Acquisition of Subsurface Scattering Objects

Diplomarbeit — Diploma Thesis

Christian Fuchs

Max-Planck-Institut für Informatik
Saarbrücken, Germany

Betreuer — Supervisors
Dr. Michael Goesele,
Max-Planck-Institut für Informatik, Saarbrücken, Germany

Prof. Dr. Hans-Peter Seidel,
Max-Planck-Institut für Informatik, Saarbrücken, Germany

Gutachter — Reviewers
Prof. Dr. Hans-Peter Seidel,
Max-Planck-Institut für Informatik, Saarbrücken, Germany

Prof. Dr. Philipp Slusallek,
Universität des Saarlandes, Saarbrücken, Germany

Christian Fuchs
Max-Planck-Institut für Informatik
Stuhlsatzenhausweg 85
66123 Saarbrücken, Germany
cfuchs@mpi-sb.mpg.de

09. Februar 2006

Eidesstattliche Versicherung

Hiermit erkläre ich an Eides statt, dass ich diese Arbeit selbstständig verfasst und keine anderen als die im Literaturverzeichnis angegebenen Quellen benutzt habe.

Christian Fuchs
Saarbrücken, 30. September 2004

Contents

Chapter 1

Introduction

Light interacts with ideal opaque objects purely locally, i.e., it is scattered at the point of incidence. This behavior can be described using a bidirectional reflectance distribution function (BRDF) [Nicodemus et al. 1977]. In contrast, light passes straight through transparent objects (e.g., glass or water) and is only reflected or refracted at material boundaries. Many daily life objects (e.g., milk, skin, marble or alabaster) are translucent and belong to neither of these categories. The objects are characterized by multiple light scattering inside the object.

This *subsurface scattering* behavior leads to a very distinct appearance which is visually important: Light shines through objects as seen in Figure 1.1 where an alabaster block is illuminated by a narrow laser beam. Furthermore, translucent objects appear smooth and surface details are hidden. This is evident in Figure 1.2 which shows an example of the same object with and without subsurface scattering under identical lighting conditions. Subsurface scattering was suppressed by covering the object with a fine layer of white dust. Much of the geometric detail in the mane and the face are washed out by the subsurface scattering.

Translucent objects can be rendered using a variety of techniques. Most recent rendering systems are based on the dipole approximation to a diffusion model introduced by Jensen et al. [2001]. Like all physically-based rendering techniques, this method requires knowledge of the properties of the material that is going to be rendered. Jensen et al. [2001] described therefore a method to acquire physically correct material parameters with a point-based measurement setup. Since this method is limited to homogeneous materials we develop a system tailored to capture the necessary input data to render a real, translucent object with spatially varying material properties (like the alabaster block in Figure 1.1 with its veins) using the rendering system proposed by Lensch et al. [2003b].

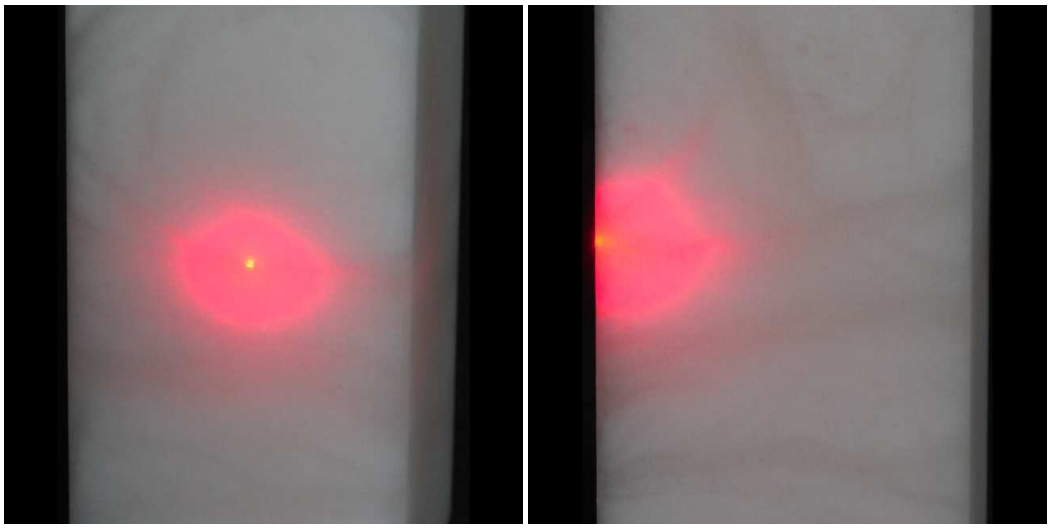


Figure 1.1: Subsurface scattering in an alabaster block. *Left:* Front view of the alabaster block while it is illuminated from the front. *Right:* Side view of the block while it is illuminated by a laser from the front. Note how the light penetrates the object and how it is scattered even deeper into the object along the veins in the alabaster.

