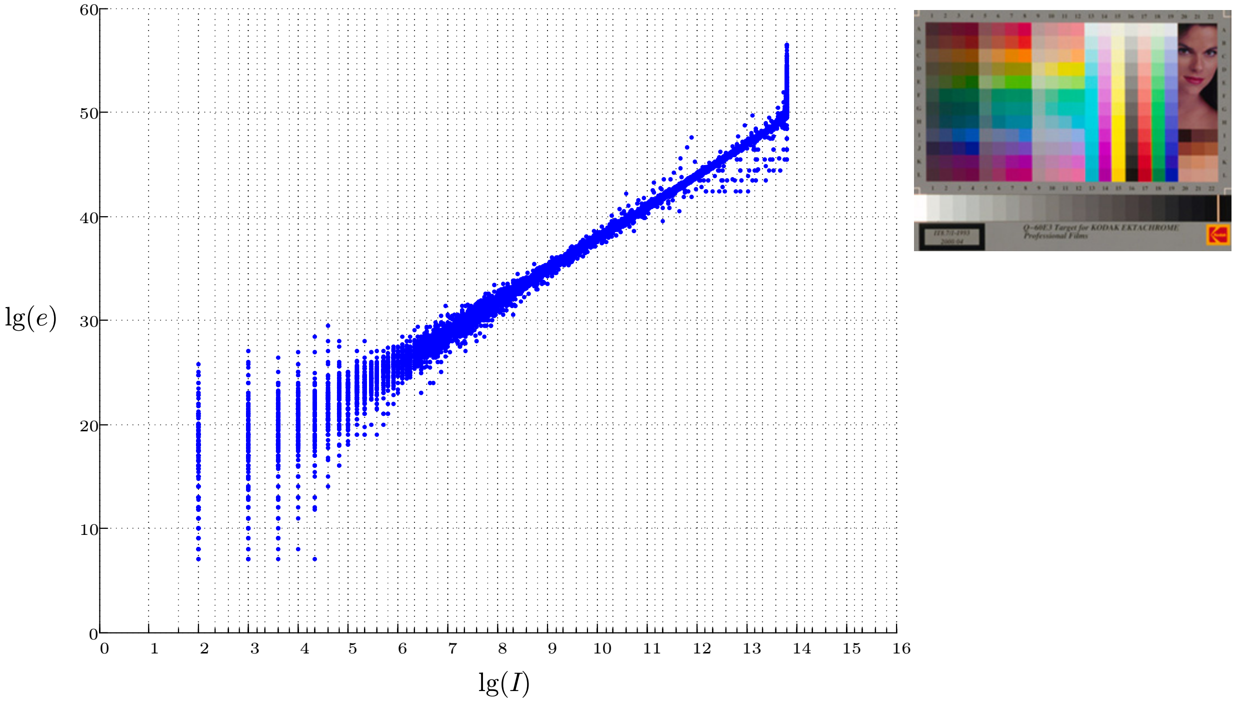

Figure 4.8: Response plot of the Cannon EOS-1D camera in the green channel. The plot has been created using fifteen images of the Q-60 calibration target (shown on the right) with shutter speeds varying from 1 *sec* to $^1/_{25}$ *sec*. The logarithm of the exposure e is shown versus the logarithm of the camera-recorded pixel brightness I.

are fixed. As the light source is moved, changes in measured irradiance then remain directly proportional to changes in the incoming radiance. In experiments we have been using the Canon EOS-1D camera [31]. The radiance response function has been sampled by taking a series of images of the Kodak Q-60 calibration target [127] at varying exposure levels. Images were captured using the camera's RAW mode that preserves the "raw" brightness values captured by the sensor with no electronic processing that would typically follow. Figure 4.8 shows the distribution of the measured values forming the camera's response. The best-fit response function can be computed by minimizing the objective function from Equation C.5. The distribution of data points for this camera suggests a nearly linear response. For small exposure values the response is very noisy due to random noise of the camera. For large exposures the sensor becomes saturated. Only the relative values on the vertical axis are meaningful (refer to [56]). Note, that the range of camera brightness values is from 2^2 to 2^{14} (EOS-1D is a 12-bit camera with a 2-bit offset in RAW mode).

Any irradiance readings recovered using the inverse transfer function will always be corrupted by the camera noise. In general, there are two main types of noise in a digital sensor: the pattern noise and the random noise (see Appendix C.2). The pattern noise refers to the spacial bias and variation in the photo response over the sensor's surface. It can be effectively eliminated by subtracting a bias dark frame captured at the same exposure settings. In experiments we have recorded a dark frame after each data acquisition session. The random noise is generally the result of random quantum processes and is influenced by temperature. It becomes especially problematic at low exposure levels, as demonstrated in Figure 4.8. Averaging several frames is one way to reduce it, although it appropriately lengthens the acquisition time. An alternative is to perform spatial averaging of pixels. The obvious side effect is a reduction in spatial resolution. This may be an acceptable sacrifice, especially when using high-resolution cameras.

Finally, as with all photometric methods, special care must be taken to keep the camera pixel-registered. Recall, that relatively small differences in recorded irradiance have direct effect on estimated depth. At the same time, shifting the camera view just slightly on a sub-pixel level can have a drastic effect when a pixel's extent crosses a brightness discontinuity, either a physical edge, or a textural discontinuity. For scenes with many brightness discontinuities, sub-pixel registration is essential and it can be very challenging for a high-resolution camera. Spatial averaging of pixels can also help in this case.

4.4 Experimental Setup

The experimental setup is shown in Figure 4.9. It used a digital camera to take a series of images of objects placed on a static stand under controlled illumination. A white LED was used as a light source and moved

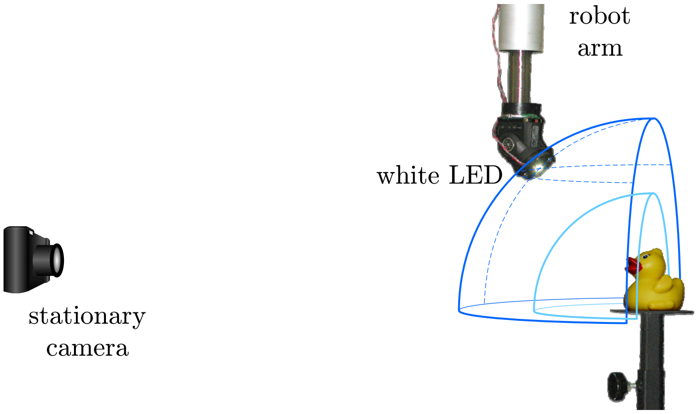

Figure 4.9: Experimental setup diagram.

using a robot. The effective diameter of the LED used in experiments was approximately 10 mm. Physical workspace limitations of the robot allowed the LED to be positioned at a maximum radius of only 190 mm. In order to comply with the point-source approximation condition (Section 4.3.3), this limited the minimum distance from the object's surface to the light source to 90 mm. The radial pattern of the LED illuminator radiation was also non-uniform, with intensity falling off away from the central axis (see Figure 4.7). This was accouted for using the procedure from Section 4.3.3. An AdeptOne-MV robot [2] was used to move the light source over two spherical surfaces with irradiance sampled uniformly in the (ϕ, θ) space. Only the front upper quarter spheres of light directions were sampled (see Figure 4.9).

Experiments used a low-noise 12-bit Canon EOS-1D camera and all images were captured in RAW camera mode with fixed exposure. The camera was calibrated using the procedure from Section 4.3.4. A mosaic of a full set of images is shown in Figure 4.10. Of note are robot arm occlusions for light source locations near the camera view. Since unoccluded parts of these images contained useful data, occluded parts were masked-out during processing rather than discarding these images completely.

4.5 Depth Reconstruction Results

Initial evaluation evaluation of the algorithm was done on objects with known geometry and simple reflectance in order to compare it to existing photometric methods. Wooden blocks covered with white paper (nearly Lambertian) were used as models for that purpose. Two reconstruction examples are shown in

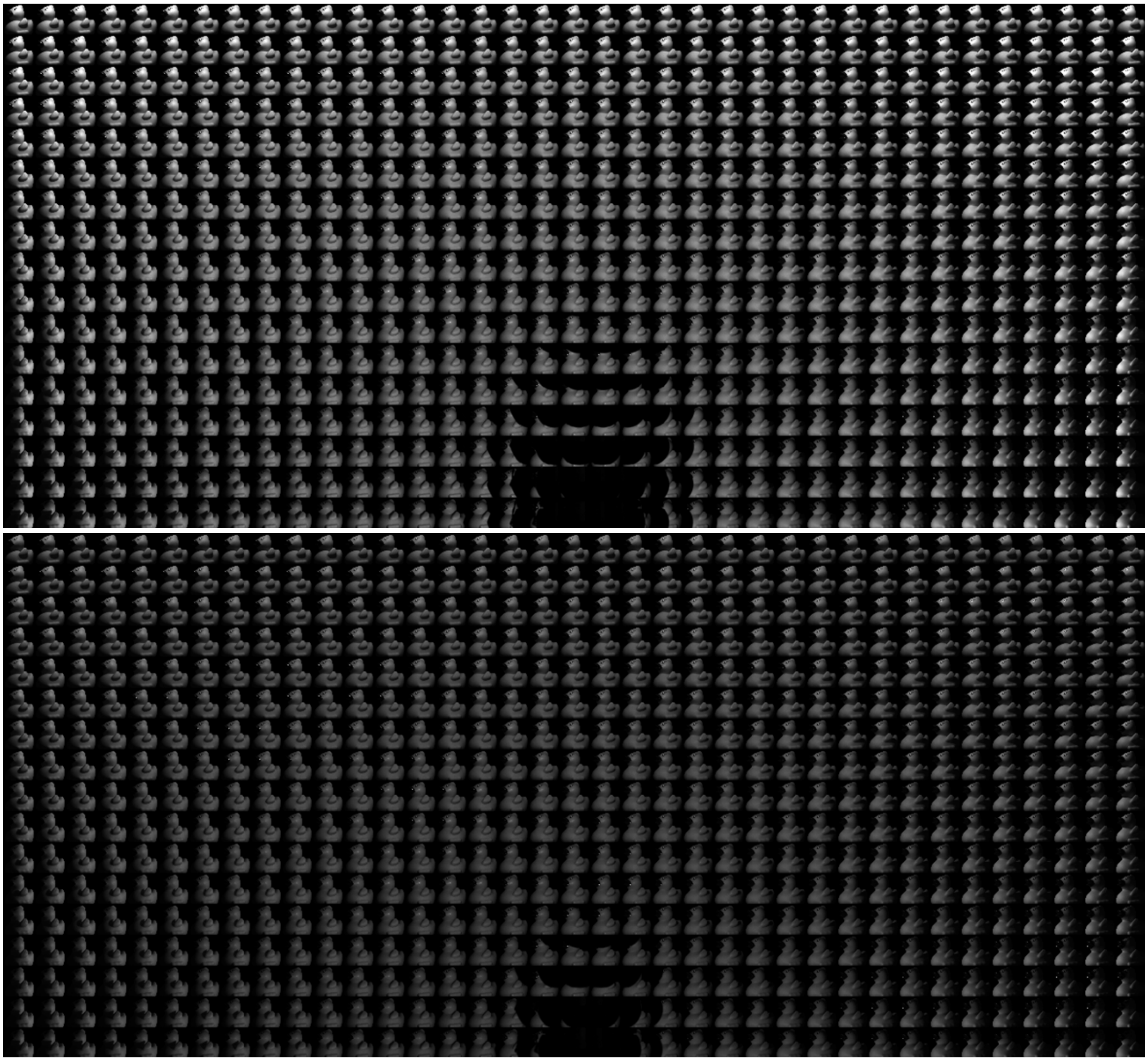

Figure 4.10: A sample set of images captured for inner (top image) and outer (bottom image) spherical light sampling. The vertical axis corresponds to light elevation (θ), ranging from 5° to 85°, while the horizontal axis is the azimuth direction (ϕ), ranging from $-90°$ to 90°. Individual images were roughly 1 megapixel crops and were scaled-down to fit the figure.

Figure 4.11. The top example tests the algorithm on a smooth cylinder and the bottom example presents a shape with a discontinuous depth. Top-down views of reconstructed profiles are shown below sample frontal views. Each thin line in the graph represents a line of pixels from a narrow 10 pixel horizontal strip across the center of the images above. They are overlaid with the approximate true geometry of each of the models marked as thick red lines. For both objects light positions on the spheres were sampled every 10°, which is half of that in Figure 4.10. The forward camera direction is misaligned slightly with the z-axis in both cases.

On average, the standard deviation of the recovered depth was 1.63 mm, or 1.2% of the inner radius of the sampling sphere. The highest recorded errors occured in a few spots, usually where the direct illumination assumption was violated through scatterng of the light in the space under the paper (the paper was only loosely attached with a tape). Note, that sharp discontinouities present in the left example would be especially difficult to recover by methods based on surface normal integration. Here reconstructions were performed on a per-pixel basis with no spatial smoothing.

Unlike reconstructions shown in Figure 4.11, reconstructions of real objects often contained some high-frequency noise, typically caused by errors in sub-pixel registration (refer to Section 4.3.4). Although a smoothness constraint in Equation 4.14 over several pixels would have helped, it would also make it difficult to preserve edges. One of the advantages of the algorithm is that it works on a per pixel basis and a spatial smoothness constraint would violate that. Therefore a choice was made to leave filtering as a possible postprocessing step. In Figure 4.12 is an example of depth map postprocessed using a median filter. Note the high-frequency noise on what would be sharp brightness discontinuities in the camera view. These include physical edges of the object and texture transition regions. After applying an edge-preserving median filter, the noise is mostly gone.

Next there are several reconstruction examples of real non-Lambertian objects. Depth maps were post-processed with a median filter in a similar fashion shown in Figure 4.12. Figure 4.13 shows depth reconstruction of the rubber duck used earlier to illustrate the algorithm. The complete set of input images for the duck is shown in Figure 4.10. The same light source sampling was used for the other reconstruction examples. The duck is made of rubber and is partially painted with glossy paint. There are multiple visible specularities on the surface. Note, that the duck's eye has not been reconstructed since the amount of reflected light captured at this exposure level was below the dynamic range of the camera. Taking multiple exposures and constructing high-dynamic range images (HDR) [56] would be a potential solution. This would not only allow for capture of objects with a large range of surface albedo, but also provide more flexibility in capture of highly specular objects, since clipping also occurs for specularities. Similar dynamic range issues are also

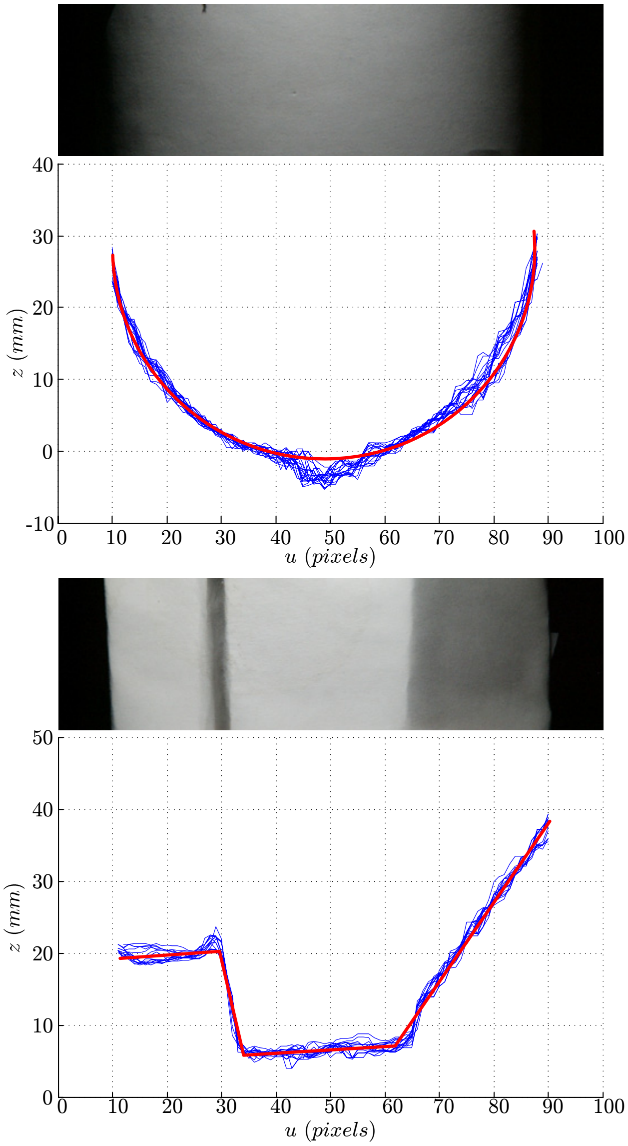

Figure 4.11: Depth reconstruction compared to the ground truth. At the top is a smooth cylinder and on the bottom is an object with sharp discontinuities. Reconstructed depth profiles (blue) are overlaid with the true profiles (red). Profiles are shown from the top view, perpendicular to the camera direction. The origin for the z axis is at the center of a sphere traced by the light source.

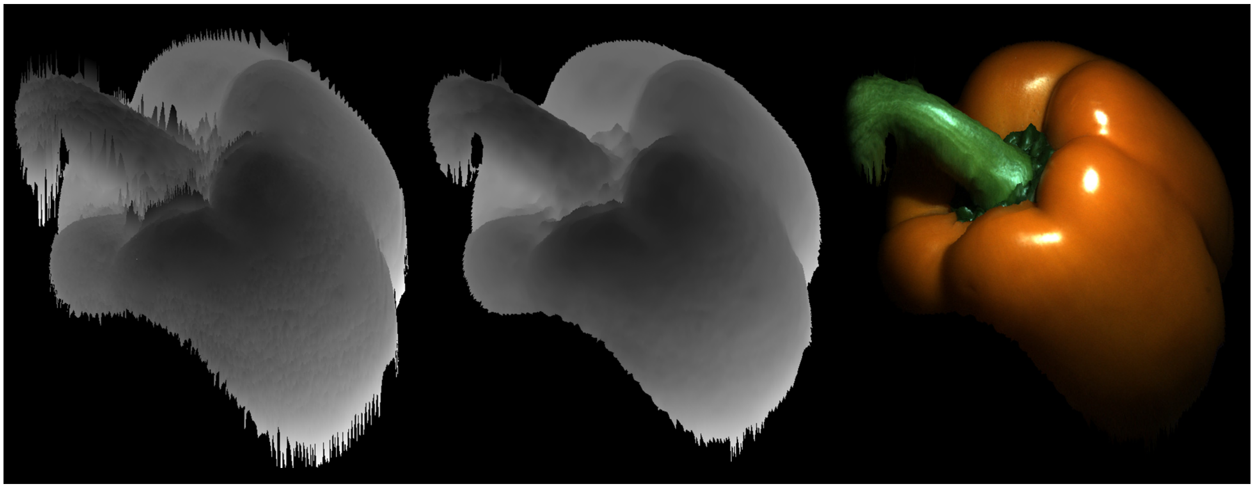

Figure 4.12: Depth map filtering. On the left is a raw depth map of a pepper containing some high-frequency noise, mostly over image brightness discontinuities. In the center is the same depth map after filtering with a 5x5 edge-preserving median filter. Final result after texturing is shown to the right. Depth maps were rotated away from the original camera view in order to better show the high-frequency noise.

present in the other examples.

Figure 4.14 shows depth reconstruction of a real pepper. The waxy surface is very glossy and partially transmissive. There is a detectable amount of subsurface scattering, however it did not significantly affect the result.

Finally Figure 4.15 shows depth reconstruction of a ceramic rabbit sculpture. In addition to specularities, the surface of the sculpture has a complicated macrostructure with a significant amount of macro-shadowing. There is also a sharp discontinuity, where the back ear is occluded by the head. While it could be a problem during the surface normal field integration step in traditional photometric stereo, here the discontinuity has no impact on the reconstruction process.

4.6 Image-based Relighting

Image-based relighting [243, 146] is a technique of image-based modeling and rendering, or realistic synthesis of new images using a pre-recorded set of images, typically photographs of a real scene. Since images can be considered samples of the plenoptic function [1], the rendering process turns into plenoptic function resampling. When the lighting is fixed, the problem reduces to that of light field rendering [141, 80]. The light field is a function representing the reflected radiance of the scene under fixed lighting, observed from any view point in space. Light field rendering uses the light field function sampled during the acquisition process to render images of the same scene under the same illumination but from novel view points. Image-

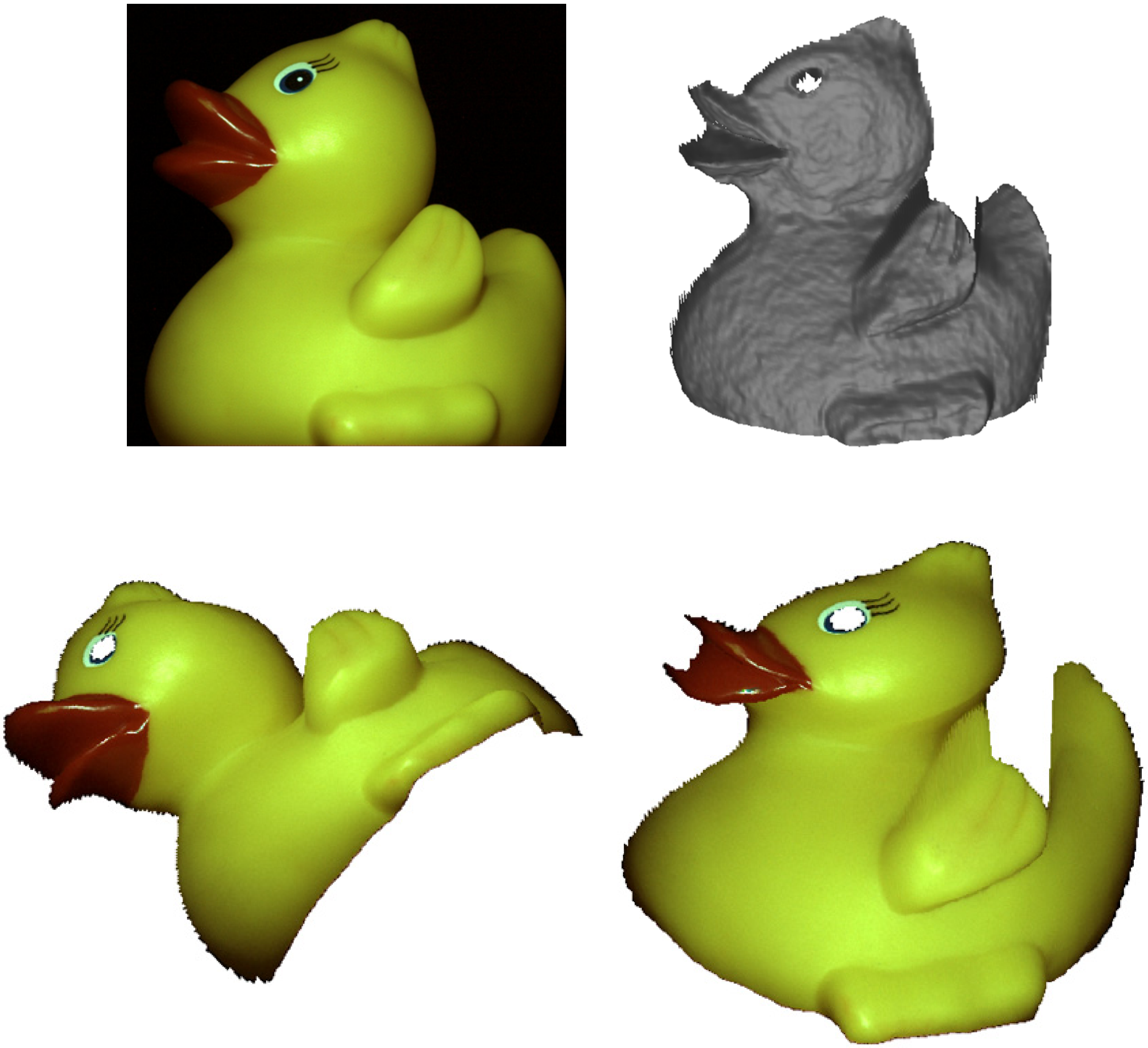

Figure 4.13: Depth reconstruction of a rubber duck. At the top is a sample input image with the recovered depth map to the right shown as a lit mesh using the Lambertian lighting model. The mesh was rotated slightly away from the camera view. At the bottom are two rendered novel camera views of depth maps textured using the sample image from the top.