Figure 1.2: Photographs of an alabaster horse model with and without subsurface light scattering under identical lighting conditions. *Left:* Original object, much surface detail is hidden by the translucency. *Right:* The object is covered with fine white dust to stop light from entering the object.

1.1 Basic Approach

We chose to build an image-based acquisition system that takes the specific properties of subsurface scattering objects (like diffuse scattering) into account. The diffuseness removes the angular dependencies of the incoming and outgoing light thus simplifying the measurement process. During acquisition the object is illuminated with a laser beam at a dense set of surface points. The object's impulse response is then captured with a high-dynamic range video camera. Once the impulse response is known for all pairs of incoming and outgoing surface locations, the object is characterized completely using the diffuse assumption.

The acquired data is resampled into a hierarchical data structure taking advantage of the specific properties of translucent material such as the exponential fall-off near the point of incidence and the smooth global response. Holes in the local and global data (mostly caused by occlusion) are consistently filled. This also includes a smoothing step that is used to suppress noise in the data caused by the acquisition process. We render the objects from arbitrary viewpoints and with arbitrary lighting conditions using the method proposed by Lensch et al. [2003b]. The incident illumination can vary per surface point (e.g., illumination with textured point lights, or projected shadow borders). This diploma thesis is part of a framework covering a complete pipeline for acquisition, processing, and rendering of translucent objects which has already been published in several conference proceedings and journals [Lensch et al. 2003b, Goesele et al. 2004].

1.2 Overview

The main focus of this diploma thesis is on the acquisition process employed in this pipeline. The remainder of this thesis is structured as follows: Chapter 2 describes the theory of subsurface scattering and our modeling assumptions. Chapter 3 summarizes previous work in this area. In Chapter 4, the acquisition process is described in detail. The resampling and postprocessing steps followed by a discussion of applicable rendering methods are presented in Chapter 5. Results of our measurement based modeling are presented in Chapter 6 before the conclusion and the discussion of future work in Chapter 7.

Chapter 2
Background

In this chapter we will review some of the basic results of physics that are necessary for a full understanding of the methods described in the remaining chapters. In addition to that we will describe the principles and terms behind the BSSRDF (Bidirectional Scattering Surface Reflectance Distribution Function). We will show briefly how the BSSRDF can be approximated by a dipole light source. This chapter will also introduce the *diffuse reflectance* function R_d and describe how this function can be stored and rendered hierarchically to make its handling more practical.

2.1 Radiometry

As a form of electromagnetic radiation, light can be interpreted either as a wave containing electrical and magnetic components at different frequencies (wave optics) or as a flow of particles (the *photons*) transporting the energy (particle optics).

In wave optics, the energy is transported by oscillating magnetic and electric fields. The oscillation directions for the magnetic and electrical fields are perpendicular to each other and to the propagation direction of the light. Light that only consists of waves that oscillate in the same direction is called *polarized*.

In particle optics, the energy is carried in the form of *photons*. Every photon carries a certain amount of energy, depending on the frequency of the light.

Radiometry is the science that deals with measurements of electromagnetic radiation, including light. The most important quantities in this field are also required for understanding the principles of digital image synthesis. These are:

Radiant Energy: The energy transported by light of all wavelengths, measured in *Joule* $\left[J = Ws = kg\ \frac{m^2}{s^2}\right]$.

Radiant Flux, Radiant Power Φ: The power (energy per unit time) of the radiation. It is measured in *Watts* $[W]$.

Radiant Intensity I: Radiant Flux per solid angle ω measured in *Watts per unit solid angle* $\left[\frac{W}{sr}\right]$.

$$I := \frac{d\Phi}{d\omega}$$

The intensity I can be used to characterize point light sources. Since a full sphere of directions has a solid angle of $4\pi \cdot sr$, an isotropic point light (a light source that emits the same amount of light in each direction) has an intensity of $\frac{\Phi}{4\pi \cdot sr}$.

Irradiance E and Radiant Exitance B: The irradiance $E := \frac{d\Phi}{dA}$ represents the radiant flux $d\Phi$ *arriving* at a surface with area dA while the radiant exitance B, which is often called *radiosity* in computer graphics, describes the radiant flux per unit area *leaving* a surface. Both are measured in $\left[\frac{W}{m^2}\right]$.

Radiance L: Radiant Flux per unit projected area per unit solid angle arriving at or leaving a point on a surface measured in $\left[\frac{W}{m^2 \cdot sr}\right]$.

$$L(x, \omega) := \frac{d^2\Phi}{cos\theta \cdot d\omega \cdot dA}$$

where x is the point on the surface and θ is the angle between the direction ω and the surface normal.

The relationship between irradiance and incoming radiance is

$$E(x) = \int_{\Omega(\vec{n})} L_i(x_i, \omega_i) cos\theta_i d\omega_i$$

where $\Omega(\vec{n})$ represents the hemisphere of directions around the surface normal $\vec{n}$ and L_i is the radiance arriving at the surface point x.

Radiance is an important quantity in computer graphics since it is constant along a ray in empty space. Therefore it is used by almost all rendering systems including ray-tracers.

2.2 Light Transport in Scattering Media

When light hits an opaque object (for example metals or chalk), a certain fraction of its energy is absorbed by the material. The remaining light is reflected according to the BRDF [Nicodemus et al. 1977].

Figure 2.1: Example for single scattering inside an object. *Left:* A glass of honey illuminated with a red laser beam. *Right:* Shorter exposure time in order to improve visibility of the laser beam.

Translucent objects however allow the light to enter the material, where it is scattered around in a manner that depends on the physical properties of the material like the absorption coefficient (σ_a) and the scattering coefficient (σ_s). For a detailed explanation of these and other properties, see [Jensen et al. 2001] or [Ishimaru 1978].

2.2.1 Single vs. Multiple Scattering

Light scattering inside an object can be divided into two subclasses: In the case of *single scattering*, light reaches the observer after only one scattering event. Single scattering inside the object depends strongly on the direction of incident light and the viewing direction of the observer. Figure 2.1 shows an example of a single scattering object. To measure it properly would require to illuminate the object at all surface points from all possible directions while at the same time observing all surface points from all possible directions. If light undergoes *multiple scattering* events (see Figure 2.2 for an illustration) in a highly scattering material, it is distributed uniformly around the point of incidence. The *diffuse reflection* R_d of light that is scattered multiple times can be measured more easily since it does no longer depend on the direction of the incoming light and on the viewing direction of the observer.

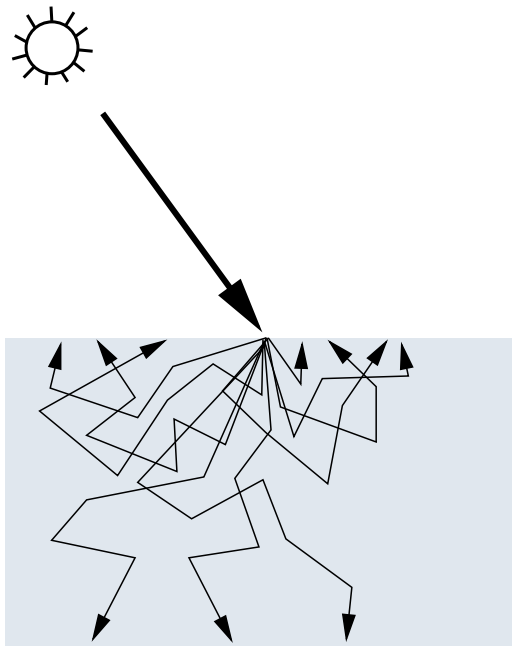


Figure 2.2: Subsurface scattering. Light penetrates the surface of translucent objects and is scattered around (multiple times) before it leaves the object (probably at some other location).