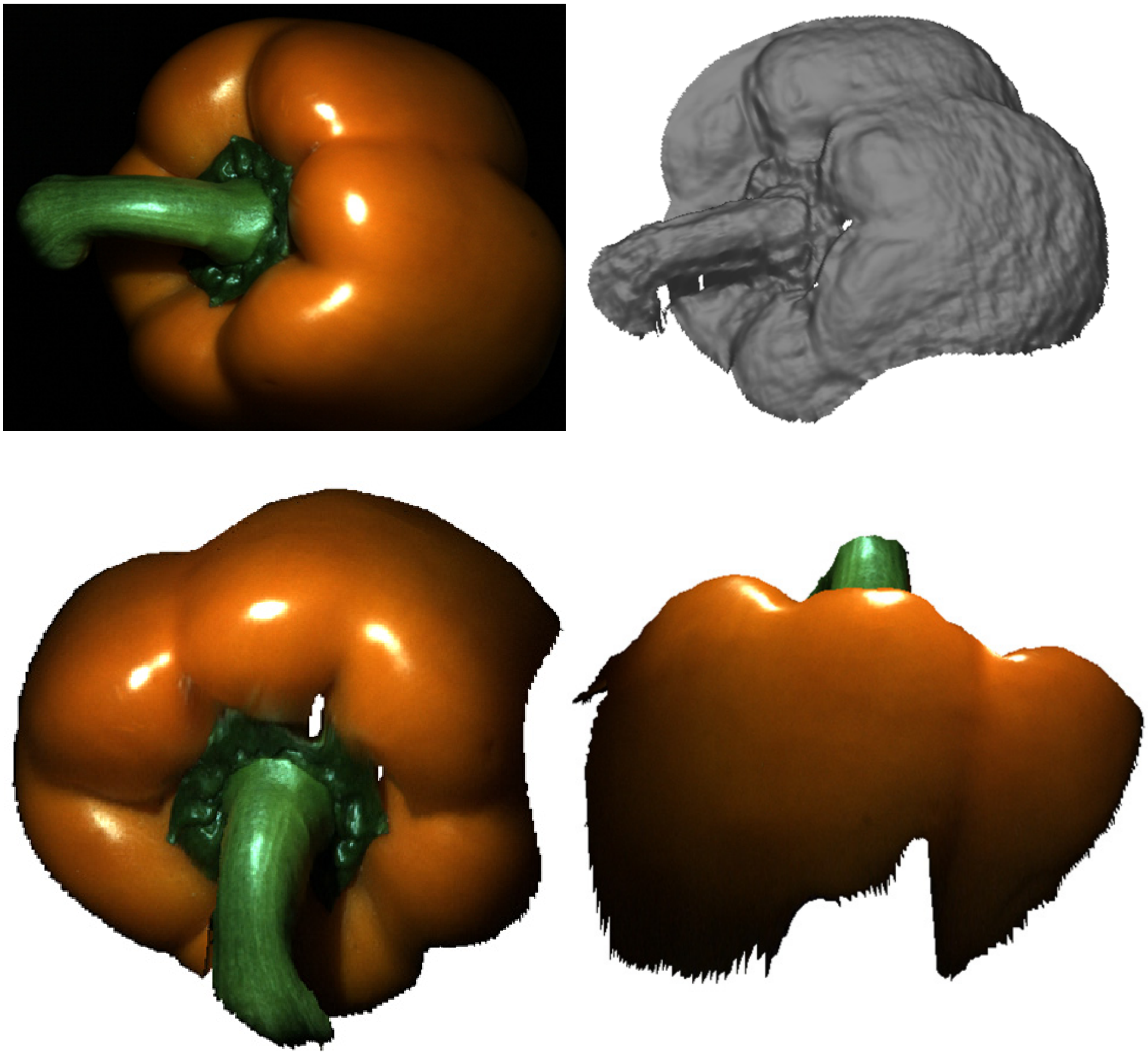

Figure 4.14: Depth reconstruction of an orange pepper. At the top is a sample input image with the recovered depth map to the right shown as a lit mesh using the Lambertian lighting model. The mesh was rotated slightly away from the camera view. At the bottom are two rendered novel camera views of depth maps textured using the sample image from the top.

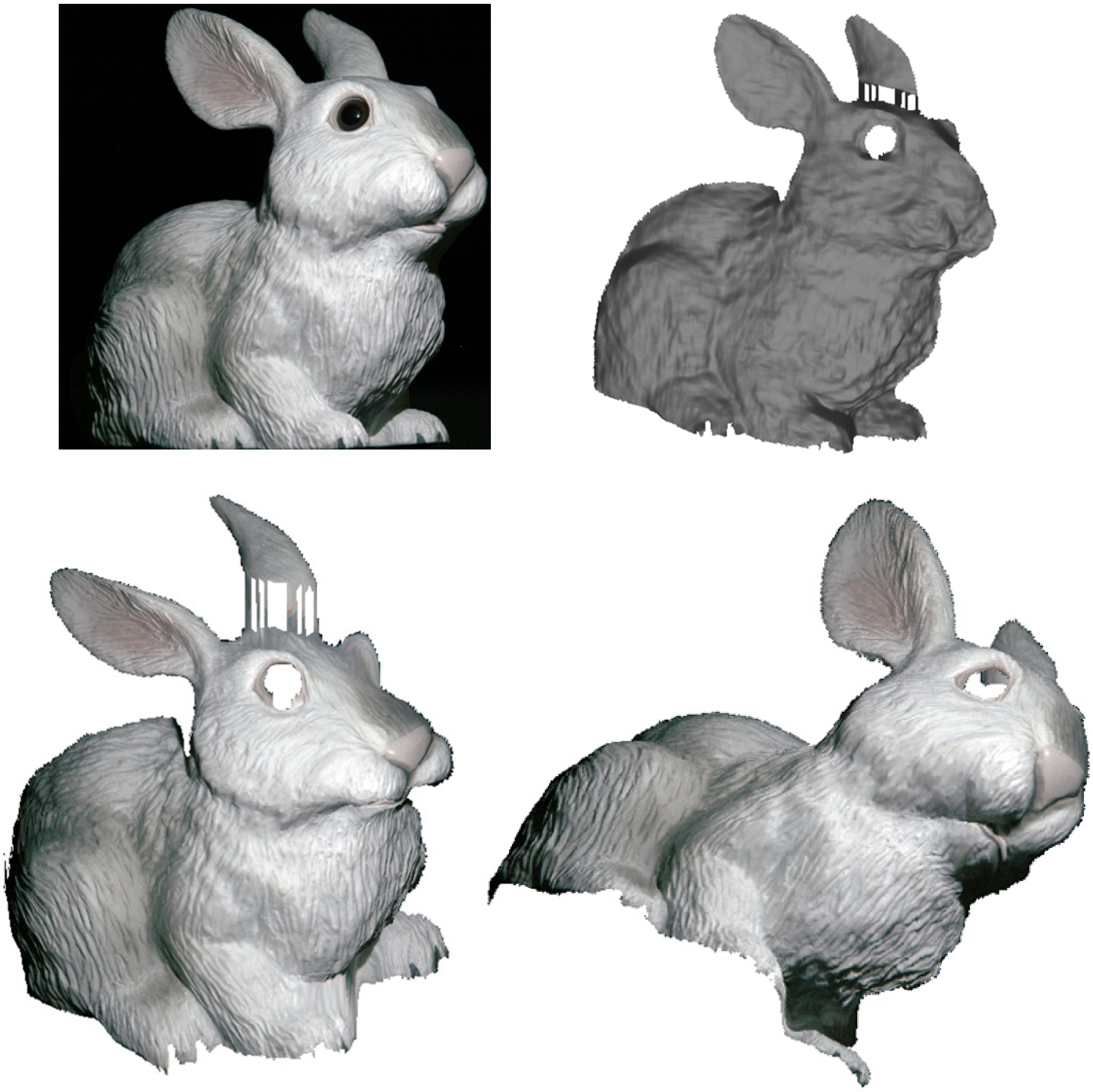

Figure 4.15: Depth reconstruction of a rabbit sculpture. At the top is a sample input image with the recovered depth map to the right shown as a lit mesh using the Lambertian lighting model. The mesh was rotated slightly away from the camera view. At the bottom are two rendered novel camera views of depth maps textured using the sample image from the top.

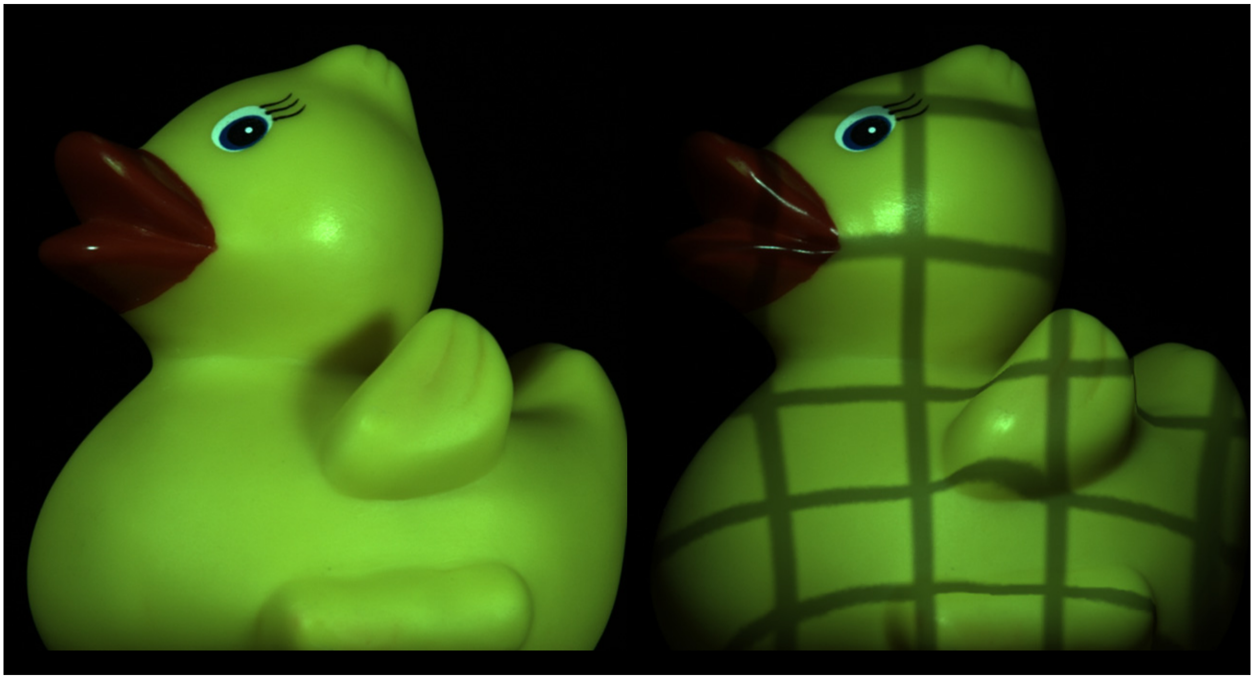

Figure 4.16: Two relighting examples of the rubber duck with nearby light sources using the recovered geometry. On the left is an example of a novel illumination using two nearby point sources placed on the two sides of the model. On the right is a synthesized image of the rubber duck with a square grid placed near the surface. The grid casts a shadow on the geometry by blocking one of two light sources.

based relighting extends this concept by sampling the incident illumination. This way new images of the scene can be rendered given new illumination conditions.

Our choice of a spherical configuration for the incident light sampling described in Section 4.3.2 is not arbitrary, as our aim is to capture the effects of illumination for two purposes: depth reconstruction and relighting. The spherical configuration is most convenient for scene relighting, as it efficiently covers all possible incoming light directions. At the same time, it can also be used to recover the depth. Knowing the depth of the scene is advantageous, as it allows relighting the scene with nearby light sources, when the effects of radiant energy density falloff with distance became apparent. Note that this is not very feasible with image-based relighting, as it would require dense sampling of the incident light in three dimensions, substantially increasing the size of the original data and the acquisition time. Therefore in image-based relighting lights are typically distant [243, 55, 156]. Instead, we sample the incident illumination at two distances to the surface and first recover the depth of the scene. Then we use one of the original sets and the recovered depth to render the scene under a novel illumination.

For convenience, we have chosen to define the new lighting using an illumination function $\mathcal{I}_{new}(\phi, \theta)$ with a corresponding radial distance function $l_{new}(\phi, \theta)$. The illumination function $\mathcal{I}_{new}(\phi, \theta)$ defines the intensity of a virtual point source placed along the direction (ϕ, θ) at a distance specified by $l_{new}(\phi, \theta)$. Both

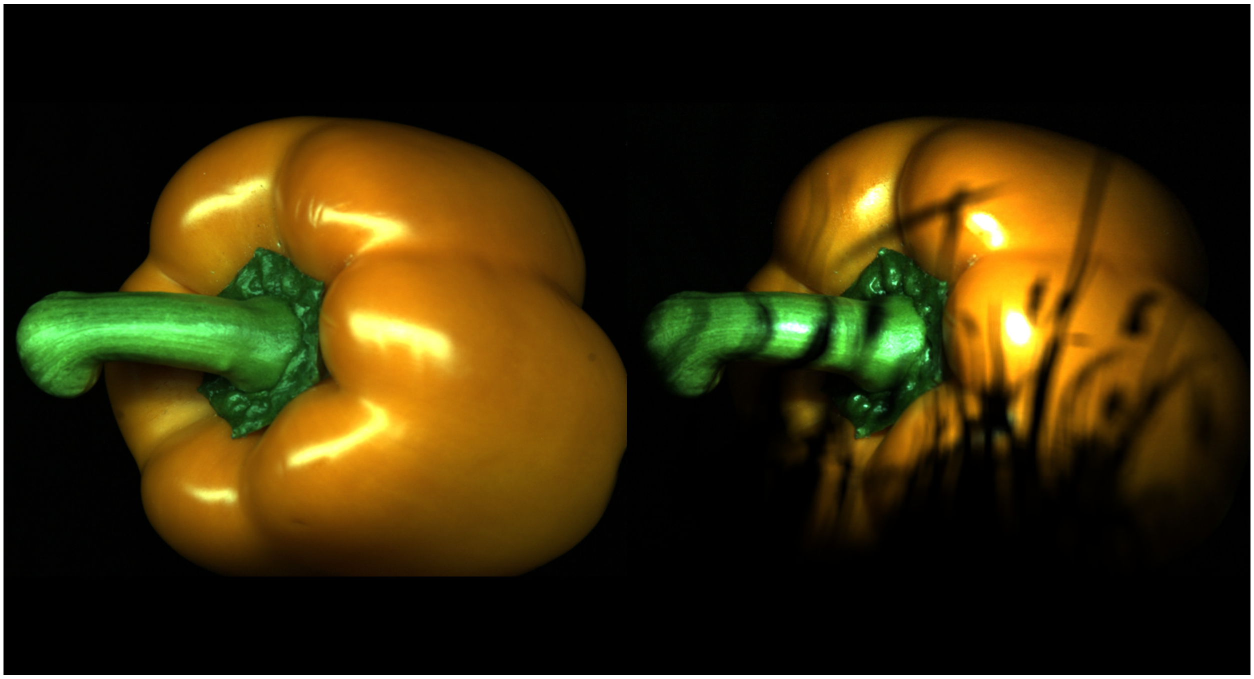

Figure 4.17: Two relighting examples of the orange pepper with nearby light sources using the recovered geometry. On the left is an example of a novel illumination using a line source extending horizontally. On the right is a synthesized image of the orange pepper with a field of grass illuminated from behind by an area light source . The grass casts shadows on the geometry by blocking the area source (note the blurring effect).

are defined in the global coordinates of the scene. Note, that any area sources can be also be approximated by resampling them into virtual point sources. This way we can specify an arbitrary proximate lighting environment that can be resampled using uniform point light sources[2].

In order to recover the depth $\hat{s}(u,v)$ for a particular pixel (u,v) we have earlier captured two global radiance functions: $L_1(\phi,\theta;u,v)$ and $L_2(\phi,\theta;u,v)$. When L_1 and L_2 are remapped into the local coordinates corresponding to the point $(u,v,\hat{s})$, they satisfy Equation 4.14 and therefore can be used to compute the product of reflectance and intensity from Equation 4.6:

$$f_{r\to c}(\phi_{\hat{s}},\theta_{\hat{s}};u,v)\mathcal{I}_0 = L_1(\phi_{\hat{s}},\theta_{\hat{s}};u,v){l_1}^2(\phi_{\hat{s}},\theta_{\hat{s}}) = L_2(\phi_{\hat{s}},\theta_{\hat{s}};u,v){l_2}^2(\phi_{\hat{s}},\theta_{\hat{s}}) \tag{4.18}$$

Note, that we could use either L_1 or L_2. We can now compute the radiance seen by the camera under the new illumination $\mathcal{I}_{new}$:

$$L_{new}(u,v) = \int_{\Phi}\int_{\Theta} f_{r\to c}(\phi_{\hat{s}},\theta_{\hat{s}};u,v)\mathcal{I}_0 \frac{\mathcal{I}_{new}(\phi_{\hat{s}},\theta_{\hat{s}})}{{l_{new}}^2(\phi_{\hat{s}},\theta_{\hat{s}})} d\theta_{\hat{s}} d\phi_{\hat{s}} \tag{4.19}$$

---

[2] More complex lighting that cannot be approximated using uniform point sources requires additional two dimensions to describe the directional power distribution pattern of each point source.

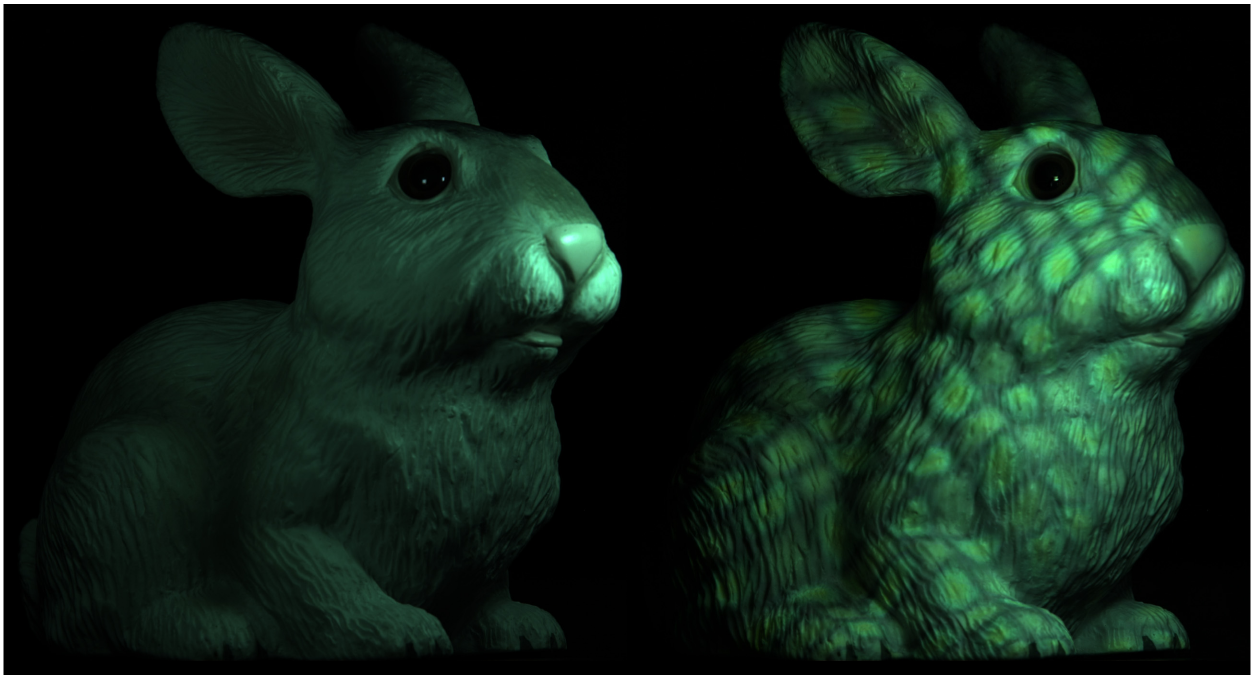

Figure 4.18: Two relighting examples of the rabbit sculpture with nearby light sources using the recovered geometry. On the left is an example of a novel illumination using two point sources one near the surface and the second far away. On the right is a synthesized image of the rabbit sculpture illuminated with light passing through a stained glass.

If the original point source intensity $\mathcal{I}_0$ is unknown, $\mathcal{I}_{new}$ has to be specified as a fraction of $\mathcal{I}_0$. Equation 4.19 uses the same remapping procedure as used by the depth reconstruction algorithm from Section 4.2.1. The intensity function $\mathcal{I}_{new}$ is mapped the same way as radiance functions, that is $\mathcal{I}_{new_P}(\phi_P, \theta_P) = \mathcal{I}_{new}(\phi, \theta)$, where P is the new origin point for the operator $\mathcal{W}_P$.

Several examples of objects rendered using this technique are shown next. All were synthesized for the same camera view, but under novel illumination conditions. In Figure 4.16 are two examples using the rubber duck model. The left image uses two nearby light sources to demonstrate the effects of geometry, and the right one shows a shadow cast by a regular rectangular grid placed in front of one of the lights. Figure 4.17 demonstrates a non-point nearby light source (a strip source in this case) illuminating the orange pepper model. In addition, the second image shows soft-shadows cast by a shadow mask modeling a field of grass in front of a nearby square area light source. Finally Figure 4.18 has an example of an image rendered with a light source positioned very close to the surface of an object. The second image shows the rabbit sculpture illuminated with light passing through an artificial stained glass window. Note, that none of these images would be possible to render correctly if the geometry of the scene was unknown.

Chapter 5

Capturing the Appearance of Volumetric Surface Textures

This chapter describes a new technique for capturing volumetric appearance of a 3-D texture with unknown reflectance and structure. From the acquired reflectance data in the form of a 6-D Bidirectional Texture Function (BTF), the proposed algorithm creates an efficient volumetric representation in the form of a stack of semi-transparent layers each representing a slice of the original material. In addition to negligible storage requirements, this representation is ideally suited for hardware-accelerated real-time rendering.

5.1 Introduction

The focus of this algorithm is on the most complex form of surface materials which can only be accurately reproduced with a volumetric representation (refer again to Figure 1.3 for some examples). Although efficient algorithms for rendering volumetric textures have been known for years, capturing the richness of real volumetric materials remains a challenging problem. Currently volumetric acquisition options are limited. This is especially true, if the aim is to capture the reflectance of a material as well. Existing rendering algorithms typically use computer-generated representations [158, 60, 119, 170, 51, 140, 237, 73]. There are many acquisition techniques that can recover the geometry or the density of the material, but are not suited for capturing of the reflectance. These include any of the volumetric scanning techniques discussed in Chapter 3. In particular, computed tomography methods (see Section 3.1.1) could work quite well and have been used before [36].

The algorithm described here is quite similar in nature to tomography. It also relies on the principle of non-diffusive transmission of radiation and uses a similar integration-like formulation of light passing through a volumetric material. A material sample is modeled using a stack of vertically offset texture layers. At each layer point a vector of BRDF parameters models the local reflectance and the geometry is represented with the volumetric attenuation. At runtime, layers can be synthesized over a surface giving an impression of a 3-D texture. Similar volumetric representation have been shown to be useful for a number of volumetric textures [158, 170, 140]. The texture model representation used by the proposed algorithm will be referred

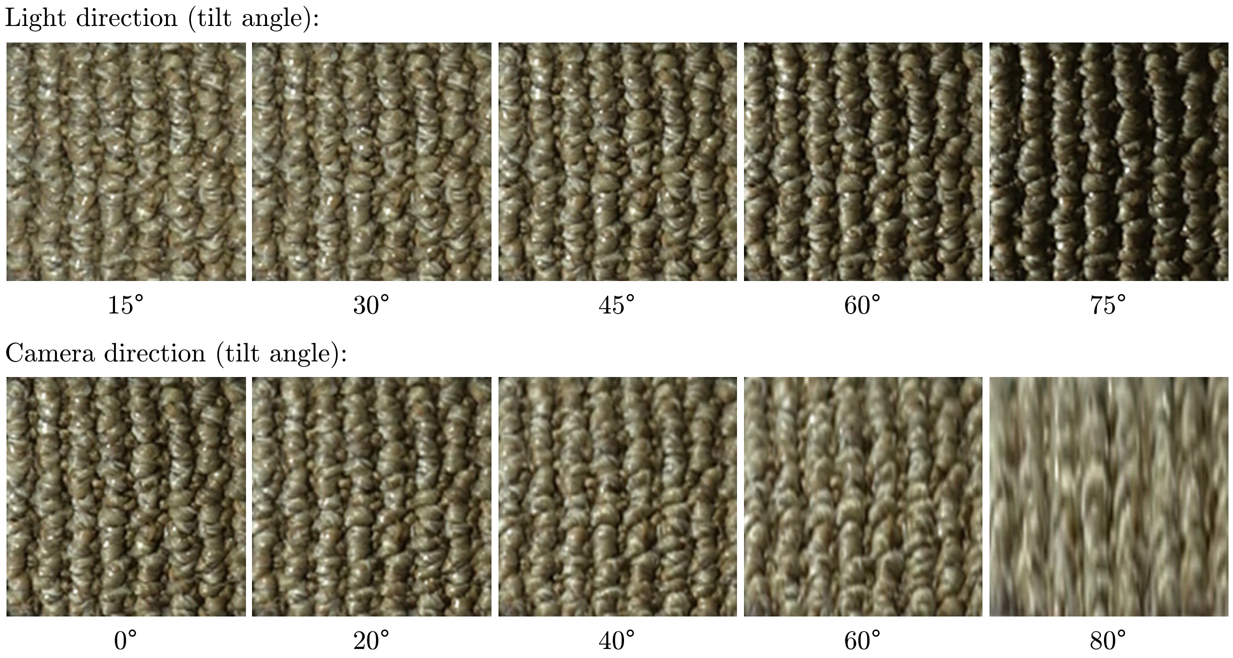

Figure 5.1: A BTF example for a woven carpet sample. The figure shows several samples for changing light tilt direction (away from the sample normal direction, top row) and changing camera tilt direction (bottom row).

to as the Layered Volumetric Surface Texture (LVST) model.

5.1.1 Bidirectional Texture Functions

In general, the appearance of a surface material can be described with the *bidirectional texture function* (BTF). The concept was proposed by Dana et al. [50]. This 6-D function of light and viewing directions describes the apparent reflectance of a surface when reprojected onto a fixed plane. As the result, a material sample can be described using a BTF implicitly encapsulating both the geometry and the reflectance. In Figure 5.1 are a few samples of a BTF representing a sample of a woven carpet. The top row shows the changing appearance as the light source direction changes. On the bottom are shown the effects of changing camera direction. Note the blurriness caused by the decreasing sampling rate in the vertical direction as the camera direction is near the grazing angle. This is only the effect of the limited resolution of the camera. The sampling rate ideally should stay the same.

5.1.2 Related Work

There are several related methods that aim to infer the geometric structure of a texture. Liu et al. [148] and Haindl et al. [84] used shape-from-shading to construct a surface displacement map. Shape-from shading methods work for BTFs with simple reflectance for which the Lambertian reflectance assumption applies. Another approach taken by Yu and Chang [252] uses shadow graphs for height map reconstruction. Their method is a combination of shape-from-shading and shape-from-shadow with constraints enforcing consistency between the two. Wang and Dana [234] used a library of geometric textons to reconstruct an approximation of a BTF height map. Although the focus of their method is actually better reconstruction of shadowing effects rather than geometry, it generates a height map as well. It also works for non-Lambertian surfaces for which the reflectance model can be approximated by a known model. Unfortunately, displacement mapping methods are not very useful for textures with a more complex geometry.

Although volumetric reconstruction into layers has not attracted much attention in the BTF setting, it has been explored in the area of stereo reconstruction, which was discussed earlier in Section 3.2.2. Stereo algorithms aim to estimate the depth of a scene using two or more images taken from different locations. This information can then be used to create a volumetric model of the scene. Capturing transparency has recently attracted more attention for two reasons: some scenes actually contain transparent objects (glass windows, glass containers, etc.), but also in order to make boundaries of opaque regions smoother. Of particular interest is the work by Szeliski and Golland [224]. Given a stereo sequence, their algorithm extracts layers with semi-transparent boundary regions. By performing a global optimization on the color and transparency value of boundaries, they can obtain smooth, semi-transparent edges. However, as with most stereo reconstruction algorithms, their algorithm does not model reflectance.

Volumetric decomposition of BTFs can also be viewed as a compression scheme. Storage requirements typically make it impractical to use BTFs directly in real-time rendering. Over the past few years there has been an ongoing work on better compression and parametrization. Most compression solutions use some variation of eigen decomposition or principal component analysis [163, 130, 231, 147, 150]. For additional compression, some also utilize Laplacian pyramids [84, 150]. A different approach is to view BTFs as a 2-D apparent BRDF (ABRDF) function [243], that also encapsulates the effects of geometry, reflections, transparency, and shadowing. Per-pixel view and lighting variations can then be approximated by ABRDF clustering [47, 227, 234] or modeled. Modeling can be done using traditional parametric BRDF representations [157, 70, 124] or factorization techniques [222].

5.2 The LVST Model

The volumetric model used by the algorithm consists of a stack of N layers, as pictured in Figure 5.2a. Each layer is parametrized with (u_n, v_n), where u and v are the 2-D texture dimensions, and n is the layer

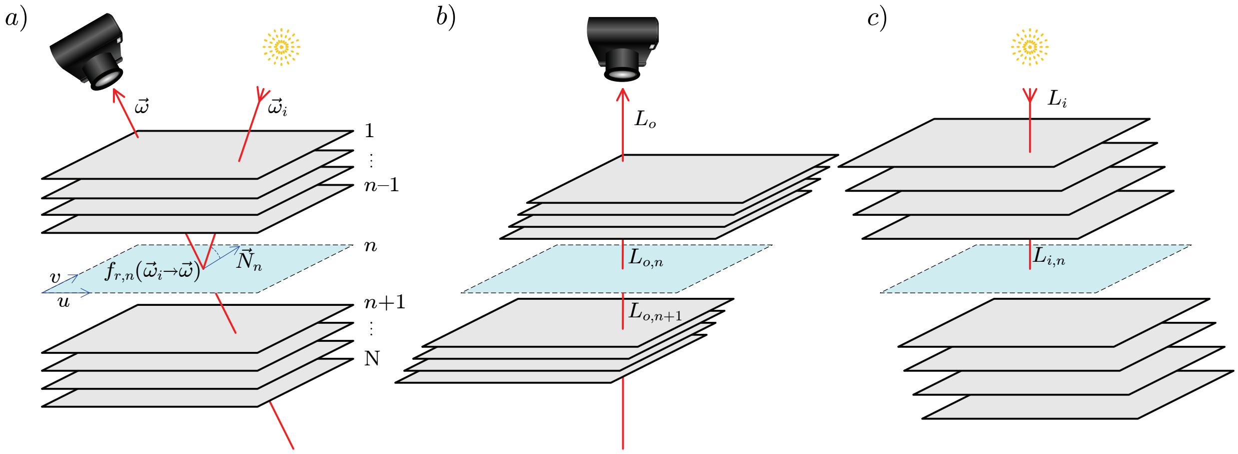

Figure 5.2: The LVST model. (a) Layer parameters relate the total outgoing radiance in terms of the reflected illumination and the total outgoing radiance from layers below. (b) Layers warped into the camera direction. (c) Layers warped into the light source direction.

index. We number layers starting at one for the top layer. The bottommost layer is always fully opaque. Rendering involves superposition of each layer by re-projection into the view direction and computation of the reflectance values for each of them. This composing process can be expressed as a backward warping operation $\mathcal{W}_b$ applied to each layer projecting them into the viewing direction. This is accomplished using the shear-warp algorithm [134]. Under parallel projection this operation reduces to simple parallax shifts of layers relative to each other. The inverse of the backward warping operator $\mathcal{W}_f$, maps the camera view along $\vec{\omega}$ into coordinates of a layer n as:

$$\begin{bmatrix} u_n & v_n \end{bmatrix}^{\mathbf{T}} = \mathcal{W}_f(\vec{\omega}, n) \begin{bmatrix} \tilde{u} & \tilde{v} \end{bmatrix}^{\mathbf{T}} \tag{5.1}$$

where $(\tilde{u}, \tilde{v})$ are the coordinates in the camera view. In the composing process of N layers, the total camera-observed brightness I_o for the current viewing direction along $\vec{\omega}$ is the sum of brightness contributions from each individual layer I_l attenuated by the product of the projected attenuation factors a:

$$I_o(\tilde{u}, \tilde{v}, \vec{\omega}) = \sum_{n=1}^{N} I_{l,n}(u_n, v_n, \vec{\omega}) \prod_{n'=1}^{n-1} a_{n'}(u_{n'}, v_{n'}, \vec{\omega}) \tag{5.2}$$

Note, that this composing expression can be made into a discretized approximation to Equation 2.33. By substituting radiance for the brightness in Equation 5.2, the total outgoing radiance L_o becomes the sum of the incoming radiance $L_{i,n}$ in-scattered at each layer n according to the bidirectional distribution function (BDF) f:

$$L_o(\vec{\omega}) = \sum_{n=1}^{N} \int_{4\pi} f_n\big(\vec{\omega}' \rightarrow \vec{\omega}\big) L_{i,n}\big(\vec{\omega}'\big)\, \vec{N}_n \cdot \vec{\omega}'\, d\vec{\omega}' \prod_{n'=1}^{n-1} a_{n'} \tag{5.3}$$

$\vec{N}_n$ is the normal vector. All parameters in Equation 5.3 are also functions of u and v which will be omitted from now on from equations for better readability.

The operation described by Equation 5.3 is essentially equivalent to computed tomography (CT) (see Section 3.1.1). The light is attenuated on its way to the surface as it travels in the volume and the attenuation is discretized at each layer. In CT attenuation is typically modeled as a continuous function. Here light interactions along the direction $\vec{\omega}$ are limited to layer intersection points and increasing the number of layers has the net effect of increasing the sampling rate along the ray. As in tomography, a measurement contains the effects of all light interactions along a ray. The source of radiation is on the same side of the sample as the detector, same as in reflection tomography (see Section 3.1.1). Still, the geometry is not estimated by timing the reflected echoes, but rather as in transmission tomography by estimating the absorption.

Equation 5.3 accounts for light contributions from all directions and therefore could potentially be used to approximate an arbitrary volumetric material. Still, a more application-specific approximation can lead to a more efficient implementation. In order to utilize the model for real-time rendering, the number of reflection computations has to be kept at a reasonable level. The algorithm presented here limits itself to just a single reflection and a diffuse correction term $L_{d,n}$ that approximates secondary reflections. It was designed specifically for non-diffusive volumetric materials with strong light source direction correlation. Under these conditions, light interactions are dominated by absorption (as opposed to scattering) and light arriving at the camera is mainly due to the first scattering reflection event. The integral of the incoming radiance $L_{i,n}$ at each layer is then approximated with just the primary incoming light direction $\vec{\omega}_i$:

$$\begin{aligned} &L_o(\vec{\omega}) = \\ &\sum_{n=1}^{N} f_{r,n}(\vec{\omega}_i \rightarrow \vec{\omega}) \Big[L_{i,n}(\vec{\omega}_i) + L_{d,n}\Big]\, \vec{N}_n \cdot \vec{\omega}_i \prod_{n'=0}^{n-1} a_{n'} \end{aligned} \tag{5.4}$$

We replace the BDF with the more appropriate reflectance term (BRDF) (see Appendix A.3) f_r, as we assume scattering is inconsequential. For a BRDF, we are free to choose any parametric reflectance model that best suits a particular material. BRDF models with fewer parameters have the advantage of faster conversion. In particular, we have been using the Lafortune model [135] which gives a good trade-off between generality and the number of required parameters.

The proposed LVST model can be summed up as consisting of a stack of layers with spatially-varying parameters. Each layer location (u_n, v_n) is described by several parameters:

- The surface normal $\vec{N}_n$.
- Parameters of the chosen BRDF $f_{r,n}$.
- The attenuation factor a_n.
- The diffuse radiance term $L_{d,n}$.

In the next section we will show how to iteratively solve for these unknown parameters.

5.3 Estimating the LVST Parameters from a BTF Dataset

Equation 5.4 expresses the total outgoing radiance along a single ray. A captured BTF samples the outgoing radiance values L_o for discrete light and camera directions $(\vec{\omega}_i, \vec{\omega})$. The total incoming radiance L_i can also be measured. For each layer n, we would like to find the best estimate for the BRDF $f_{r,n}$, attenuation a_n, surface normal $\vec{N}_n$, and the diffuse light term $L_{d,n}$, so that the observed error between rendered and originally measured radiance values is minimized. Globally we are trying to solve the inverse image synthesis problem with unknown reflectance and geometry. This problem appears in vision applications and, in general, is difficult to solve. For a particular BRDF model with f parameters, we have to solve for $n \cdot u \cdot v \cdot (f + 4)$ unknowns. For the trichromatic RGB model, the number of unknows is even greater. Solving for them directly, even for models with just a few layers, can be prohibitive. Instead, we solve this high dimensionality problem by expressing it as a series of local optimization steps. In particular, we iteratively perform local optimizations on individual layers.

5.3.1 Layer Optimization

Consider the radiance $L_{o,n}$ leaving a particular layer n toward the camera along $\vec{\omega}$ (see Figure 5.2). By looking back at Equation 5.4 we can see, that it is equal to the sum of the incoming radiance $L_{i,n}(\vec{\omega}_i) + L_{d,n}$ reflected by the layer and the radiance transmitted from the layers below $L_{o,n+1}$. We will denote the reflected radiance with $L_{l,n}$:

$$\begin{aligned} L_{o,n}(\vec{\omega}) &= L_{l,n}(\vec{\omega}) + a_n L_{o,n+1}(\vec{\omega}) \\ L_{l,n}(\vec{\omega}) &= f_{r,n}\big(\vec{\omega}_i \rightarrow \vec{\omega}\big)\Big[L_{i,n}(\vec{\omega}_i) + L_{d,n}\Big]\,\vec{N}_n \cdot \vec{\omega}_i \end{aligned} \tag{5.5}$$

Lets assume, that all incoming and outgoing radiance values are known and we are given a set of $(L_{i,n}, L_{o,n+1}, L_{o,n})$ triplets for K different light and viewing directions. For each discrete position on a layer n we want to solve for the unknown parametric BRDF function $f_{r,n}$, the diffuse radiance constant $L_{d,n}$, the surface normal $\vec{N}$, and the attenuation coefficient a_n. The least-squares solution can be obtained by minimizing objective function $\mathcal{O}$ of the form:

$$\mathcal{O} = \sum_{k=1}^{K} \Big[L_{l,n,k} + a_n L_{o,n+1,k} - L_{o,n,k} \Big]^2 w_{n,k} \tag{5.6}$$

We have dropped some notation from Equation 5.5 for readability.

The weighting factor $w_{n,k}$ in the Equation 5.6 is crucial, as it provides the means to distribute radiance into appropriate layers. It is used to penalize data when a layer point is not in the direct line of sight of the camera and represents visibility. It can be written as a product of attenuations:

$$w_{n,k} = \prod_{n'=1}^{n-1} a_{n'} \tag{5.7}$$

Unlike the global problem, the local objective function is only non-linear in the BRDF model parameters. We will directly take advantage of that fact later on. The optimization process can broken down into a series of layer updates. For each layer, the algorithm uses the current estimate of the model to compute necessary layer radiance values. From each input image in the sampled BTF, we can obtain the outgoing radiance from the top layer $L_{o,1} = L_o$. Similarly, the total incoming radiance from the light source is the same as the incoming radiance to the top layer $L_{i,1} = L_i$. If the current state of the LVST model is known, we can also compute radiance values for other layers. In general, the outgoing radiance $L_{o,n}$ from layer n is equal to the difference between the measured radiance L_o and radiance contributions from layers above n:

$$L_{o,n} = \frac{1}{a_{n-1}} \Big[L_o - \sum_{n'=1}^{n-1} L_{l,n'} \prod_{n''=1}^{n'-1} a_{n''} \Big] \tag{5.8}$$

In a similar way, we can express the radiance incoming from below the layer n as the sum:

$$L_{o,n+1} = L_{l,n+1} + \sum_{n'=n+2}^{N} L_{l,n'} \prod_{n''=n+1}^{n'-1} a_{n''} \tag{5.9}$$

Finally, to compute the radiance due to direct illumination, we first perform a warping operation similar to $\mathcal{W}_f$, but with respect to the light source direction (see Figure 5.2c). We can express the direct illumination reaching the layer n as:

$$L_{i,n} = L_i \prod_{n_i=1}^{n-1} s_{n_i} \tag{5.10}$$

The shadowing coefficient s_{n_i} is eventually equal to the attenuation a_{n_i}. Initially model optimization is done without shadowing, which is then estimated in the final pass. The reasons for that will be explained in Section 5.3.5.

In summary, the algorithm begins from the bottommost layer and iterates through all layers until convergence is achieved. All layers are initialized to be fully transparent. For each BTF sample image, the layer stack undergoes viewer and camera direction warping operation $\mathcal{W}_f$ and the three layer radiance values are computed. These can then used to solve the least-squares problem from Equation 5.6.

5.3.2 BRDF Model Optimization

The majority of BRDF models are non-linear and therefore minimization of the objective function from Equation 5.6 generally requires a non-linear solution. Unfortunately non-linear minimization algorithms are also much slower than their linear counterparts, as they typically require multiple iterations over the input data. This is problematic, as typical input BTF data sets are very large, containing on the order of thousands of samples while Equation 5.6 must be solved for each individual pixel separately. Iterating over the whole input data is unpractical.

When dealing with large data sets, it is better to use incremental algorithms, such as the Extended Kalman Filter [16], conceptually equivalent to an incremental version of the Gauss-Newton method. The idea behind incremental solvers is to deal with large sets by incrementally iterating over smaller chunks of data. In particular, an incremental version of the Gauss-Newton method works by progressively updating the current gradient covariance matrix $\mathcal{C}_m$, as new data vectors $\mathbf{x}$ become available. A single iteration over the parameter vector $\mathbf{P}$ is performed every M such updates. It can be conceptually written it as:

$$\mathcal{C}_m = \lambda \mathcal{C}_{m-1} + \nabla \mathbf{x}_m \nabla \mathbf{x}_m^T \tag{5.11}$$

$$\mathbf{P}_i = \mathbf{P}_{i-1} - \mathcal{C}_{Mi}^{-1} \sum_{m'=M(i-1)}^{Mi} \nabla \mathbf{x}_{m'} \mathbf{x}_{m'} \tag{5.12}$$

The parameter vector $\mathbf{P}_i$ is updated every M data vectors $\mathbf{x}$. Instead of directly computing the inverse of $\mathcal{C}_m$, the algorithm actually uses the Sherman-Morrison update formula [79]. This saves some computation time. The smoothing parameter λ controls “forgetfulness” with respect to the past data. A value of one signifies that all observed data has equal weight. It is desirable to smoothly increase the value of λ after each update from some initial constant less than one approaching unity. This results in faster initial convergence. Given radiance values computed using Equations 5.8, 5.9, and 5.10, the algorithm evaluates the gradient of Equation 5.6 and subsequently updates the incremental solver. For more details on the incremental optimization algorithm refer to the work of Bertsekas [16].

5.3.3 An Example: Incrementally Solving for the Lafortune Model Parameters

As several LVSTs in the paper use the Lafortune BRDF model (see Section A.3.2), it will be used here as an example to show how the incremental optimization performs. Similar analysis can be done for other BRDF models.

The algorithm uses the simplified version of the Lafortune model with non-zero values only along the diagonal of the specular lobe matrix $\mathbf{C}_l$ (see Equation A.15). One-lobe Lafortune BRDF model with this simplification can be written as:

$$\boldsymbol{f}_r = \boldsymbol{K}_d + \boldsymbol{K}_s(\vec{\omega}^{\mathbf{T}} \boldsymbol{C}_l \vec{\omega}_i)^{k_l} \tag{5.13}$$

,where $\boldsymbol{K}_d$ and $\boldsymbol{K}_s$ are the diffuse and the specular coefficients and k_l is the specular exponent for the lobe. Trichromatic color parameters are indicated in bold. When the specular coefficient is zero, BRDF reduces to the constant diffuse coefficient.

There is a total of 15 unknown parameters, including the attenuation and the diffuse radiance correction term. In order to minimize the objective function from Equation 5.6 by using the incremental method, the gradient vector $\nabla\mathbf{x}$ of the objective function must be evaluated for each input sample data. The residual in Equation 5.6 has the form:

$$\boldsymbol{R}_n = \boldsymbol{f}_r \, \vec{N}_n \cdot \vec{\omega}_i \, [L_{i,n} + L_{d,n}] + a\boldsymbol{L}_{o,n+1} - \boldsymbol{L}_{o,n} \tag{5.14}$$

The gradient vector $\nabla\mathbf{x}$ therefore is:

$$\frac{\partial \mathcal{O}}{\partial \vec{N}} = \vec{\omega}_i \Big[\boldsymbol{K}_d + \boldsymbol{K}_s (\vec{\omega}^{\mathbf{T}} \boldsymbol{C}_l \vec{\omega}_i)^{k_l} \Big] \, [L_i + L_d] \cdot \boldsymbol{R} \, w \tag{5.15}$$

$$\frac{\partial \mathcal{O}}{\partial \boldsymbol{K}_d} = \vec{N} \cdot \vec{\omega}_i \, [L_i + L_d] \;\, \boldsymbol{R} \, w \tag{5.16}$$

$$\frac{\partial \mathcal{O}}{\partial \boldsymbol{K}_s} = (\vec{\omega}^{\mathbf{T}} \boldsymbol{C}_l \vec{\omega}_i)^{k_l} \; \vec{N} \cdot \vec{\omega}_i \, [L_i + L_d] \;\, \boldsymbol{R} \, w \tag{5.17}$$

$$\frac{\partial \mathcal{O}}{\partial \boldsymbol{C}_l} = \vec{\omega}^{\mathbf{T}} \vec{\omega}_i \; k_l \boldsymbol{K}_s (\vec{\omega}^{\mathbf{T}} \boldsymbol{C}_l \vec{\omega}_i)^{k_l - 1} \; \vec{N} \cdot \vec{\omega}_i \, [L_i + L_d] \cdot \boldsymbol{R} \, w \tag{5.18}$$

$$\frac{\partial \mathcal{O}}{\partial k_l} = \ln(\vec{\omega}^{\mathbf{T}} \boldsymbol{C}_l \vec{\omega}_i) \, \mathbf{K}_s (\vec{\omega}^{\mathbf{T}} \boldsymbol{C}_l \vec{\omega}_i)^{k_l} \; \vec{N} \cdot \vec{\omega}_i \, [L_i + L_d] \cdot \boldsymbol{R} \, w \tag{5.19}$$

The layer subscript n has been dropped for better readability. Recall, that w is the weighting factor defined in Equation 5.7.

The incremental algorithm requires several other parameters to be defined. These are best found experimentally. This was done using an artificial data set of a generated Lafortune model reflection. Samples were taken uniformly over a hemisphere of light and camera directions and then randomly ordered. Even though the whole set consisted of 10800 samples, the algorithm typically converged for the Lafortune model parameters after $10 - 20\%$ of samples. This convergance rate depended somewhat on the noise level, as shown in Figure 5.3. Four examples are shown for different levels of added Gaussian noise.

Even with a noise level of 20%, the algorithm performed reasonably well with parameter values slowly approaching the true values by the 5000th sample. The exponential parameter k_l took the longest to converge and therefore always limited the minimum required number of iterations. In general, that would be true of any exponential BRDF parameteres. An interesting observation about the effects of noise is that small amounts actually improve the convergance speed. Notice that with $\sigma = 0.05$ parameters (in particular the specular exponent k_l) converge faster than for $\sigma = 0.025$. This is related to the smoothing parameter λ which controls the weighting of data samples processed on earlier iterations. The more weight is put on the previously seen data the slower the algorithm is to change the parameter values. Noise acts in a similar way, as it obscures the true values essentially decreasing their impact on the parameters.

The effect of changing the smoothing parameter λ in the presence of noise is demonstrated in Figure 5.4. In the algorithm, the parameter λ is smoothly increased from some initial value λ_0 to asymptotically approach unity using the formula:

$$\lambda = 1 - \frac{\lambda_0}{1 + \lambda_d N m^2} \tag{5.20}$$

The parameter N is the data set size and m is the current sample number. λ_d controls the rate of decay

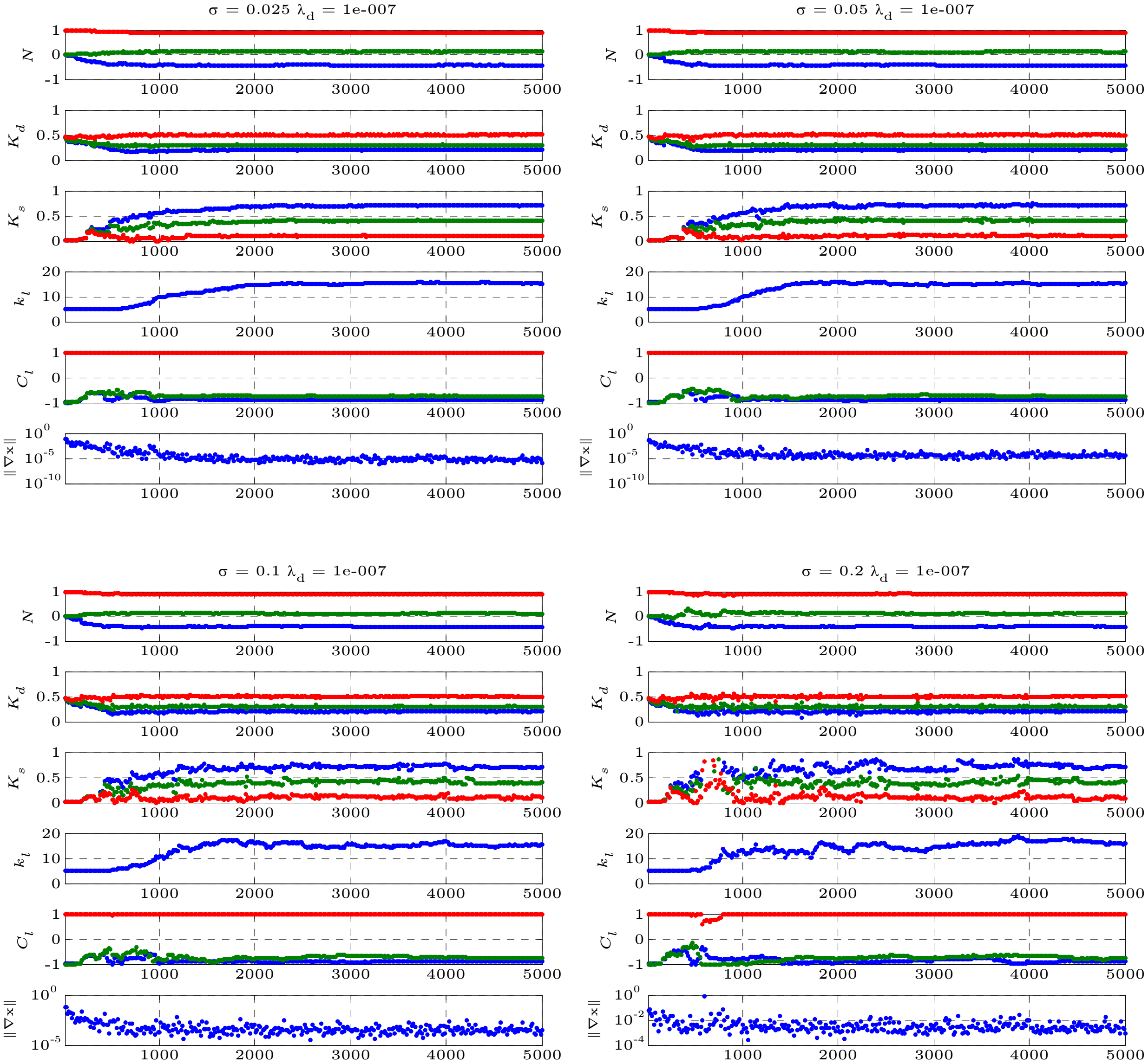

Figure 5.3: Incremental BRDF model optimization. Plots show the convergence for Lafortune model parameters for a single pixel using an artificial data set containing 10800 uniformly sampled and randomized data points. The parameters in each of the four plots listed from top to bottom are: the surface normal vector $\vec{N}$, the color diffuse term $\boldsymbol{K}_d$, the color specular term $\boldsymbol{K}_s$, the specular lobe exponent k_l, and the three specular lobe coefficients $\boldsymbol{C}_l$ (refer to Equation A.15). The bottom plot shows the convergence of the magnitude of the gradient vector $\|\nabla \mathbf{x}\|$. The four plots show the convergence for different standard deviation values σ of additive Gaussian noise. Color terms are kept in the range $[0, 1]$.

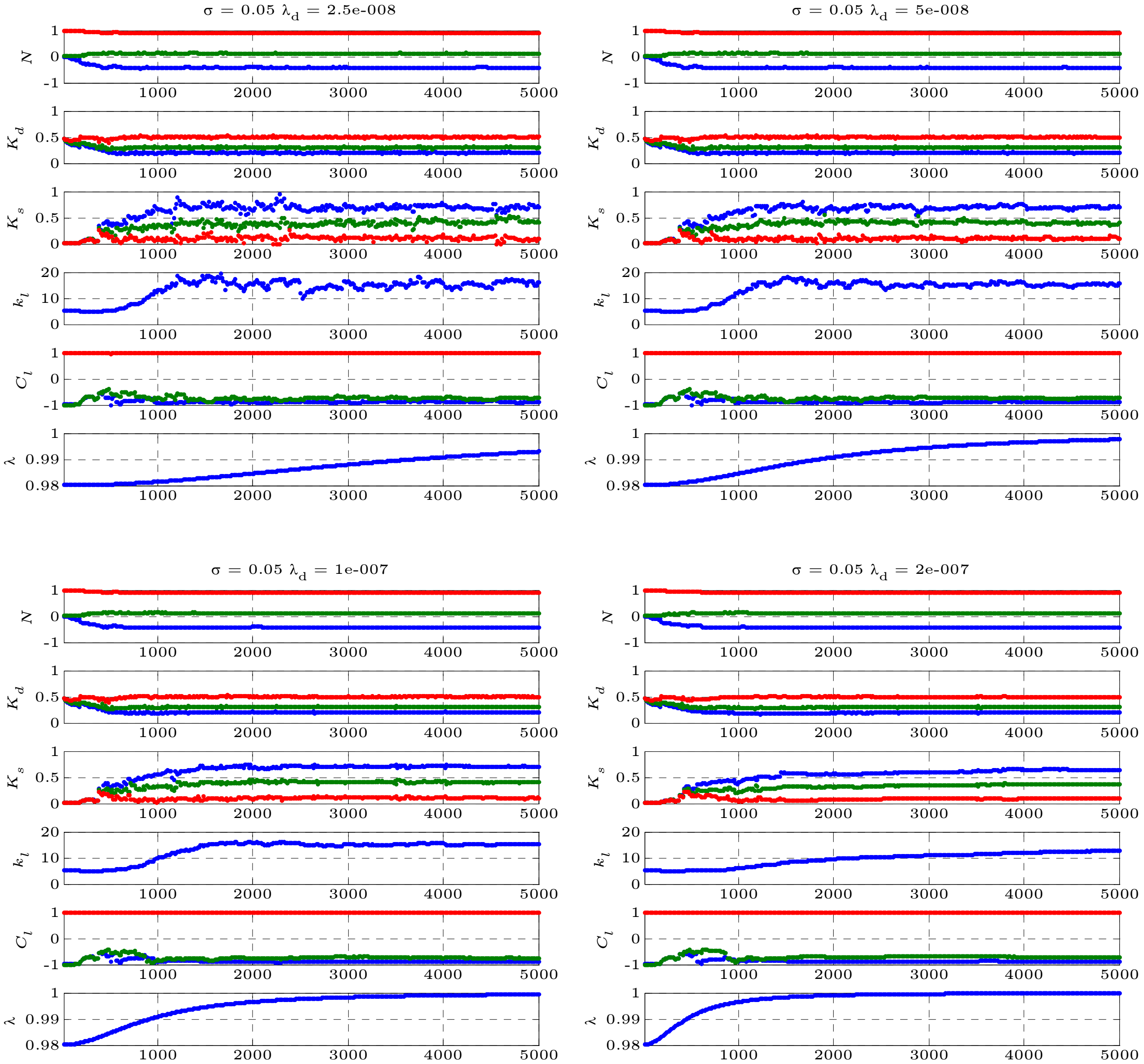

Figure 5.4: Incremental BRDF model optimization. Plots show the convergence for Lafortune model parameters for a single pixel using an artificial data set containing 10800 uniformly sampled and randomized data points. The parameters in each of the four plots listed from top to bottom are: the surface normal vector $\vec{N}$, the color diffuse term $\boldsymbol{K_d}$, the color specular term $\boldsymbol{K_s}$, the specular lobe exponent k_l, and the three specular lobe coefficients $\boldsymbol{C_l}$ (refer to Equation A.15). The bottom plot shows the convergence of the smoothing parameter λ. The four plots show the convergence for different decay rates of the smoothing parameter λ_d. Color terms are kept in the range $[0, 1]$.

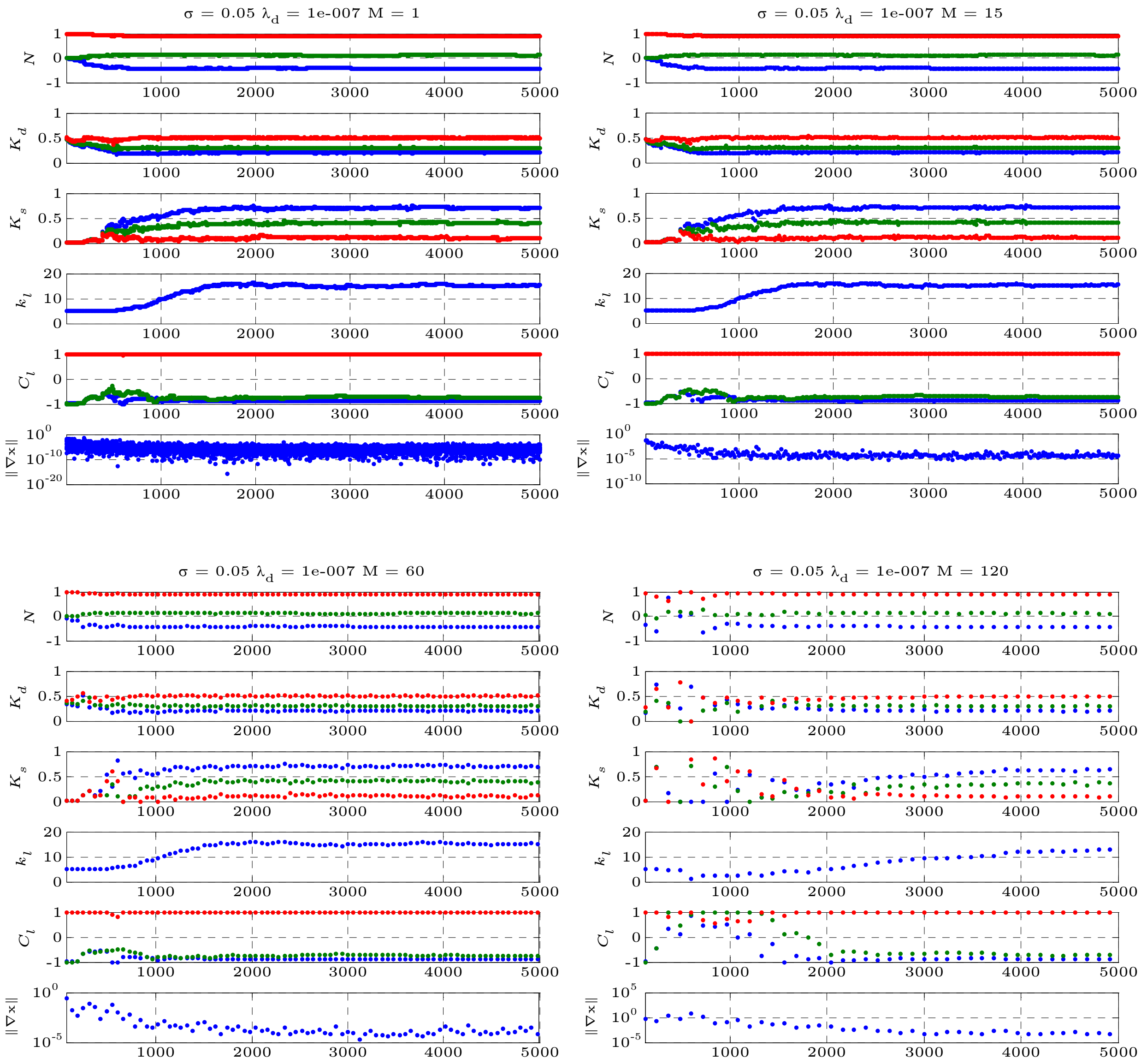

Figure 5.5: Incremental BRDF model optimization. Plots show the convergence for Lafortune model parameters for a single pixel using an artificial data set containing 10800 uniformly sampled and randomized data points. The parameters in each of the four plots listed from top to bottom are: the surface normal vector $\vec{N}$, the color diffuse term $\boldsymbol{K}_d$, the color specular term $\boldsymbol{K}_s$, the specular lobe exponent k_l, and the three specular lobe coefficients $\boldsymbol{C}_l$ (refer to Equation A.15). The bottom plot shows the convergence of the magnitude of the gradient vector $\|\nabla \mathbf{x}\|$. The four plots show the convergence for different parameter update periods M. Parameters are updated every M data vectors $\mathbf{x}_m$. Color terms are kept in the range $[0, 1]$.

and Figure 5.4 shows the effect of changing its value. Varying it affects the trade-off between algorithmic stability and the speed of convergence. Note, that for large values ($\lambda_d = 2 \cdot 10^{-7}$) the exponential specular parameter k_l does not reach a convergence. On the other hand, small values of λ_d produce unstable results.

Finally the update period of BRDF parameters is controlled by the parameter M. Its purpose is mainly to improve the computational performance of the algorithm, as we could potentially update parameters on every new data sample. However, updates are expensive, as they involve computation of the inverse of the gradient covariance matrix $\mathcal{C}$ (see Equation 5.11). Increasing the update period M can significantly improve the running performance of the algorithm. Figure 5.5 shows the effect of changing the update period on the convergence. Note that with $M = 15$ samples there is no visible degradation in convergence performance, while with an update period of $M = 60$ samples the algorithm becomes initially a little less stable. Doubling the period to $M = 120$ significantly degrades the stability.

5.3.4 Recovering the Attenuation Parameter

In general, the spatially varying attenuation parameter a_n is difficult to recover using the incremental solver directly. The problem is ill-posed, as multiple combinations of BRDF parameters and the attenuation can result in many local minima. Since Equation 5.6 in linear in the attenuation parameter, we will take advantage of it. We use the non-linear incremental solver to initially estimate the layer radiance $L_{l,n}$, while a_n is set to maximum opacity. We use this result to subsequently solve a linear problem of the form:

$$a_n = \underset{\alpha_2}{\operatorname{argmin}} \sum_{k=1}^{K} \Big[\alpha_1 L_{l,n,k} + \alpha_2 L_{o,n+1,k} - L_{o,n,k}\Big]^2 w_{n,k} \tag{5.21}$$

This equation can be interpreted as a blending optimization between two one-layer BTF representations at slightly different offsets. α_1 is a dummy blending factor for $L_{l,n}$ that we do not have to solve for. Once we solve Equation 5.21 for α_2, $L_{l,n}$ must be updated making α_1 redundant.

Recovering geometry from images can at most be guaranteed visual consistency with the originals. Recovering the attenuation parameter can be viewed as solving a blending equation for two BTF representations at different depth offsets. Yet visual consistency can often be achieved in a non-unique way, allowing multiple solutions. For example, a constant radiance offset present in the input data could be arbitrary distributed between layers. This is the case for flat, featureless texture regions in the BTF. Although visually consistent with the input data, unexpected artifacts may appear, when rendered for viewing directions that were not part of the input. Moreover, as the distance between layers is decreased, so will the differences in their appearance, until they become dominated by noise and re-projection error. Increasing the number of layers can therefore be beneficial only up to a certain point. We can conclude, that the resolution limit, defined as the minimum useful separation distance between layers, is directly related to the strength of the noise present in the BTF and inversely related to the local texture gradient measured against the relative warping displacement for that separation distance. Below this threshold, the attenuation parameter cannot be meaningfully estimated. Alternatively, we could modify the algorithm to achieve better consistency. For example, adding a smoothness constraint based upon the local neighborhood information, such as the surface gradient continuity, could lead to a more consistent solution and is an interesting area to explore.

5.3.5 Shadows and the Diffuse Radiance Term

Unlike some other phenomena, shadows cannot be easily parametrized. Shadows are created when light is attenuated on its way to a surface and depend on geometry. Although the LVST model stores spatially-varying attenuation parameters at each layer, shadows generated that way will rarely be consistent with the original shadows due to geometry being flattened into layers. Yet sharp shadow boundaries that do not match exactly may still look more "natural" to an observer. Following this argument, we model shadows using the actual attenuation values stored at each layer. As this is not consistent with least-squares optimization used to recover other parameters, shadowing is initially ignored and is only taken into account in the final pass of the algorithm.

Note, that in a shadowed area, the diffuse radiance $L_{d,n}$ becomes the primary source of illumination. In the final step of the algorithm we solve for the diffuse radiance term, which essentially serves as a correction factor for shadows. We again linearize Equation 5.6 solving for $L_{d,n}$:

$$\begin{aligned} L_{d,n} = \underset{\alpha}{\operatorname{argmin}} \sum_{k=1}^{K} w'_{n,k} \\ \left[f_{r,n} \left[L_{i,n,k} + \alpha \right] \vec{N}_n \cdot \vec{\omega}_i + L_{o,n+1,k} - L_{o,n,k} \right]^2 \end{aligned} \tag{5.22}$$

The main difference this time is that we only consider data points for which $L_{i,n,k} < 1$, i.e. at least some shadow is cast by layers above. Many other ways of rendering shadows are definitely possible and it can be an interesting area to investigate.

Summing it up, the algorithm consists of a series of local optimization steps that are repeated until convergence is achieved:

1. *Initialize all layers and make them fully transparent.*
2. *Starting from the deepest layer N, iterate over the layer stack (with no shadowing):*

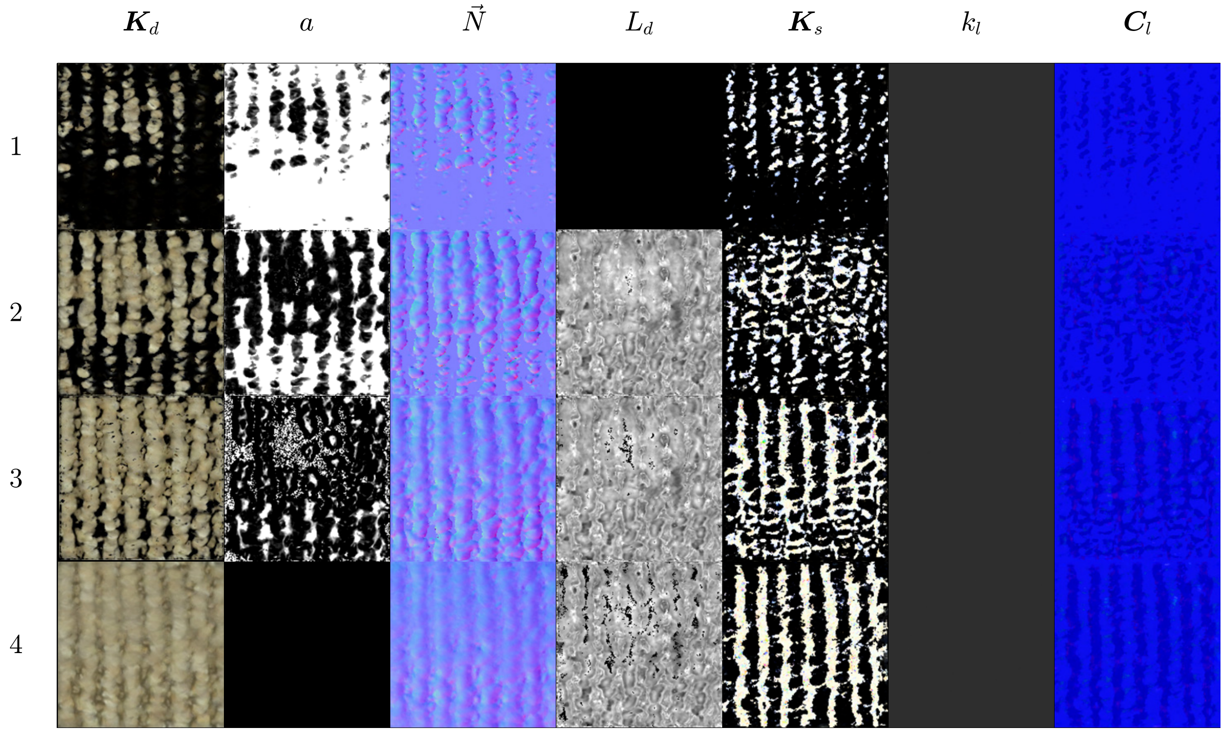

Figure 5.6: LVST parameter textures recovered from the carpet BTF set using the Lafortune reflectance model. Each row corresponds to one layer. Parameters are in columns. From left to right, they are: the color diffuse term $\boldsymbol{K}_d$, the attenuation coefficient a, the normal vector $\vec{N}$ (RGB color represents (x, y, z) vector components), the diffuse radiance factor L_d, the color specular term $\boldsymbol{K}_s$, the specular lobe exponent k_l, and the three specular lobe coefficients $\boldsymbol{C}_l$. Attenuation, diffuse radiance, and specular exponent are scalar values.

   (a) *Solve for the attenuation coefficient (Section 5.3.4):*
      i. *Generate an opaque layer for the current layer location.*
      ii. *Solve Equation 5.21 for the attenuation.*
   (b) *Estimate the BRDF and surface normal parameters (Section 5.3.2).*

3. *Solve for the diffuse radiance component and update the model for shadowing (Section 5.3.5).*

At the end, the algorithm produces a stack of multidimensional layers containing all LVST model parameters. An example of an LVST representation recovered using this algorithm is shown in Figure 5.6. Notice, that for the bottom layer (number 4) the attenuation is set for maximum opacity and that there is no diffuse radiance in the topmost layer. There is also very little visible differentiation in the specular lobe parameter, as it actually ended up very close to the initial value used by the algorithm.

5.4 Input BTF Data Sets

Textures used in the experiments come from the UIUC BTF database [130]. The database contains some complex volumetric surface textures and a few are shown in Figure 1.3. Each texture was sampled over a range of discrete viewing and lighting directions resulting in over 10,000 images.

The data required some additional post-processing, most importantly, inversion of the gamma correction that was automatically applied by the video camera that was used. Some image artifacts, which could not be corrected for, included video compression artifacts and camera noise. Errors in geometric alignment, including camera and light orientations, lens distortions, approximation of the perspective with parallel projections, and errors in the BTF rectification process were also problematic.

5.5 Results

Although the algorithm could be applied to BTFs with simple geometry, it does not provide any advantage over other compression schemes mentioned in Section 5.1.2. Since it was designed to specifically handle textures with complex volumetric structures, we have tested the algorithm on several captured BTFs with such characteristics. In addition to BTF data, the algorithm requires a choice of number of layers and the height bounds. These can be manually selected by examining the data. Since all layer updates are independent of each other in the (u, v) space, we take an advantage of this parallelism and solve for all layer parameters simultaneously. The number of iterations required before convergence conditions are detected depends on the complexity of the texture and the chosen BRDF, but is on the order of a few thousands

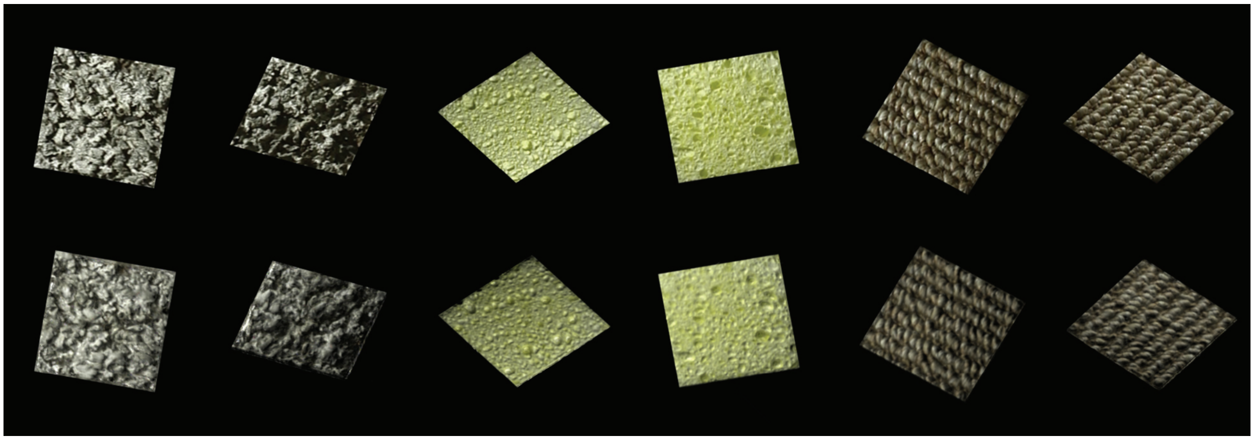

Figure 5.7: Images synthesized using the LVST model. The original BTF samples are at the top with corresponding synthesized equivalents at the bottom.

input BTF samples. Quite often convergence of a single layer is achieved before the majority of the input data is even evaluated, suggesting significant data redundancy. Depending on the number of layers and their dimensions, our C++ implementation typically takes a few hours to fully converge when run on a desktop PC.

Several examples of synthesized views along with the original data are shown in Figure 5.7. All textures in the figure have the same spatial resolution as the original BTF data (192x192, with a possible margin of a few pixels). They are modeled with four to five layers, except the lichen texture, which uses six layers. Since the algorithm relies on parametric modeling, it cannot represent phenomena not represented by the model. The resulting reconstructions are typically blurred versions of the originals. However, a major advantage of this representation is a more natural interpolation and even extrapolation from the original data.

The LVST representation can be used to render BTFs on polygonal meshes. Figure 5.8 demonstrates smooth changes in texture appearance on a smoothly-varying surface with tiled LVSTs. Please note silhouettes visible near the upper edge of the cylinder. To further improve the visual quality at grazing angles, the algorithm can be trivially extended to additionally construct vertical layers.

5.5.1 Compression and Real-Time Rendering

The result of the algorithm is a compact texture representation highly suitable for real-time applications. From the original sets of over 10,000 images, corresponding LVST representations are on the order of 10-20 textures. Such high compression ratio and easily parallelizable rendering process is perfectly suitable for graphics hardware acceleration. The rendering process simply involves evaluation of Equation 5.4. For each layer, we have to evaluate the BRDF model stored in several textures and modulate it with the shadow cast

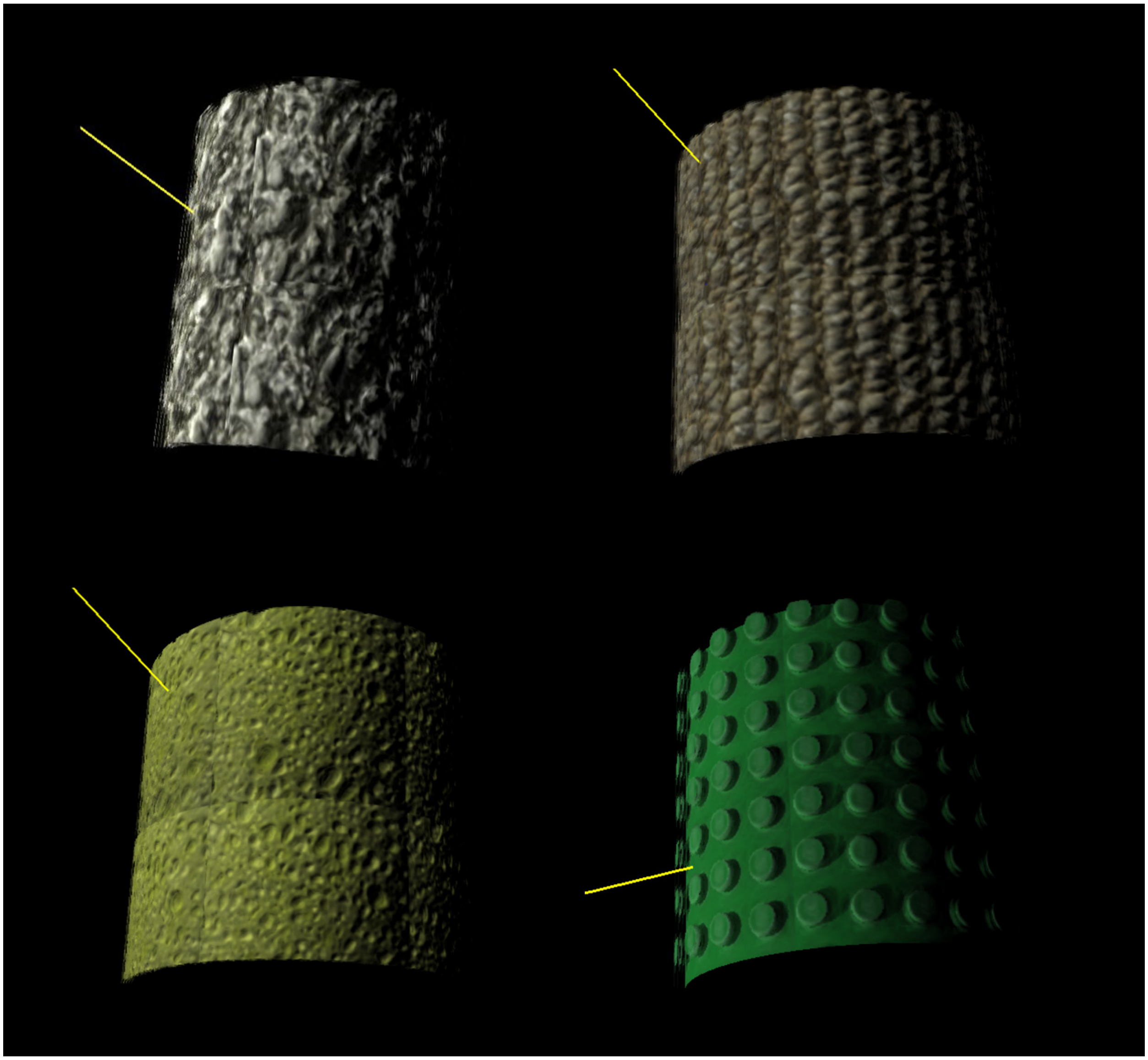

Figure 5.8: LVSTs tiled on a cylinder and rendered in real-time. The two examples in the left column were fitted using the diffuse model and the two on the right use the Lafortune model. The yellow line indicates the light direction. Note the 3-D texture on the silhouettes.

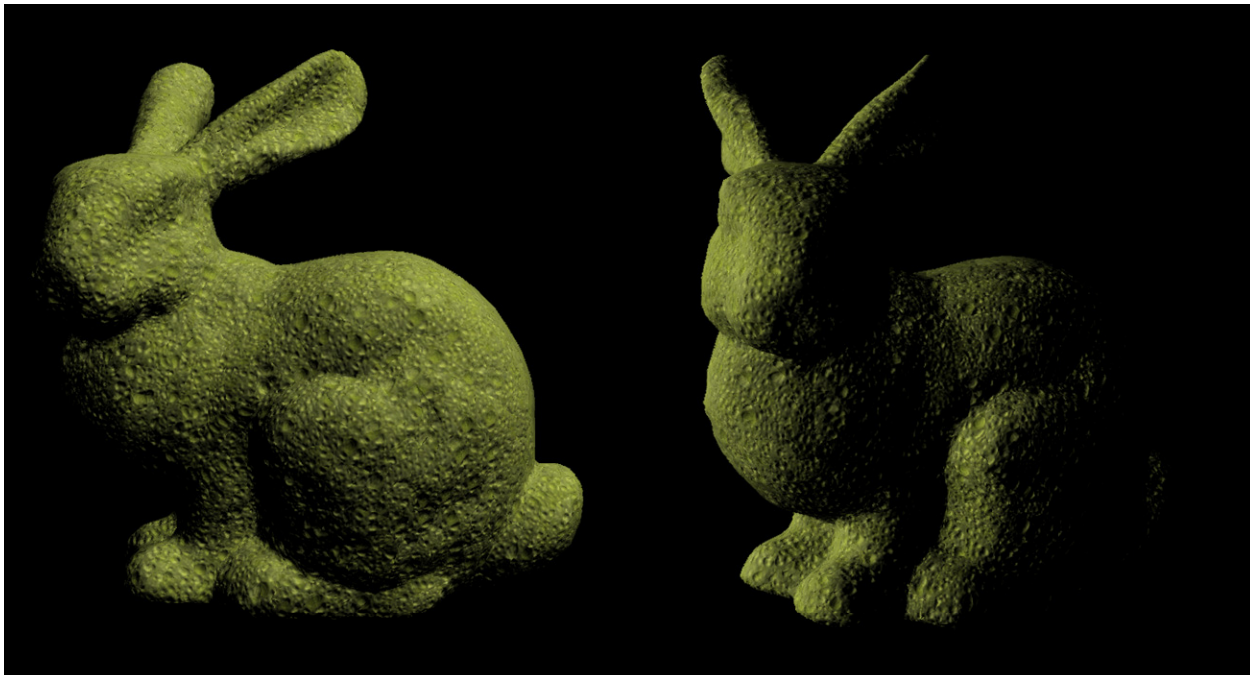

Figure 5.9: Two views of a mesh with a synthesized sponge texture lit and rendered in real-time.

by the layers above. In practice, it can be performed in a few simple steps:

1. *Starting from the bottommost layer N, evaluate the BRDF for the given light and camera directions.*
2. *Compute the shadow cast by all layers above the current layer and use it to compute the final layer radiance.*
3. *Render the layer into the screen buffer blending it according to the attenuation factor.*

These three steps can be evaluated using graphics hardware in a pixel shader. We have implemented a real-time application to demonstrate this. In a vertex program it computes layer offsets and evaluates the BRDF model in a fragment program. A patch-based surface texture synthesis algorithm [151] uses vectors of LVST parameters in a fashion similar to regular texture synthesis, that operates on vectors of RGB color values. In LVST synthesis, we have used vectors of concatenated values of attenuation coefficients a_n and some number of BRDF parameters from every layer. Typically the diffuse term $K_{d,n}$ alone was sufficient. Note, that the texture synthesis algorithm from [151] trades-off synthesis quality for speed and therefore we expect even better visual results from other synthesis algorithms.

In Figure 5.9 are two screen captures of a mesh lit and rendered in real-time with s synthesized sponge texture. Following figures show the results for several other textures: lichen (Figure 5.10), carpet (Figure 5.11),

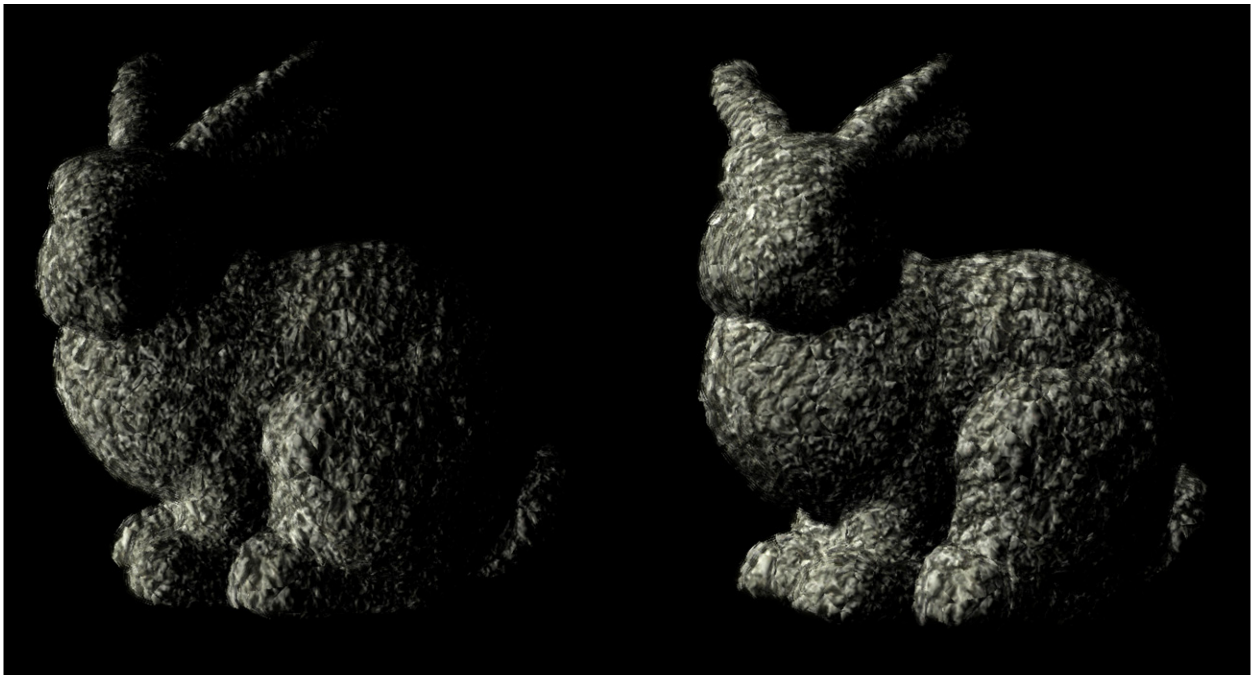

Figure 5.10: Two views of a mesh with a synthesized lichen texture lit and rendered in real-time.

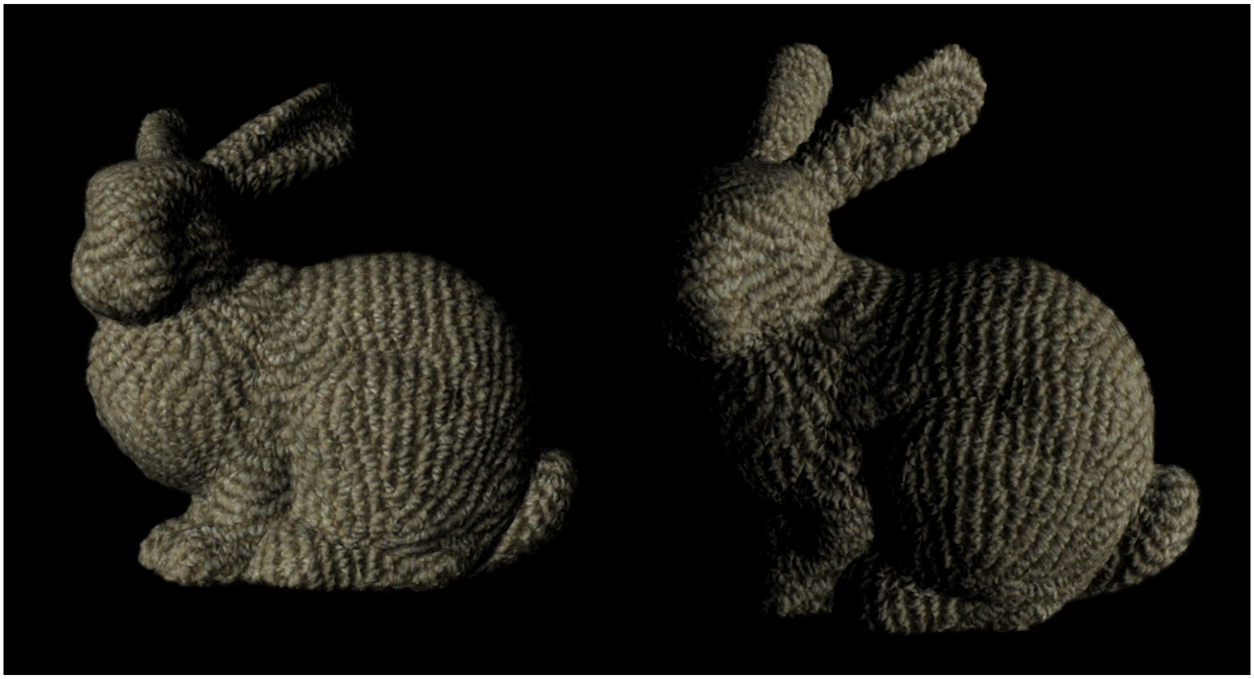

Figure 5.11: Two views of a mesh with a synthesized carpet texture lit and rendered in real-time.

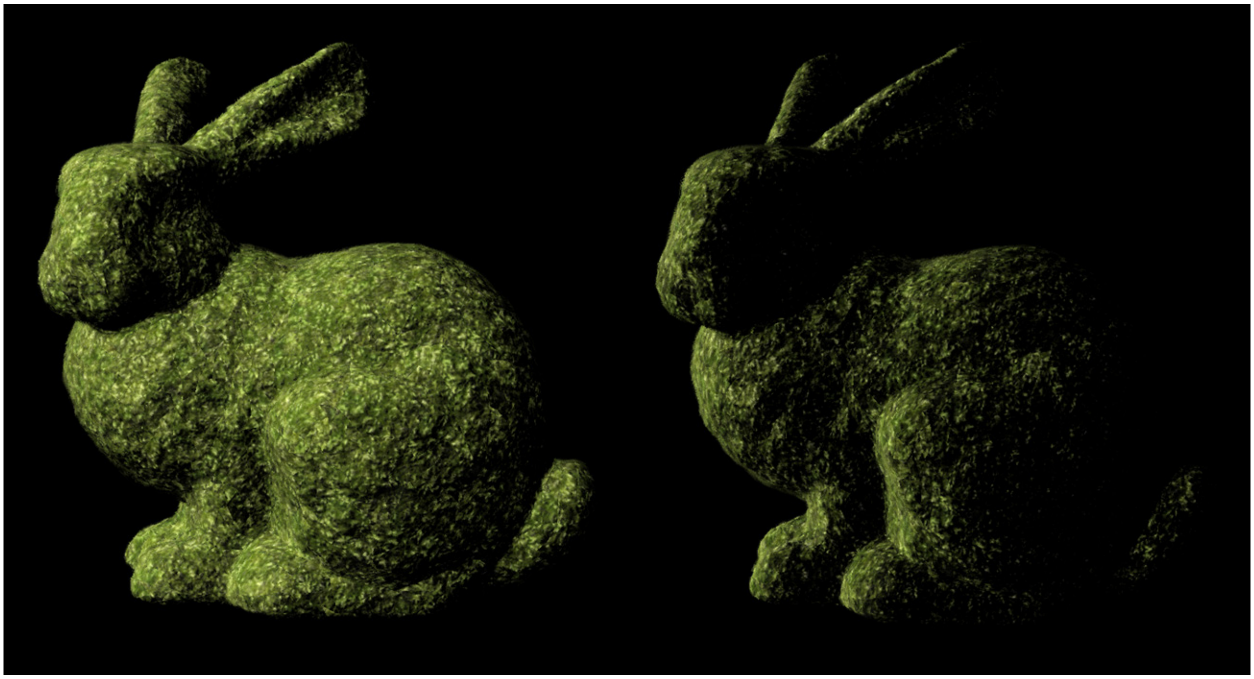

Figure 5.12: Two views of a mesh with a synthesized moss texture lit and rendered in real-time.

and moss (Figure 5.12). The carpet and the moss textures use the Lafortune model.

Chapter 6

Conclusions

The research work presented here results in two novel algorithms for reconstruction of 3-D geometry with complex reflectance observed in non-diffusive media. Although both algorithms use similar assumptions, they are intended for two different purposes. The first algorithm presented in Chapter 4 is intended for accurate reconstruction of large-scale opaque geometry. The second algorithm presented in Chapter 5 is designed to capture the appearance of small-scale geometry and reflectance of volumetric surface materials. When used in conjunction, they could provide an effective framework for capturing different aspects of a real scene.

6.1 Discussion

6.1.1 The Large-Scale Geometry Reconstruction Algorithm

The photometric ranging algorithm presented in Chapter 4 is designed for surfaces with complex, spatially varying, and unknown reflectance properties. Photometric ranging exploits the law of radiant energy density falloff with distance from a point light source and, unlike other photometric methods, the depth is estimated directly and independently at each pixel using reflectance field alignment, rather than through surface normal field integration. Using the method of double-covering the incident light in a spherical configuration, the method have been shown to work for a variety of surfaces with unknown and varying non-Lambertian reflectance. Relatively dense sampling typically required of this method makes it well suited for image-based rendering and allows for image synthesis with lights at finite distances, something that is not feasible using existing image-based relighting techniques. As synthesized examples have shown, this method could be applied directly in rendering applications.

Several aspects of this method could benefit from a deeper investigation. Experiments have shown that achieving sub-pixel registration is of utmost importance, otherwise camera shifts over brightness discontinuities can significantly bias brightness shifts due to the changing distance to the light source. Further investigation is required on how to achieve an improved registration, either through a better mechanical design, or better registration algorithms. Also, the benefits of increasing the dynamic range of the data by capturing high dynamic range images, or using the direct illumination separation algorithm proposed by Nayar et al. [168] to eliminate interreflections have not been verified experimentally. In Section 4.3.2 an observation was made that only changes of the reflected radiance resulting from the light movement along the camera ray affect the depth measured from the camera view. This would argue for positioning a light coincidental with the camera center. It could possibly be best accomplished with a half-silvered mirror positioned at a 45° angle in the path of the camera ray and reflecting light from a source into the path. Moving the light perpendicular to the camera ray would then be equivalent to moving the light along the ray.

On a higher level, double-covering the incident lighting may contain additional information that the algorithm is currently not taking advantage of. It is not clear, for example, if that could be used to simultaneously recover the geometry of the scene, surface normals, and light locations. Finally, this method shows a promise for a way to effectively sample the full 4-D BRDFs of objects composed of several distinct materials, similar to Goldman et al. [78], but without a need for an explicit BRDF model.

6.1.2 The Small-Scale Volumetric Appearance Modeling Algorithm

The algorithm for appearance capture and modeling of volumetric textures proposed in Chapter 5 works by separating the geometric information from the reflectance data present in a bidirectional texture function. It is especially destined for volumetric surface textures with complex, opaque microstructure, shadowing, and unknown reflectance properties. These types of textures are especially difficult to compress and reconstruct by existing BTF compression schemes. By contrast, the layered representation naturally offers very high compression of the original data and is fit for real-time rendering applications. On top of that, by providing a geometric representation, it allows for realistic rendering for textures at grazing angles. Even higher coherence of the rendered texture could be achieved by trivially extending the algorithm to re-sample the BTF into vertical layers. At a negligible additional storage cost, textures could then be fully rendered at any orientation to the viewer.

Although we have applied the underlying algorithm to decomposition of BTFs into layers, its applications could potentially be useful in other areas. The concept of recursively updating data slices could be used in other volume reconstruction applications, where transparency, shadowing, and complicated reflectance properties play an important role. Finally, it could likely be extended to voxelized representations as well.

6.2 Final Remarks

In recent years the prospect of reconstructing geometry with complex reflectance has become very appealing. As the rendering capabilities of graphics hardware continue to soar, the demand for reconstruction algorithms that can handle such geometry continues to grow as well. While a number of algorithms has been recently proposed, there is still no universal solution for scanning of arbitrary scenes. Imagine for example a foggy scene at a lake. Complex reflectance of plant leaves, water with its transparency and refractions, diffusive fog, complex large-scale geometry of trees as well as small-scale geometry of bark and moss. All the complexity occurs simultaneously. Although currently we are not yet capable of digitizing a lake scene, the recent progress of research gives a hope that it will be possible in the future. The two algorithms proposed here hopefully bring us closer to that ultimate goal.

Appendix A

Basic Radiometric Concepts

The field of *radiometry* is dedicated to measuring and describing various properties and interactions of light, including the infrared and the ultraviolet radiation. Roughly, it spans the electromagnetic spectrum from 10 to 10^6 nm. Note that the visible light is only in the range of about 360 to 830 nm. There is also a related field of *photometry*, which deals with the perception of light by the human eye. It takes into account the eye's wavelength-dependent sensitivity, rather than the actual power. This work is based on radiometric principles.

One basic property used to describe light is its *radiant energy* Q_e. The amount of energy transmitted per unit time is called the *radiant power*:

$$\Phi = \frac{dQ_e}{dt} \; [\mathrm{W} = {}^{\mathrm{J}}\!/_{\mathrm{s}}] \tag{A.1}$$

Radiant power of a surface is described in terms of the area power density W. It is a measure of the amount of energy moving through an imaginary unit area in space per unit time. Depending on whether the energy is seen as arriving at or leaving a surface, its respective power area density is called either *irradiance* E, or *radiosity* B:

$$W \equiv E \equiv B = \frac{d\Phi}{dA} \; [{}^{\mathrm{W}}\!/_{\mathrm{m}^2}] \tag{A.2}$$

A.1 Solid Angles

The distribution of radiant power at some point, either emitted or received, usually requires a radial description. Power density is then a measure of the energy flux passing through a *solid angle* per unit time.

Solid angle is a 2-dimensional angular measure of a surface area of an area projection onto a unit sphere. It is usually easier to think about it with an analogy of an angle in 2-dimensions. An angle can be defined as ratio of the length of a circular arc l to the radius r of the circle:

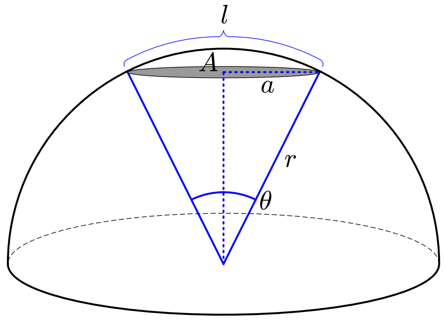

Figure A.1: Solid angle approximation.

$$\theta = \frac{l}{r} \equiv 2\pi\alpha \ \ [\mathrm{rad}] \tag{A.3}$$

α represents the fraction of the circle's circumference. Equivalent solid angle definition represents the ratio of an area of a spherical patch S to the squared radius of a sphere r^2:

$$\omega = \frac{S}{r^2} \equiv 4\pi\alpha \ \ [\mathrm{sr}] \tag{A.4}$$

Here α represents the fraction of the surface area of the sphere. Solid angles are typically used in a vector representation as a *solid angle vector*: $\vec{\omega}(\omega, \theta, \phi)$, where (θ, ϕ) designate the direction in spherical coordinates. ω is the solid angle centered along that direction.

Angular power density is very convenient for describing the power radiated by a point light source. For an ideal point light source the angular power density is identical in all directions:

$$\mathcal{I} = \frac{d\Phi}{d\vec{\omega}} \ \ [{}^{\mathrm{W}}\!/_{\mathrm{sr}}] \tag{A.5}$$

The angular power density $\mathcal{I}$ is often called the source *intensity*. Note that intensity is often a casually overused term to mean many different things, including the total power of a light source, or the brightness of a pixel recoded by a digital camera.

In radiometric calculations it fequently becomes necessary to determine the solid angle subtended by some distant surface patch. For small surface patches the subtended solid angle can be typically approximated by

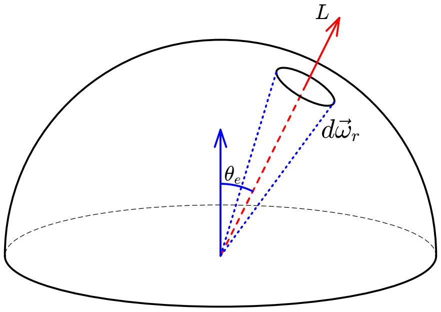

Figure A.2: The concept of radiance.

the area of the patch itself:

$$\omega = \frac{S}{r^2} \approx \frac{A}{r^2} \tag{A.6}$$

Here A refers to the area of the patch. For a distant circular patch of radius a, it can approximated further as:

$$\omega \approx \frac{\pi a^2}{r^2} = \frac{\pi\Big(r\sin(\frac{\theta}{2})\Big)^2}{r^2} \approx \frac{\pi}{4}\theta^2 \tag{A.7}$$

where θ is the angle subtended by the patch as shown in Figure A.1. Equation A.7 can be used as long as θ is small, so that the approximation $\sin(\theta) \approx \theta$ can be used.

A.2 Radiance

One of the most useful terms in radiometry is *radiance*. It is commonly used to describe the transport of radiant power between surfaces. For example, a surface radiates $B = \frac{d\Phi}{dA}$ watts per unit area and we would like to know how much of that power is received by another surface at some distance away (see Figure A.2). Radiance measures the radiant power density leaving, or arriving at a surface per unit projected *solid angle*:

$$L = \frac{dB}{d\vec{\omega}_r \cos(\theta_e)} \equiv \frac{dE}{d\vec{\omega}_e \cos(\theta_r)} \quad [\mathrm{W/m^2 \cdot sr}] \tag{A.8}$$

θ_e denotes the angle between the emitter surface normal and the receiver surface solid angle direction $\vec{\omega_r}$. θ_r and $\vec{\omega}_e$ have equivalent meaning in the irradiance equation. The usefulness of radiance comes from its solid angle weighting. Radiance is independent of the distance between surfaces [1]. It can be thought of as a measure of radiant power transmitted along light rays. This fundamental *ray law* [211] is the core of all raytracing and radiosity algorithms in computer graphics.

A.3 Bidirectional Distribution Functions

When light hits a surface, it is either absorbed, or scattered. This transfer of light energy or radiance is typically characterized using a *bidirectional distribution function* (BDF) [89]:

$$f(\vec{\omega}_i \to \vec{\omega}_o, \lambda) = \frac{dL_{dE(\vec{\omega}_i,\lambda)}(\vec{\omega}_o, \lambda)}{dE(\vec{\omega}_i, \lambda)} \;\; [{}^1\!/\mathrm{sr}] \tag{A.9}$$

BDF defines the spherical distribution of the radiance (parametrized by $\vec{\omega}_o$) scattered by a surface due to the irradiance arriving from some solid angle vector $\vec{\omega}_i$ at a wavelength λ. The wavelength dependance can often be ignored. In general, BTF can also be used to model light polarization and wavelength shifts. For a broader model refer to Equation 2.34.

BDF combines the effects of both reflection and transmisison. For a solid material, for which only reflection plays a role, BTF reduces to a *bidirectional reflectance distribution function* (BRDF) [171]. Similarly transmission of light can be described by a *bidirectional transmission distribution function* (BTDF). BTF can therefore be viewed as a combination of BRDF and BSDF.

BTF becomes useful in computations involving transfer of light. For example, the total radiance scattered from a point in a particular direction is calculated by integrating the BTF for all incoming irradiance:

$$L(\vec{\omega}_o, \lambda) = \int_{\Omega_i} f(\vec{\omega}_i \to \vec{\omega}_o, \lambda) E(\vec{\omega}_i, \lambda)\, d\vec{\omega}_i \tag{A.10}$$

This equation is commonly referred to as the *reflectance equation* [40].

A.3.1 Reflectance and Transmittance

Although radiance distribution functions are very convenient in radioimetric equations, they are very inconvenient to measure directly. They describe radiance over infinitesimal solid angles, that can only be approximated by inherently finite measurements. Reflection measurements typically use a related quantity

---

[1] This is true only if we assume a vacuum environment.

called *reflectance* [2] [171]. It simply specifies what fraction of the incident flux incoming over some finite solid angle is reflected into another solid angle:

$$\rho(\vec{\Omega}_i \to \vec{\Omega}_o) = \frac{\int_{\vec{\Omega}_o} E(\vec{\omega}_o)\, d\vec{\omega}_o}{\int_{\vec{\Omega}_i} E(\vec{\omega}_i)\, d\vec{\omega}_i} = \frac{\int_{\vec{\Omega}_o} \int_{\vec{\Omega}_i} f(\vec{\omega}_i \to \vec{\omega}_o) E(\vec{\omega}_o)\, d\vec{\omega}_i\, \cos(\theta_o)\, d\vec{\omega}_o}{\int_{\vec{\Omega}_i} E(\vec{\omega}_i)\, d\vec{\omega}_i} \tag{A.11}$$

θ_o denotes the direction of the outgoing solid angle vector $\vec{\omega}_o$. Typical reflectance measurements use a calibrated beam of light uniform over the entire measurement range of directions. The incoming radiance therefore is a constant and the Equation A.11 reduces to:

$$\rho(\vec{\Omega}_i \to \vec{\Omega}_o) = \frac{1}{\int_{\vec{\Omega}_i} d\vec{\omega}_i} \int_{\vec{\Omega}_o} \int_{\vec{\Omega}_i} f(\vec{\omega}_i \to \vec{\omega}_o) \cos(\theta_i) \cos(\theta_o)\, d\vec{\omega}_i\, d\vec{\omega}_o \tag{A.12}$$

Note that any measurement of reflectance should carry a description of the setup that includes the light beam configuration and the solid angle of the sensor.

A.3.2 Some Popular BRDF Models

The simplest BRDF model assumes the light reflects uniformly in all directions:

$$f(\vec{\omega}_i \to \vec{\omega}_o, \lambda) = \frac{K_d(\lambda)}{\pi} \tag{A.13}$$

$K_d(\lambda)$ is a wavelength-dependent reflectance coefficient. Surprisingly, many surfaces commonly referred to as “matte” behave this way. Other surfaces can often be approximated using this model as well. Due to its simplicity, it is widely used in computer vision algorithms. The model is called Lambertian after Johann Heinrich Lambert, who first mathematically described this phenomena in his book “Photometria”, published in 1760.

A more complex model aims to account for Lambertian and also for specular reflections. It is based on an observation that many surfaces posses mirror-like characteristics by reflecting the illumination about the surface normal. The reflection intensity falloff away from the reflection direction can be approximated by a function:

$$f(\vec{\omega}_i \to \vec{\omega}_o, \lambda) = \frac{K_d(\lambda)}{\pi} + K_s(\lambda)\Big(\vec{\omega}_o^{\mathbf{T}}(2\vec{n}\vec{n}^{\mathbf{T}} - \mathbf{I})\vec{\omega}_i\Big)^k \tag{A.14}$$

---

[2] There is also an equivalent term for refraction called *transmittance*.

where $K_s(\lambda)$ is the specular coefficient. $2\vec{n}\vec{n}^{\mathbf{T}} - \mathbf{I}$ is the reflection matrix for the light direction $\vec{\omega}_i$ about the surface normal $\vec{n}$. The exponent k is the "shininess" parameter that controls the width of the specular peak. For a perfect mirror, k approaches infinity. The model is known as the Phong model after Bui Tuong Phong who proposed the model [181].

Phong model in its simplicity is not very good at representing more complicated BRDFs with off-specular peaks and anisotropic behavior. A more generalized form of the Phong model developed by Lafortune et al. [135] allows for much greater flexibility. The basic idea is to represent the BRDF as a sum of L lobes, which are not necessarily symmetric:

$$f(\vec{\omega}_i \to \vec{\omega}_o, \lambda) = \frac{K_d(\lambda)}{\pi} + \sum_{l=1}^{L} K_{s,l}(\lambda)(\vec{\omega}_o^{\mathbf{T}}\mathbf{C}_l(\lambda)\vec{\omega}_i)^{k_l} \tag{A.15}$$

Here $\mathbf{C}_l(\lambda)$ is a 3×3 matrix[3] that relates the incoming and outgoing directions. It generalizes the Phong reflection matrix $2\vec{n}\vec{n}^{\mathbf{T}} - \mathbf{I}$. However, in many cases, not all 9 parameters of $\mathbf{C}_l$ are necessary. Except for very complex anisotropic BRDFs, only the three diagonal parameters usually need to be non-zero, while isotropic BRDFs require only two independent diagonal parameters [135].

There is a large number of other more complex BRDF models in existance, for example: *Ward*, *Torrance-Sparrow*, *Cook-Torrance*, *Minnaert*, *Oren-Nayar*, *Schlick*, etc. For details on some of these, refer to the book by Glassner [77].

A.3.3 Scattering Phase Function

BTFs are very convenient in describing material interfaces, where the surface normal can be easily identified. In scattering media such reference surface does not exist. Instead, the distribution of scattered light at any point in a medium is defined using a *scattering phase function* (SPF) [213, 103]. When normalized, this unitless function can be interpreted as a probability density function:

$$\int_{4\pi} p(\mathbf{r}, \vec{\omega}_i \to \vec{\omega}_o, \lambda)\, d\vec{\omega}_i = 1 \tag{A.16}$$

A SPF specifies the directional distribution of scattered light with a wavelength λ that arrived from a direction $\vec{\omega}_i$. In general, this definition can also be expanded to include phase (wavelength) shifts.

In most homogenous materials (for example: air, water, or milk), SPF is isotropic with respect to the incoming direction and can be parametrized simply with the cosine of the angle between the incoming and outgoing directions. One notable exception are crystals, where SPF is usually defined with respect to the

---

[3] Note that the notation used here expresses $\vec{\omega}$ in Cartesian coordinates (x, y, z).

crystal lattice planes.

Appendix B

First-Order Differential Equations

Consider a first-order differential equation of the form:

$$y'(x) + a(x)\, y(x) = b(x) \tag{B.1}$$

We would like to solve this equation for $y(x)$ over the interval $[x_0, x_1]$. The key for the solution is to find an *integrating factor* [24] $f(x)$, such that, when multiplied with the the left side of the equation B.1, it can be reduced as shown:

$$f(x)\big[y'(x) + a(x)\, y(x)\big] = \big[f(x)\, y(x)\big]' \tag{B.2}$$

We can solve for $f(x)$ by expanding both sides using the rules of differentiation:

$$\begin{aligned}
f(x)\, y'(x) + f(x)\, a(x)\, y(x) &= f'(x)\, y(x) + f(x)\, y'(x) \\
f(x)\, a(x)\, y(x) &= f'(x)\, y(x) \\
f(x)\, a(x) &= f'(x) \\
\int_x^{x_1} \frac{1}{f(u)} f'(u)\, du &= \int_x^{x_1} a(u)\, du \\
\ln f(x) &= \int_x^{x_1} a(u)\, du \\
f(x) &= \exp\Big(\int_0^{\|x_1 - x\|} a(u)\, du\Big) \qquad \text{(B.3)}
\end{aligned}$$

Therefore we can simplify B.1 by multiplying both sides by $f(x)$ and using the reduction B.2:

$$\begin{aligned} f(x)\big[y'(x)+a(x)\,y(x)\big] &= f(x)\,b(x) \\ \big[f(x)\,y(x)\big]' &= f(x)\,b(x) \\ \int_{x_0}^{x}\big[f(v)\,y(v)\big]'dv &= \int_{x_0}^{x} f(v)\,b(v)\,dv \\ f(v)\,y(v)\big|_{x_0}^{x} &= \int_{x_0}^{x} f(v)\,b(v)\,dv \\ f(x)\,y(x)-f(x_0)\,y(x_0) &= \int_{x_0}^{x} f(v)\,b(v)\,dv \qquad \text{(B.4)} \end{aligned}$$

The solution is especially simple when solving for $y(x_1)$:

$$y(x_1) = f(x_0)\,y(x_0) + \int_{x_0}^{x_1} f(v)\,b(v)\,dv \qquad \text{(B.5)}$$

Appendix C

Light Sensing with Digital Cameras

A digital camera is an image sensing device that takes advantage of the *photoelectric effect* to convert light into electric charge. When a photon strikes the surface of a sensor, it can sometimes knock-out an electron by transferring its energy into the kinetic energy of the electron. This can build up a charge over time on the sensor, that act as a capacitor cell. A number of these capacitor cells arranged in an array make up a picture. Currently the most common sensor types are either a charge-coupled device (CCD) or a complementary-metal-oxide-semiconductor (CMOS) [138]. While CCDs are typically more sensitive to light than CMOS based sensors, they are also more expensive than the CMOS chips.

To create an image in a digital camera the sensor is placed behind an optical lens. The irradiance falling on a photosensitive element depends on the lens geometry and can be expressed as [96]:

$$E_o = L_o \frac{\pi}{4} \left(\frac{D}{f}\right)^2 \cos(\alpha)^4 \tag{C.1}$$

where D and f are the lens diameter and the focal length. α is the angle that the pixel's principal ray makes with the optical axis of the camera. Assuming, that the image plane of the camera is perpendicular to the camera view vector the measured pixel brightness value I corresponds to the total pixel irradiance E_o received over a time t. I is related to the irradiance through some unknown transfer function $\mathcal{T}$:

$$I = \mathcal{T}\Big(E_o\, t\Big) \tag{C.2}$$

The brightness I can be expressed directly in terms of the incoming radiance L_o, subject to exposure e [160]:

$$I = \mathcal{T}\Big(L_o k\, e\Big) \tag{C.3}$$

where:

$$k = \frac{\cos(\alpha)^4}{f^2} \;\; e = \frac{\pi d^2}{4} t$$

The exposure is an adjustable value representing the amount of light that reaches the photosensitive cell over a period of time. It can be controlled by changing the aperture d and/or the acquisition time t. The exposure is usually set to allow the photosensitive element capture the maximum amount of information while minimizing the noise.

C.1 Nonlinearity

In order to measure the irradiance, it is necessary to recover the inverse transfer function $\mathcal{T}^{-1}$. Making a logical assumption that the response curve is monotonically increasing, we can rewrite Equation C.2 as:

$$\begin{aligned} \lg E_o &= \lg \mathcal{T}^{-1}(I) - \lg t \\ \lg E_o &= \mathcal{G}(I) - \lg t \end{aligned} \tag{C.4}$$

The range of brightness values I is limited by the maximum capacitance of a sensor cell and the resolution of the analog to digital conversion.

One way to solve for the unknown function $\mathcal{T}^{-1}$ is to setup an optimization problem that minimizes an objective function over a discrete range of exposure values [160, 56]:

$$\mathcal{O} = \sum_{i=1}^{M} \sum_{j=1}^{N} \left[\mathcal{G}(I_{ij}) - \lg E_i - \lg t_j\right]^2 + c \sum_{i=I_{min}}^{I_{max}} \mathcal{G}''(i)^2 \tag{C.5}$$

M is the number of chosen pixel locations and N is the number of exposures. The measurement data consists of a set of images of a static scene taken at varying exposures. Note the second smoothing term used to relate the discrete values of $\mathcal{G}$ to each other. For details, refer to the paper by Debevec and Malik [56], or Mitsunaga and Nayar [160].

C.2 Sensor Noise

As with any sensing device, the signal detected by a digital camera is corrupted by noise.The noise in digital cameras can be classified as either the *random noise*, or the *pattern noise* [185].

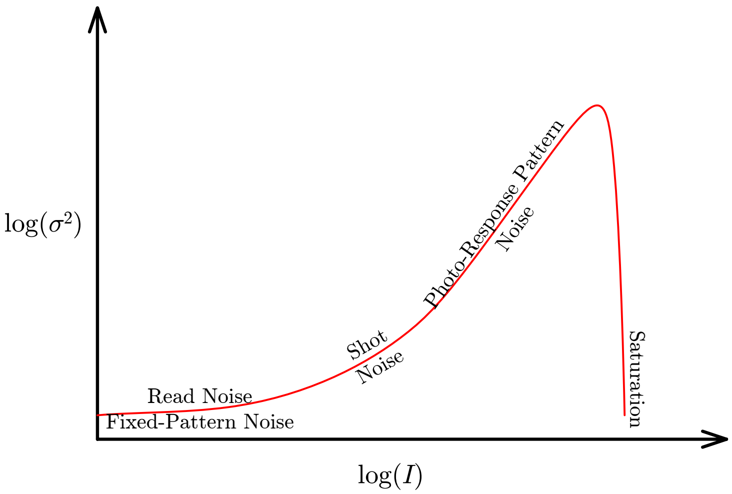

Figure C.1: Typical photon transfer curve for a digital camera showing the noise level for different brightness values up to the saturation point. The graph shows the dominant noise contributors for different sections on the curve.

C.2.1 Random Noise

This type of noise is the result of random processes occurring during image acquisition. It is often referred to as the *temporal noise*, since it's temporally random and is responsible for variations in images of an identical scene. The most important components of random noise is the *shot noise*, the *thermal noise*, and the *readout noise*. Its effects can be reduced by frame averaging.

C.2.2 Pattern Noise

Pattern noise comes from different defects and physical variations present in sensing cells across the sensor. It is therefore often referred to as the *spatial noise*. The most important components of the spatial noise are the *dark current non-uniformity* and the *photo response non-uniformity*. The effects of pattern noise can be partially eliminated by dark frame subtraction.

References

[1] Edward H. Adelson and James R. Bergen. *Computational Models of Visual Processing*, chapter The Plenoptic Function and the Elements of Early Vision, pages 3–20. MIT Press, 1991.

[2] Adept Technology, Inc. *AdeptOne-MV / AdeptThree-MV Robot, Rev C*, 1997.

[3] G. Agin and T. Binford. Computer description of curved objects. *Transactions on Computers*, C-25:439–449, 1976.

[4] Markus-Christian Amann, Thierry M. Bosch, Marc Lescure, Risto A. Myllylae, and Marc Rioux. Laser ranging: a critical review of unusual techniques for distance measurement. *Optical Engineering*, 40(1):10–19, 2001.

[5] Bjørn A. J. Angelsen. *Ultrasound imaging. Waves, signals and signal processing.*, volume I and II. Emantec AS, Trondheim, Norway, 2000.

[6] M. Asada, H. Ichikawa, and S. Tsuji. Determining surface orientation by projecting a stripe pattern. *Transactions on Pattern Analysis and Machine Intelligence*, 10(5):749–754, 1988.

[7] P. Aschwanden and W. Guggenbuhl. Experimental results from a comparative study on correlation-type registration algorithms. In *Robust Computer Vision*. Wickmann, 1993.

[8] Nicholas Ayache and Francis Lustman. Trinocular stereo vision for robotics. *Transactions on Pattern Analysis and Machine Intelligence*, 13(1):73–85, 1991.

[9] Réjean Baribeau, Marc Rioux, and Guy Godin. Color reflectance modeling using a polychromatic laser range sensor. *Transactions on Pattern Analysis and Machine Intelligence*, 14(2):263–269, 1992.

[10] Stephen T. Barnard and Martin A. Fischler. Computational stereo. *ACM Computing Surveys*, 14(4):553–572, 1982.

[11] A. R. Barringer. The use of audio and radio frequency pulses for terrain sensing. In *Selected papers on remote sensing of the environment*, pages 215–226. American Society of Photogrammetry, 1966.

[12] Svetlana Barsky and Maria Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1239–1252, 2003.

[13] M. Bashkansky, C. Adler, and J. Reintjes. Coherently amplified Raman polarization gate for imaging through scattering media. *Optics Letters*, 19:350–352, 1994.

[14] D. A. Benaron and D. K. Stevenson. Optical time-of-flight and absorbance imaging of biologic media. *Science*, 259:1463–1466, 1993.

[15] R. Berg, O. Jarlman, and S. Svanberg. Medical transillumination imaging using short-pulse diode lasers. *Applied Optics*, 32:574–579, 1993.

[16] Dimitri P. Bertsekas. Incremental least-squares methods and the extended Kalman filter. *SIAM Journal on Optimization*, 6:807–822, 1996.

[17] Dinkar N. Bhat and Shree K. Nayar. Ordinal measures for image correspondence. *Transactions on Pattern Analysis and Machine Intelligence*, 20(4):415–423, 1998.

[18] Frank Bignone, Olof Henricsson, Pascal Fua, and Markus A. Stricker. Automatic extraction of generic house roofs from high resolution aerial imagery. In *European Conference on Computer Vision*, volume 1, pages 85–96, 1996.

[19] Stan Birchfield and Carlo Tomasi. Depth discontinuities by pixel-to-pixel stereo. In *International Conference on Computer Vision*, pages 1073–1080, 1998.

[20] Franc̀ois Blais. Review of 20 years of range sensor development. *Journal of Electronic Imaging*, 13(1):231240, 2004.

[21] F. Bloch, W. W. Hansen, and M. Packard. The nuclear induction experiment. In *Physical Review*, volume 70, pages 474–485, 1946.

[22] Craig F. Bohren and Donald R. Huffman. *Absorption and Scattering of Light by Small Particles*. John Wiley & Sons, Inc., New York, NY, April 1998.

[23] Thierry Bosch and Marc Lescure, editors. *SPIE Milestone Series: Selected Papers on Laser Distance Measurements*, volume 115. SPIE Optical Engineering Press, 1995.

[24] William E. Boyce and Richard C. DiPrima. *Elementary Differential Equations*. John Wiley & Sons, Inc., New York, 7th edition, 2001.

[25] K.L. Boyer and A.C. Kak. Color-encoded structured light for rapid active ranging. *Transactions on Pattern Analysis and Machine Intelligence*, 9(1):14–28, 1987.

[26] G. Bradski and T.E. Boult. Guest editorial: Stereo and multi-baseline vision. *International Journal of Computer Vision*, 47(1-3):5–5, 2002.

[27] Myron Z. Brown, Darius Burschka, and Gregory D. Hager. Advances in computational stereo. *Transactions on Pattern Analysis and Machine Intelligence*, 25(8):993–1008, 2003.

[28] G. Brownell and H. Sweet. Localization of brain tumors with positron emitters. *Nucleonics*, 11:40–45, 1953.

[29] P. Bruscaglioni, G. Zaccanti, and Q. Wei. Transmission of a pulsed polarized light beam through thick turbid media: numerical results. *Applied Optics*, 32:6142–6150, 1993.

[30] E.C. Burrows and K.-Y. Liou. High resolution laser LIDAR utilizing two-section distributed feedback semiconductor laser as a coherent source. *Electronics Letters*, 26(9):577–579, 1990.

[31] Canon, Inc. *EOS-1D Digital*, September 2001.

[32] K. P. Chan, M. Yamada, B. Devaraj, and H. Inaba. Optical imaging through highly scattering media by use of heterodyne detection in the 1.3 μm wavelength region. *Optics Letters*, 20:492–494, 1995.

[33] B. Chance, E. Anday, S. Nioka, S. Zhou, H. Long, K. Wordenand C. Li, T. Turray, Y. Ovetsky, D. Pidikiti, and R. Thomas. A novel method for fast imaging of brain function, non-invasively, with light. *Optics Express*, 2:411–423, May 1998.

[34] Subramanyan Chandrasekhar. *Radiative Transfer*. Clarendon Press, Oxford, UK, 1950.

[35] H. Chen, Y. Chen, D. Dilworth, E. Leith, J. Lopez, and J. Valdmanis. Two-dimensional imaging through diffusing media using 150-fs gated electronic holography techniques. *Optics Letters*, 16:487–489, 1991.

[36] Yanyun Chen, Xin Tong, Jiaping Wang, Stephen Lin, and Baining Guoand Heung-Yeung Shum. Shell texture functions. In *SIGGRAPH*, pages 343–353, 2004.

[37] Wai-Fung Cheong, Scott A. Prahl, and Ashley J. Welch. A review of the optical properties of biological tissues. *Journal of Quantum Electronics*, 26(12):2166–2185, December 1990.

[38] James J. Clark. Active photometric stereo. In *Conference on Computer Vision and Pattern Recognition*, pages 29–34, 1992.

[39] James J. Clark and Holly Pekau. An integral formulation for differential photometric stereo. In *Conference on Computer Vision and Pattern Recognition*, volume 1, pages 119–124, 1999.

[40] Michael F. Cohen, John Wallace, and Pat Hanrahan. *Radiosity and realistic image synthesis.* Academic Press Professional, Inc., 1993.

[41] Arthur H. Compton. A quantum theory of the scattering of X-Rays by light elements. *Physical Review*, 21(5):483–502, 1923.

[42] J. C. Cook. Proposed monocycle-pulse VHF radar for airborne ice and snow measurements. In *Transactions on Communication and Electronics*, volume 79, pages 588–594. AIEE, 1960.

[43] Thomas T. Cormen, Charles E. Leiserson, and Ronald L. Rivest. *Introduction to algorithms.* MIT Press, Cambridge, MA, USA, 1990.

[44] Ingemar J. Cox, Sunita L. Hingorani, Satish B. Rao, and Bruce M. Maggs. A maximum likelihood stereo algorithm. *Computer Vision and Image Understanding*, 63(3):542–567, 1996.

[45] Antonio Criminisi, Ian D. Reid, and Andrew Zisserman. Single view metrology. *International Journal of Computer Vision*, 40(2):123–148, 2000.

[46] Crytek GmbH, http://www.crytek.com/technology/index.php?sx=eng2. *CryENGINE™ 2 Specifications*, 2006.

[47] Oana G. Cula and Kristin J. Dana. Compact representation of bidirectional texture functions. In *Conference on Computer Vision and Pattern Recognition*, volume 1, pages 1041–1047, 2001.

[48] M. Cutler. Transillumination as an aid in the diagnosis of breast lesions. *Surgery, Gynecology and Obstetrics*, 48:721–729, 1929.

[49] Raymond Damadian. Apparatus and method for detecting cancer in tissue. Technical report, The Patent Office, 1974.

[50] Kristin J. Dana, Bram van Ginneken, Shree K. Nayar, and Jan J. Koenderink. Reflectance and texture of real-world surfaces. In *SIGGRAPH*, pages 1–34, 1999.

[51] Katja Daubert, Hendrik P. A. Lensch, Wolfgang Heidrich, and Hans-Peter Seidel. Efficient cloth modeling and rendering. In *Eurographics Workshop on Rendering*, pages 63–70, London, UK, 2001. Springer-Verlag.

[52] Anthony B. Davis and Yuri Knyazikhin. *A Primer in Three-Dimensional Radiative Transfer*, chapter 3, pages 71–153. Springer Verlag, 2004 (in print).

[53] Louis-Victor de Broglie. *Recherches sur la théorie des Quanta.* PhD thesis, University of Paris, 1924.

[54] E. B. de Haller and C. Depeursinge. Resolution of time resolved breast transillumination. In *Advances in Optical Imaging and Photon Migration*, volume 21, pages 88–92. Optical Society of America, 1994.

[55] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *SIGGRAPH*, pages 145–156, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[56] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH*, pages 369–378, 1997.

[57] C. H. Desch, D. O. Sproule, and W. J. Dawson. The detection of cracks in steel by means of supersonic waves. *Journal of the Iron and Steel Institute*, 153:319, 1946.

[58] A. Dieckmann. FMCW-LIDAR with tunable twin-guide laser diode. *Electronics Letters*, 30(4):308–309, 1994.

[59] K. A. Dines. Imaging of ultrasonic reflectivity. presented at the Symposium on Computers in Ultrasound, Philadelphia, PA, 1985.

[60] Jean-Michel Dischler. Efficiently rendering macro geometric surface structures with bi-directional texture functions. In *Eurographics Workshop on Rendering*, pages 169–180. Springer, 1998.

[61] Robert William Ditchburn. *Light*. Dover Publications, Inc., 1991.

[62] James J. Duderstadt and William R. Martin. *Transport Theory*. John Wiley & Sons Inc., New York, 1979.

[63] Charles R. Dyer. *Foundations of Image Understanding*, chapter Volumetric Scene Reconstruction from Multiple Views, pages 469–489. Kluwer, Boston, 2001.

[64] I. Edler and C. H. Hertz. The use of ultrasonic reflectoscope for the continuous recording of movements of heart walls. *Kungliga Fysiografiska Sällskapets i Lund Förhandlingar*, 24(5):1–19, 1954.

[65] J. Ens and P. Lawrence. An investigation of methods for determining depth from focus. *Transactions on Pattern Analysis and Machine Intelligence*, 15(2):97–108, 1993.

[66] C.H. Esteban and F. Schmitt. Silhouette and stereo fusion for 3D object modeling. *Computer Vision and Image Understanding*, 96(3):367–392, 2004.

[67] G. W. Faris and M. Banks. Upconverting time gate for imaging through highly scattering media. *Optics Letters*, 19:1813–1815, 1994.

[68] Olivier Faugeras, Bernard Hotz, Herve Mathieu, Thierry Viéville, Zhengyou Zhang, Pascal Fua, Eric Théron, Laurent Moll, Gerard Berry, Jean Vuillemin, Patrice Bertin, and Catherine Proy. Real time correlation based-stereo: algorithm, implementations and applications. Technical Report 2013, INRIA, 1993.

[69] A. F. Fercher, K. Mengedoht, and W. Werner. Eye-length measurement by interferometry with partially coherent light. *Optics Letters*, 13:186–188, 1988.

[70] Jiri Filip and Michal Haindl. Non-linear reflectance model for bidirectional texture function synthesis. In *International Conference on Pattern Recognition*, volume 1, pages 80–83, Washington, DC, 2004. IEEE Computer Society.

[71] F. A. Firestone. The supersonic reflectoscope, an instrument for inspecting the interior of the solid parts by means of sound waves. *Journal of the Acoustical Society of America*, 17:287, 1946.

[72] J. B. Fishkin and E. Gratton. Propagation of photon-density waves in strongly scattering media containing an absorbing semi-infinite plane bounded by a straight edge. *Journal of the Optical Society of America A*, 10:127–140, 1993.

[73] Ryo Furukawa, Masaharu Harada, Yasuaki Nakamura, and Hiroshi Kawasaki. Synthesis of textures with intricate geometries using BTF and large number of textured micropolygons. In *International Workshop on Texture Analysis and Synthesis*, pages 77–82, 2005.

[74] Athinodoros S. Georghiades. Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In *International Conference on Computer Vision*, pages 816–823, 2003.

[75] A. P. Gibson, J. C. Hebden, and S. R. Arridge. Recent advances in diffuse optical imaging. *Physics in Medicine and Biology*, 50(4):R1–R43, 2005.

[76] R. G. Giovanelli. Radiative transfer in non-uniform media. *Australian Journal of Physics*, 12:164–170, June 1959.

[77] Andrew S. Glassner. *Principles of Digital Image Synthesis*. Morgan Kaufmann Publishers, Inc., 1994.

[78] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying BRDFs from photometric stereo. In *International Conference on Computer Vision*, pages I: 341–348, 2005.

[79] Gene H. Golub and Charles F. Van Loan. *Matrix Computations, Johns Hopkins Studies in Mathematical Sciences*. Johns Hopkins University Press, 1996.

[80] Steven J. Gortler, Radek Grzeszczuk, Richard Szeliski, and Michael F. Cohen. The lumigraph. In *SIGGRAPH*, pages 43–54, 1996.

[81] J. F. Greenleaf, S. A. Johnson, S. L. Lee, G. T. Herman, and E. H. Wood. Algebraic reconstruction of spatial distribution of acoustic absorption within tissue from their two dimensional acoustic projections. *Acoustical Holography*, 5:591–603, 1974.

[82] P. M. Griffin, L. S. Narasimhan, and S. R. Yee. Generation of uniquely encoded light patterns for range data acquisition. *Pattern Recognition*, 25(6):609–616, June 1992.

[83] W. E. L. Grimson. A computer implementation of a theory of human stereo vision. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, 292(1058):217–253, 1981.

[84] Michal Haindl, Jiří Filip, and Michael Arnold. BTF image space utmost compression and modeling method. In *International Conference on Pattern Recognition*, volume 3, pages 194–197, Washington, DC, 2004. IEEE Computer Society.

[85] J. C. Hebden, S. R. Arridge, and D. T. Delpy. Optical imaging in medicine I: Experimental techniques. *Physics in Medicine and Biology*, 42(5):825–840, 1997.

[86] J. C. Hebden, D. J. Hall, M. Firbank, and D. T. Delpy. Time resolved optical imaging of a solid tissue-equivalent phantom. *Applied Optics*, 35:8038–8047, 1995.

[87] J. C. Hebden, R. A. Kruger, and K. S. Wong. Time resolved imaging through a highly scattering medium. *Applied Optics*, 30:788–794, 1991.

[88] J. C. Hebden and K. S. Wong. Time resolved optical tomography. *Applied Optics*, 32:372–380, 1993.

[89] Paul S. Heckbert. Adaptive radiosity textures for bidirectional ray tracing. In *SIGGRAPH*, pages 145 – 154, 1990.

[90] M. R. Hee, J. A. Izatt, E. A. Swanson, and J. G. Fujimoto. Femtosecond transillumination tomography in thick tissues. *Optics Letters*, 18:1107–1109, 1993.

[91] Louis G. Henyey and Jesse L. Greenstein. Diffuse radiation in the galaxy. *Astrophysical Journal*, 93:70–83, 1941.

[92] G. T. Herman. *Image reconstruction from projections: The fundamentals of computerized tomography*. Academic Press, 1980.

[93] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying BRDFs. *Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, August 2005.

[94] Aaron Hertzmann and Steven M. Seitz. Shape and materials by example: A photometric stereo approach. In *Conference on Computer Vision and Pattern Recognition*, volume 1, pages 533–540, Jun 2003.

[95] C. K. Hitzenberger. Optical measurement of the axial eye length by laser Doppler interferometry. *Investigative Ophthalmology & Visual Science*, 32:616–624, 1991.

[96] Berthold K. P. Horn. *Robot Vision.* MIT Press, Cambridge, MA, 1986.

[97] Berthold K. P. Horn. Height and gradient from shading. *International Journal of Computer Vision*, 5(1):37–75, 1990.

[98] Berthold K. P. Horn and Brian G. Schunck. Determining optical flow. Technical report, Massachusetts Institute of Technology, Cambridge, MA, USA, 1980.

[99] Youichi Horry, Ken-Ichi Anjyo, and Kiyoshi Arai. Tour into the picture: Using a spidery mesh interface to make animation from a single image. In *SIGGRAPH*, 1997.

[100] G. Hounsfield. A method of an apparatus for examination of a body by radiation such as X-ray or gamma radiation. Technical Report 1283915, The Patent Office, 1972.

[101] D. H. Howry and W. R. Bliss. Ultrasonic visualization of soft tissue structures of the body. *Journal of Laboratory and Clinical Medicine*, 40:579–592, 1952.

[102] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinsonand W. Chang, M. R. Hee, T. Flotte, K. Gregory, and C. A. Puliafitoand J. G. Fujimoto. Optical coherence tomography. *Science*, 254:1178–1181, 1991.

[103] Hendrik C. van de Hulst. *Light scattering by small particles.* Dover Publications, New York, NY, 1982.

[104] D. C. D. Hung. 3D scene modeling by sinusoid encoded illumination. *Image and Vision Computing*, 11(5):251–256, 1993.

[105] M. Idesawa and T. Yatagai. 3-D shape input and processing by moirѐ technique. In *International Conference on Pattern Recognition*, pages 1085–1090, 1980.

[106] Katsushi Ikeuchi and Berthold K. P. Horn. Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 17(1-3):141–184, 1981.

[107] H. Inaba, M. Toida, and T. Ichimua. Optical computer-assisted tomography realized by Coherent Detection Imaging (CDI) incorporating laser heterodyne method for biomedical applications. *Optical Systems in Adverse Environments*, 1399:108–115, 1990.

[108] Akira Ishimaru. *Wave Propagation and Scattering in Random Media.* Academic Press, New York, 1978.

[109] Yuji Iwahori, Hidezumi Sugie, and Naohiro Ishii. Reconstructing shape from shading images under point light source illumination. In *International Conference on Pattern Recognition*, volume 1, pages 83–87, 1990.

[110] P. C. Jackson, P. H. Stevens, J. H. Smith, D. Kear, H. Key, and P. N. T. Wells. The development of a system for transillumination computed tomography. *British Journal of Radiology*, 60:375–380, 1987.

[111] J. R. Jago and T. A. Whittingham. Experimental studies in transmission ultrasound computed tomography. *Physics in Medicine and Biology*, 36(11):1515–1527, November 1991.

[112] J. A. Jalkio, R. C. Kim, and S. K. Case. Three dimensional inspection using multistripe structured light. *Optical Engineering*, 24:966–974, 1985.

[113] G. Jarry, S. Ghesquiere, J. M. Maarek, F. Fraysse, S. Debrayand M. H. Bui, and D. Laurent. Imaging mammalian tissues and organs using laser collimated transillumination. *Journal of Biomedical Engineering*, 6:70–74, 1984.

[114] H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson. Simultaneous reconstruction of optical absorption and scattering maps in turbid media from near-infrared frequency-domain data. *Optics Letters*, 20:2128–2130, 1995.

[115] Hailin Jin, Stefano Soatto, and Anthony J. Yezzi. Multi-view stereo beyond Lambert. In *Conference on Computer Vision and Pattern Recognition*, volume 1, pages 171–178, Jun 2003.

[116] L. Jofre, M. S. Hawley, A. Broquetas, E. de los Reyes, and M. Ferrandoand A. R. Elias-Fuste. Medical imaging with a microwave tomographic scanner. *Transactions on Biomedical Engineering*, 37(3):303–312, 1990.

[117] Bernard Journet and Stephane Poujouly. High-resolution laser range finder based on a phase-shift measurement method. *SPIE: Three-Dimensional Imaging, Optical Metrology, and Inspection IV*, 3520:123–132, 1998.

[118] North E. Coleman Jr. and Ramesh Jain. Obtaining 3-dimensional shape of textured and specular surface using four-source photometry. *Computer Vision, Graphics, and Image Processing*, 18(4):309–328, 1982.

[119] J. T. Kajiya and T. L. Kay. Rendering fur with three dimensional textures. In *SIGGRAPH*, pages 271–280, 1989.

[120] Avinash C. Kak and Malcolm Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, 1988.

[121] L. L. Kalpaxis, L. M. Wang, P. Galland, X. Liang, P. P. Ho, and R. R. Alfano. Three-dimensional temporal image reconstruction of an object hidden in highly scattering media by time-gated optical tomography. *Optics Letters*, 18:1691–1693, 1993.

[122] T. Kanade. Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence*, 17(1-3):409–460, 1981.

[123] M. Kaschke, H. Jess, G. Gaida, J.-M. Kaltenbach, and W. Wrobel. Transillumination imaging of tissue by phase modulation techniques. In *Advances in Optical Imaging and Photon Migration*, volume 21, pages 88–92. OSA, 1994.

[124] Jan Kautz, Mirko Sattler, Ralf Sarlette, Reinhard Klein, and Hans-Peter Seidel. Decoupling BRDFs from surface mesostructures. In *Conference on Graphics Interface*, pages 177–182. Canadian Human-Computer Communications Society, 2004.

[125] Greg Kay and Terry Caelli. Estimating the parameters of an illumination model using photometric stereo. *Graphical Models and Image Processing*, 57(5):365–388, 1995.

[126] D. R. Kirkby and D. T. Delpy. Measurement of tissue TPSF by use of a cross-correlation technique with an avalanche photodiode detector. In *Optical Tomography, Photon Migration, and Spectroscopy of Tissue*, volume 2389, pages 190–197. SPIE, 1995.

[127] Kodak Professional Division. *KODAK Q-60 Color Input Targets*, June 2003.

[128] Vladimir Kolmogorov and Ramin Zabih. Computing visual correspondence with occlusions via graph cuts. In *International Conference on Computer Vision*, volume 2, pages 508–515, 2001.

[129] Melissa L. Koudelka, Sebastian Magda, Peter N. Belhumeur, and David J. Kriegman. Image-based modeling and rendering of surfaces with arbitrary BRDFs. In *Conference on Computer Vision and Pattern Recognition*, volume 1, pages 568–575, 2001.

[130] Melissa L. Koudelka, Sebastian Magda, Peter N. Belhumeur, and David J. Kriegman. Acquisition, compression and synthesis of bidirectional texture functions. In *International Workshop on Texture Analysis and Synthesis*, pages 59–64, 2003.

[131] D. E. Kuhl and R. Q. Edwards. Image separation radioisotope scanning. *Radiology*, 80:653–661, 1963.

[132] K. N. Kutulakos. Approximate N-view stereo. In *European Conference on Computer Vision*, pages I: 67–83, 2000.

[133] K. N. Kutulakos and S. M. Seitz. A theory of shape by space carving. *International Journal of Computer Vision*, 38(3):199–218, 2000.

[134] Philippe Lacroute and Marc Levoy. Fast volume rendering using a shear-warp factorization of the viewing transformation. In *SIGGRAPH*, pages 451–458, 1994.

[135] Eric P. F. Lafortune, Sing-Choong Foo, Kenneth E. Torrance, and Donald P. Greenberg. Non-linear approximation of reflectance functions. In *SIGGRAPH*, pages 117–126, 1997.

[136] A. Laurentini. The visual hull concept for silhouette-based image understanding. *Transactions on Pattern Analysis and Machine Intelligence*, 16(2):150–162, 1994.

[137] Paul C. Lauterbur. Image formation by induced local interactions; examples employing nuclear magnetic resonance. *Nature*, 242:190–191, 1973.

[138] Hsien-Che Lee. *Introduction to Color Imaging Science*. Cambridge University Press, 2005.

[139] E. Leith, H. Chen, Y. Chen, D. Dilworth, J. Lopez, R. Masriand, J. Rudd, and J. Valdmanis. Electronic holography and speckle methods for imaging through tissue using femtosecond gated pulses. *Applied Optics*, 30:4204–4210, 1991.

[140] Jerome Lengyel, Emil Praun, Adam Finkelstein, and Hugues Hoppe. Real-time fur over arbitrary surfaces. In *Symposium on Interactive 3D Graphics*, pages 227–232. ACM Press, 2001.

[141] Marc Levoy and Pat Hanrahan. Light field rendering. In *SIGGRAPH*, pages 31–42, New York, NY, USA, 1996. ACM Press.

[142] R. A. Lewis and A. R. Johnston. A scanning laser rangefinder for a robotic vehicle. In *International Joint Conference on Artificial Intelligence*, pages 762–768, 1977.

[143] X. Liang, L. Wang, P. P. Ho, and R. R. Alfano. Two-dimensional Kerr-Fourier imaging of translucent phantoms in thick turbid media. *Applied Optics*, 34:3463–3467, 1995.

[144] H. Y. Lin and C. Y. Ho. 3D model acquisition based on projections of level curves. In *International Conference on Pattern Recognition*, pages IV: 853–856, 2006.

[145] Stephen Lin and Heung-Yeung Shum. Separation of diffuse and specular reflection in color images. In *Conference on Computer Vision and Pattern Recognition*, volume 1, pages 341–346, 2001.

[146] Zhouchen Lin, Tien-Tsin Wong, and Heung-Yeung Shum. Relighting with the reflected irradiance field: Representation, sampling and reconstruction. *International Journal of Computer Vision*, 49(2-3):229–246, 2002.

[147] Xinguo Liu, Yaohua Hu, Jingdan Zhang, Xin Tong, and Baining Guoand Heung-Yeung Shum. Synthesis and rendering of bidirectional texture functions on arbitrary surfaces. *Transactions on Visualization and Computer Graphics*, 10(3):278–289, 2004.

[148] Xinguo Liu, Yizhou Yu, and Heung-Yeung Shum. Synthesizing bidirectional texture functions for real-world surfaces. In *SIGGRAPH*, pages 97–106, 2001.

[149] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.

[150] Wan-Chun Ma, Sung-Hsiang Chao, Yu-Ting Tseng, Yung-Yu Chuang, Chun-Fa Chang, Bing-Yu Chen, and Ming Ouhyoung. Level-of-detail representation of bidirectional texture functions for real-time rendering. In *Symposium on Interactive 3D Graphics*, pages 187–194, New York, NY, 2005. ACM Press.

[151] Sebastian Magda and David Kriegman. Fast texture synthesis on arbitrary meshes. In *Eurographics Symposium on Rendering*, pages 82–89, 2003.

[152] Sebastian Magda, David J. Kriegman, Todd Zickler, and Peter N. Belhumeur. Beyond Lambert: Reconstructing surfaces with arbitrary BRDFs. In *International Conference on Computer Vision*, volume 2, pages 391–398. IEEE, 2001.

[153] R. Mahon, M. D. Duncan, L. L. Tankersley, and J. Reintjes. Time-gated imaging through dense scatterers with a Raman amplifier. *Applied Optics*, 32:7425–7433, 1993.

[154] Satya P. Mallick, Todd E. Zickler, David J. Kriegman, and Peter N. Belhumeur. Beyond Lambert: reconstructing specular surfaces using color. In *Conference on Computer Vision and Pattern Recognition*, volume 2, pages 619–626, 2005.

[155] A. Mansouri, A. Mitiche, and J. Konrad. Selective image diffusion: Application to disparity estimation. In *International Conference on Image Processing*, volume 3, pages 284–288, 1998.

[156] Vincent Masselus, Pieter Peers, Philip Dutré, and Yves D. Willems. Relighting with 4-D incident light fields. *ACM Trans. Graph.*, 22(3):613–620, 2003.

[157] J. Meseth, G. Müller, and R. Klein. Preserving realism in real-time rendering of bidirectional texture functions. In *OpenSG Symposium*, pages 89–96. Eurographics Association, 2003.

[158] A. Meyer and F. Neyret. Interactive volumetric textures. In *Eurographics Workshop on Rendering*, pages 157–168. Eurographics Association, 1998.

[159] G. Mitic, J. Kölzer, J. Otto, E. Plies, G. Sölkner, and W. Zinth. Time-gated transillumination of biological tissues and tissue-like phantoms. *Applied Optics*, 33:6699–6710, 1994.

[160] T. Mitsunaga and S. K. Nayar. Radiometric self calibration. In *Conference on Computer Vision and Pattern Recognition*, pages 374–380, June 1999.

[161] H. Morita, K. Yajima, and S. Sakata. Reconstruction of surfaces of 3-D objects by M-array pattern projection method. In *International Conference on Computer Vision*, pages 468–473, 1988.

[162] E. Mouaddib, J. Batlle, and J. Salvi. Recent progress in structured light in order to solve the correspondence problem in stereovision. In *International Conference on Robotics and Automation*, volume 1, pages 130–136, 1997.

[163] G. Müller, J. Meseth, and R. Klein. Compression and real-time rendering of measured BTFs using local PCA. In *International Fall Workshop on Vision, Modeling and Visualization*, pages 271–280, Berlin, Germany, 2003. Akademische Verlagsgesellschaft Aka GmbH.

[164] Jeffrey B. Mulligan and Xavier L. C. Brolly. Surface determination by photometric ranging. In *Conference on Computer Vision and Pattern Recognition Workshop*, volume 3, page 40, Washington, DC, USA, 2004. IEEE Computer Society.

[165] Risto Myllylä, Janusz Marszalec, Juha Kostamovaara, Antti Mäntyniemi, and Gerd-Joachim Ulbrich. Imaging distance measurements using TOF lidar. *Journal of Optics*, 29:188–193, 1998.

[166] K. Nakamura, T. Hara, M. Yoshida, T. Miyahara, and H. Ito. Optical frequency domain ranging by a frequency-shifted feedback laser. *Journal of Quantum Electronics*, 36(3):305–316, 2000.

[167] S. Nayar, Katsushi Ikeuchi, and Takeo Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *Transactions on Robotics and Automation*, 6(4):418–431, August 1990.

[168] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Trans. Graph.*, 25(3):935–944, 2006.

[169] R. Nevatia. Depth measurement by motion stereo. *Computer Graphics Image Processing*, 5(2):203–214, 1976.

[170] Fabrice Neyret. Modeling, animating, and rendering complex scenes using volumetric textures. *Transactions on Visualization and Computer Graphics*, 4(2):55–70, 1998.

[171] F. Nicodemus, J. Richmond, J. Hsia, I. Ginsberg, and T. Limperis. Geometrical considerations and nomenclature for reflectance. NBS Monograph No. 160, National Bureau of Standards, Washington, DC, 1977.

[172] S. J. Norton and M. Linzer. Ultrasonic reflectivity tomography in three dimensions: Reconstruction with spherical transducer arrays. *Ultrasonographic Imaging*, 1(2):210–231, 1979.

[173] Y. Ohta and T. Kanade. Stereo by intra- and inter-scanline search using dynamic programming. *Transactions on Pattern Analysis and Machine Intelligence*, 7:139–154, 1985.

[174] Y. Ohta, M. Watanabe, and K. Ikeda. Improving depth map by right-angled trinocular stereo. In *International Conference on Pattern Recognition*, pages 519–521, 1986.

[175] M. Okutomi and T. Kanade. A multiple-baseline stereo. *Transactions on Pattern Analysis and Machine Intelligence*, 15(4):353–363, 1993.

[176] M. A. OLeary, D. A. Boas, B. Chance, and A. G. Yodh. Experimental images of heterogeneous turbid media by frequency domain diffusing photon tomography. *Optics Letters*, 20:426–428, 1995.

[177] Gary R. Olhoeft. Application of ground penetrating radar. In *International Conference on Ground Penetrating Radar*, pages 1–4, September 1996.

[178] Opto Technology, Inc. *High Flux LED Illuminators OTLXXXA-5-10-66-E, Rev. X1*, January 2000.

[179] D. G. Papaioannou, S. B. Colak, and G. W. 't Hooft. Resolution and sensitivity of optical imaging in highly scattering media. In *Photon Propagation in Tissues*, volume 2626, pages 218–227. SPIE, December 1995.

[180] Panos M. Pardalos and H. Edwin Romeijn, editors. *Handbook of Global Optimization.* Kluwer Academic Publishers, 2002.

[181] Bui Tuong Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, 1975.

[182] C. Pintavirooj, A. Romputtal, A. Ngamlamiad, W. Withayachumnankul, and K. Hamamoto. Ultrasonic refractive index tomography. *Journal of WSCG*, 12(1-3):333–339, 2004.

[183] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance absorption by nuclear magnetic moments in a solid. In *Physical Review*, volume 69, page 37, 1945.

[184] Chandrasekhara Venkata Raman and Kariamanikkam Srinivasa Krishnan. A new type of secondary radiation. *Nature*, 121:501–502, 1928.

[185] Y. Reibel, M. Jung, M. Bouhifd, B. Cunin, and C. Draman. CCD or CMOS camera noise characterization. *European Physical Journal*, 21:75–80, 2003.

[186] J. Reintjes, M. Bashkansky, M. D. Duncan, R. Mahon, L. L. Tankersleyand J. A. Moon, C. L. Adler, and J. M. S. Prewitt. Time-gated imaging with nonlinear optical Raman interactions. *Optics & Photonics News*, 4:28–32, 1993.

[187] P. I. Richards. Scattering from a point-source in plane clouds. *Journal of the Optical Society of America*, 46:927–934, 1956.

[188] M. Rioux, F. Blais, J. Beraldin, and P. Boulanger. Range imaging sensors development at NRC laboratories. In *Workshop on Interpretation of 3D Scenes*, pages 154–160, 1989.

[189] A. P. Rockwood and J. Winget. 3-dimensional object reconstruction from 2-dimensional images. *Computer-Aided Design*, 29(4):279–285, 1997.

[190] Sébastien Roy and Ingemar J. Cox. A maximum-flow formulation of the N-camera stereo correspondence problem. In *International Conference on Computer Vision*, page 492. IEEE Computer Society, 1998.

[191] Alex Ryer. *Light Measurement Handbook*. Technical Publications Department, International Light, Inc., Newburyport, MA, 1998.

[192] H. Saito and T. Kanade. Shape reconstruction in projective voxel grid space from large number of images. In *Conference on Computer Vision and Pattern Recognition*, pages II: 49–54, 1999.

[193] Y. Sato and K. Ikeuchi. Temporal-color space analysis of reflection. *Journal of the Optical Society of America A*, 11(11):2990–3002, November 1994.

[194] A. Savitzky and M. J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36:1627–1639, 1964.

[195] Takeo Sawatari. Real-time noncontacting distance measurement using optical triangulation. *Applied Optics*, 15(11):2821–2824, 1976.

[196] Daniel Scharstein and Richard Szeliski. Stereo matching with non-linear diffusion. In *Conference on Computer Vision and Pattern Recognition*, page 343, 1996.

[197] Daniel Scharstein and Richard Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1-3):7–42, 2002.

[198] Karsten Schlüns and Oliver Witting. Photometric stereo for non-Lambertian surfaces using color information. In *Computer Analysis of Images and Patterns*, pages 444–454, 1993.

[199] Cordelia Schmid and Andrew Zisserman. The geometry and matching of curves in multiple views. In *European Conference on Computer Vision*, volume 1, pages 394–409, 1998.

[200] Florian E. W. Schmidt, Martin E. Fry, Elizabeth M. C. Hillman, Jeremy C. Hebden, and David T. Delpy. A 32-channel time-resolved instrument for medical optical tomography. *Review of Scientific Instruments*, 71:256–265, January 2000.

[201] J. M. Schmitt, A. H. Gandjbakhche, and R. F. Bonner. Use of polarized light to discriminate short-path photons in a multiply scattering medium. *Applied Optics*, 31:6535–6546, 1992.

[202] Joseph M. Schmitt. Optical Coherence Tomography (OCT): A review. *Selected Topics in Quantum Electronics*, 5(4):1205–1215, July/August 1999.

[203] O. Schütz, H. Siebold, L. Götz, and S. Heywang-Köbrunner. Preliminary imaging results of a 4-wavelength NIR breast scanner. Oral Presentation at the 2$^{\text{nd}}$ Workshop on NIRS & Imaging of Biological Tissue, Florence, Italy, April 1995.

[204] M. Schweiger, S. R. Arridge, M. Hiraoka, M. Firbank, and D. T. Delpy. Comparison of a finite element forward model with experimental phantom results: application to image reconstruction. In *SPIE Proceedings*, volume 1888, pages 179–190. The International Society for Optical Engineering, 1993.

[205] Steven M. Seitz, Brian Curless, James Diebel, Daniel Scharstein, and Richard Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Conference on Computer Vision and Pattern Recognition*, volume 1, pages 519–526, 2006.

[206] Steven M. Seitz and Charles R. Dyer. Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, 35(2):151–173, 1999.

[207] Steven A. Shafer. Using color to separate reflection components. *Color Research and Applications*, 10(4):210–218, 1985.

[208] J. Shah. A nonlinear diffusion model for discontinuous disparity and half-occlusions in stereo. In *Conference on Computer Vision and Pattern Recognition*, pages 34–40, 1993.

[209] Li Shen, Takashi Machida, and Haruo Takemura. Efficient photometric stereo technique for three-dimensional surfaces with unknown BRDF. In *International Conference on 3-D Digital Imaging and Modeling*, pages 326–333, 2005.

[210] Y. Shirai and M. Suwa. Recognition of polyhedrons with a range finder. In *International Joint Conference on Artificial Intelligence*, pages 80–87, 1971.

[211] Peter S. Shirley. *Physically Based Lighting Calculations for Computer Graphics*. PhD thesis, University of Illinois at Urbana Champaign, December 1990.

[212] Heung-Yeung Shum, Richard Szeliski, Simon Baker, Mei Han, and P. Anandan. Interactive construction of 3D models from panoramic mosaics. In *Conference on Computer Vision and Pattern Recognition*, pages 427–433, 1998.

[213] Robert Siegel and John R. Howell. *Thermal Radiation Heat Transfer*. Taylor & Francis, New York, NY, 4th edition, 2002.

[214] William M. Silver. Determining shape and reflectance using multiple images. Master's thesis, Massachusetts Institute of Technology, 1980.

[215] G. Simmons, D. W. Strangway, L. Bannister, R. Baker, D. Cubleyand G. la Torraca, and R. Watts. The surface electrical properties experiment. In *Lunar Geophysics Conference*, pages 258–271. Lunar Science Institute, 1972.

[216] Gregory G. Slabaugh, W. Bruce Culbertson, Thomas Malzbender, Mark R. Stevens, and Ronald W. Schafer. Methods for volumetric reconstruction of visual scenes. *International Journal of Computer Vision*, 57(3):179–199, 2004.

[217] Fredric Solomon and Katsushi Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *Transactions on Pattern Analysis and Machine Intelligence*, 18(4):449–454, 1996.

[218] K. G. Spears, J. Serafin, N. H. Abramson, X. Zhu, and H. Bjelkhagen. Chrono-coherent imaging for medicine. *Transactions on Biomedical Engineering*, 36:1210–1221, 1989.

[219] E. M. Strzelecki, D. A. Cohen, and L. A. Coldren. Investigation of tunable single frequency diode lasers for sensor applications. *Journal of Lightwave Technology*, 6(10):1610–1618, 1988.

[220] Jian Sun, Nan-Ning Zheng, and Heung-Yeung Shum. Stereo matching using belief propagation. *Transactions on Pattern Analysis and Machine Intelligence*, 25(7):787–800, 2003.

[221] Boaz J . Super and Alan C. Bovik. Shape from texture using local spectral moments. *Transactions on Pattern Analysis and Machine Intelligence*, 17(4):333–343, 1995.

[222] Frank Suykens, Karl vom Berge, Ares Lagae, and Philip Dutré. Interactive rendering with bidirectional texture functions. *Computer Graphics Forum*, 22(3):463–472, 2003.

[223] D. J. Svetkoff, P. F. Leonard, R. E. Sampson, and R. Jain. Techniques for real-time 3D feature extraction using range information. In *Intelligent Robotics and Computer Vision*, volume 521, pages 302–309, 1984.

[224] Richard Szeliski and Polina Golland. Stereo matching with transparency and matting. *International Journal of Computer Vision*, 32(1):45–61, 1999.

[225] Hemant D. Tagare and Rui J. P. deFigueiredo. A theory of photometric stereo for a class of diffuse non-Lambertian surfaces. *Transactions on Pattern Analysis and Machine Intelligence*, 13(2):133–152, 1991.

[226] Carlo Tomasi and Roberto Manduchi. Stereo matching as a nearest-neighbor problem. *Transactions on Pattern Analysis and Machine Intelligence*, 20(3):333–340, 1998.

[227] Xin Tong, Jingdan Zhang, Ligang Liu, Xi Wang, Baining Guo, and Heung-Yeung Shum. Synthesis of bidirectional texture functions on arbitrary surfaces. In *SIGGRAPH*, pages 665–672, 2002.

[228] G. Tziritas. Estimation of motion and structure of 3-D objects from a sequence of images. In *International Conference on Computer Vision*, pages 693–697, 1987.

[229] Frank A. van den Heuvel. 3D reconstruction from a single image using geometric constraints. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(3):354–368, 1998.

[230] M. B. van der Mark, G. W. 't Hooft, A. J. H. Wachters, U. H. deVries, J. P. Janssen, and M. N. J. M. Wasser. Clinical study of the female breast using spectroscopic diffuse optical tomography. In *Optical Society of America Biomedical Topical Meeting*. Optical Society of America, April 2000.

[231] M. Alex O. Vasilescu and Demetri Terzopoulos. TensorTextures: multilinear image-based rendering. In *SIGGRAPH*, pages 336–342, 2004.

[232] V. Venkateswar and R. Chellappa. Hierarchical stereo and motion correspondence using feature groupings. *International Journal of Computer Vision*, 15(3):245–269, 1995.

[233] Hermann L. F. von Helmholtz. *Treatise on Physiological Optics*, volume 1. Dover Publications, Inc., Dover, NY, 1962 edition, 1925.

[234] Jing Wang and Kristin J. Dana. Hybrid textons: modeling surfaces with reflectance and geometry. In *Conference on Computer Vision and Pattern Recognition*, volume 1, pages 372–378, 2004.

[235] L. Wang, P. P. Ho, and R. R. Alfano. Time-resolved Fourier spectrum and imaging in highly scattering media. *Applied Optics*, 32(26), Sep 1993.

[236] L. Wang, P. P. Ho, C. Liu, G. Zhang, and R. R. Alfano. Ballistic 2-D imaging through scattering walls using an ultrafast optical Kerr gate. *Science*, 253:769–771, 1991.

[237] Xi Wang, Xin Tong, Stephen Lin, Shimin Hu, Baining Guo, and Heung-Yeung Shum. Generalized displacement maps. In *Eurographics Symposium on Rendering*, pages 227–234. Eurographics Association, 2004.

[238] Y. F. Wang, A. Mitiche, and J. K. Aggarwal. Computation of surface orientation and structure of objects using grid coding. *Transactions on Pattern Analysis and Machine Intelligence*, 9(1):129–137, 1987.

[239] J. Watson, P. Georges, T. Lépine, B. Alonzi, and A. Brun. Imaging in diffuse media with ultrafast degenerate optical parametric amplification. *Optics Letters*, 20:231–233, 1995.

[240] J. J. Wild and J. M. Reid. Application of echo-ranging techniques to the determination of structure of biological tissues. *Science*, 115:226–230, 1952.

[241] A. O. Wist, P. P. Fatouros, and S. L. Herr. Increased spatial resolution in transillumination using collimated light. *Transactions on Medical Imaging*, 12:751–757, 1993.

[242] Andrew P. Witkin. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17(1-3):17–45, 1981.

[243] Tien-Tsin Wong, Pheng-Ann Heng, Siu-Hang Or, and Wai-Yin Ng. Image-based rendering with controllable illumination. In *Eurographics Workshop on Rendering*, pages 13–22, London, UK, 1997. Springer-Verlag.

[244] Robert J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.

[245] Tai-Pang Wu and Chi-Keung Tang. Dense photometric stereo using a mirror sphere and graph cut. In *International Conference on Computer Vision*, volume 1, pages 140–147, 2005.

[246] M. Yachida, Y. Kitamura, and M. Kimachi. Trinocular vision: New approach for correspondence problem. In *International Conference on Pattern Recognition*, pages 1041–1044, 1986.

[247] C. Yan and J.-C. Diels. Imaging with femtosecond pulses. *Applied Optics*, 32:6869–6873, 1992.

[248] R. Yang, M. Pollefeys, and G. Welch. Dealing with textureless regions and specular highlights: A progressive space carving scheme using a novel photo-consistency measure. In *International Conference on Computer Vision*, pages 576–584, 2003.

[249] Y. Yao, Y. Wang, Y. Pei, W. Zhu, and R. L. Barbour. Frequency-domain optical imaging of absorption and scattering distributions by a Born iterative method. *Journal of the Optical Society of America A*, 14:325–342, 1997.

[250] K. Yoshida and S. Hirose. Real-time stereo vision with multiple arrayed camera. In *International Conference on Robotics and Automation*, pages 1765–1770, 1992.

[251] Tianli Yu, Ning Xu, and Narendra Ahuja. Shape and view independent reflectance map from multiple views. In *European Conference on Computer Vision*, 2004.

[252] Yizhou Yu and Johnny T. Chang. Shadow graphs and surface reconstruction. In *European Conference on Computer Vision*, pages 31–45, London, UK, 2002. Springer-Verlag.

[253] Ramin Zabih and John Woodfill. Non-parametric local transforms for computing visual correspondence. In *European Conference on Computer Vision*, volume 2, pages 151–158, 1994.

[254] G. Zeng, S. Paris, L. Quan, and F.X. Sillion. Progressive surface reconstruction from images using a local prior. In *International Conference on Computer Vision*, pages II: 1230–1237, 2005.

[255] L. Zhang and S. Seitz. Image-based multiresolution modeling by surface deformation. In *SPIE: Videometrics and Optical Methods for 3D Shape Measurement*, 2001.

[256] Li Zhang, Guillaume Dugas-Phocion, Jean-Sebastien Samson, and Steven M. Seitz. Single view modeling of free-form scenes. In *Conference on Computer Vision and Pattern Recognition*, volume 1, page 990, 2001.

[257] Todd E. Zickler, Peter N. Belhumeur, and David J. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *International Journal of Computer Vision*, 49(2-3):215 – 227, Sep 2002.

[258] Todd E. Zickler, Jeffrey Ho, David J. Kriegman, Jean Ponce, and Peter N. Belhumeur. Binocular Helmholtz stereopsis. In *International Conference on Computer Vision*, volume 2, page 1411, Washington, DC, USA, 2003. IEEE Computer Society.

Author's Biography

Sebastian Magda was born in 1976 in Poland. In 1992 he moved with his parents to the United States of America. He received the Bachelor of Science degree in Computer Science and the Bachelor of Science degree in Electrical Engineering from the University of Illinois at Urbana-Champaign in 1999. As an outstanding Computer Science senior he was honored with the ICCP-James N. Snyder Award and the C.W. Gear Award. In his undergraduate work he developed an interest in artificial intelligence and joined the Beckman Institute Artificial Intelligence Group where he worked as a research programmer. He continued his graduate education in Computer Science at the University of Illinois at Urbana-Champaign under the guidance of Prof. David Kriegman. He obtained the Master of Science degree in 2003. In 2005 he continued his research as a visiting scholar at the University of California at San Diego. Currently he is completing his Doctorate of Philosophy. His research interests focus on computer vision, especially on geometric reconstruction, reflectance modeling, photometric calibration, as well as on efficient modeling and real-time rendering of surface appearance.