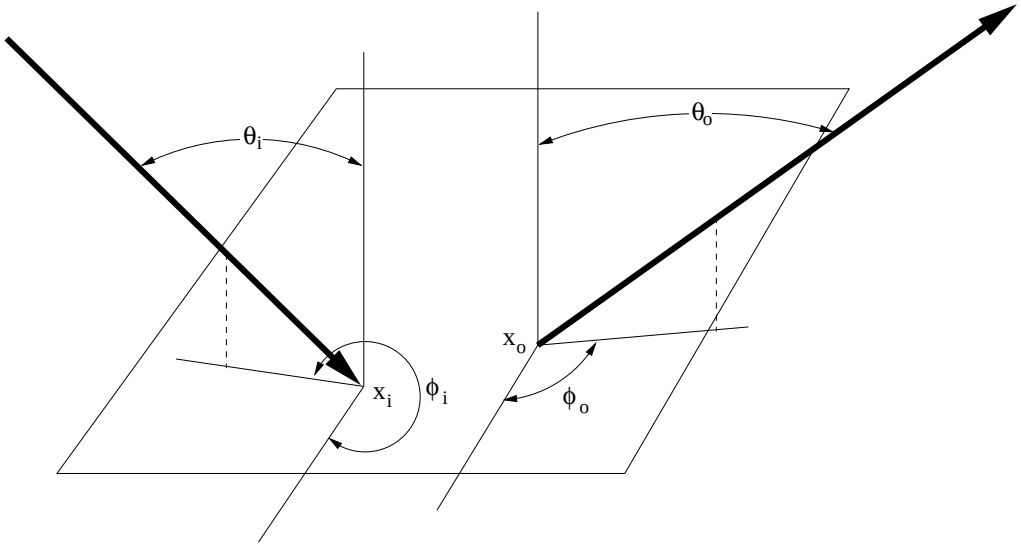


Figure 2.3: Geometry of the BSSRDF S for general cases where subsurface scattering is involved

2.2.2 Subsurface Light Transport

The BSSRDF (bidirectional scattering-surface reflectance distribution function) S provides a general model for basic light transport inside an object [1] and is defined as follows:

$$S(x_i, \omega_i; x_o, \omega_o) := \frac{dL^{\rightarrow}(x_o, \omega_o)}{d\Phi^{\leftarrow}(x_i, \omega_i)}$$

Therefore the BSSRDF S is the ratio between the outgoing radiance $L^{\rightarrow}(x_o, \omega_o)$ leaving the surface at x_o in direction ω_o and the incident flux $\Phi^{\leftarrow}(x_i, \omega_i)$ arriving at x_i from direction ω_i. The unit of the BSSRDF S is $\left[\frac{1}{m^2 \cdot sr}\right]$.

Without taking spectral effects, polarization or fluorescence into account, the outgoing radiance $L^{\rightarrow}(x_o, \omega_o)$ at a surface location x_o and a direction ω_o can be computed, given the incident radiance $L^{\leftarrow}(x_i, \omega_i)$ at all surface locations x_i and all directions ω_i by integrating over the whole surface A and the complete hemisphere of directions Ω_+:

$$L^{\rightarrow}(x_o, \omega_o) = \int_A \int_{\Omega_+(x_i)} L^{\leftarrow}(x_i, \omega_i) S(x_i, \omega_i; x_o, \omega_o) d\omega_i dx_i$$

. Figure 2.3 illustrates this for the case of one incident beam of light.

If we assume that the light is scattered multiple times inside the object, which removes the angular dependencies at incoming at exiting surface locations, we can replace S with a 4D diffuse subsurface reflectance function $R_d(x_i, x_o)$:

$$L^{\rightarrow}(x_o, \omega_o) = \frac{1}{\pi} F_{t,o}(\eta, \omega_o) \int_A R_d(x_i, x_o) \cdot \int_{\Omega_+(x_i)} L^{\leftarrow}(x_i, \omega_i) F_{t,i}(\eta, \omega_i) \langle N_i \cdot \omega_i \rangle d\omega_i dx_i. \tag{2.1}$$

The Fresnel transmittance factors $F_t(\eta, \omega)$ model what fraction of the flux or radiosity is transmitted through the surface boundary for a direction ω and a given relative index of refraction η. The factor $1/\pi$ converts radiosity into exitant radiance. The area foreshortening is represented by $\langle N_i \cdot \omega_i \rangle$.

The goal of our work is to measure $R_d(x_i, x_o)$ per color channel for all incoming and outgoing surface locations x_i and x_o. If we illuminate a target object at a single surface point x_i' with known incident radiance $L^{\leftarrow}(x_i', \omega_i)$, we can observe $L^{\rightarrow}(x_o, \omega_o)$ on the object's surface. Inverting Equation 2.1 becomes trivial and we can record the impulse response $R_d(x_i', x_o)$ for all x_o.

[1] The BSSRDF is closely related to reflectance fields [Debevec et al. 2000]

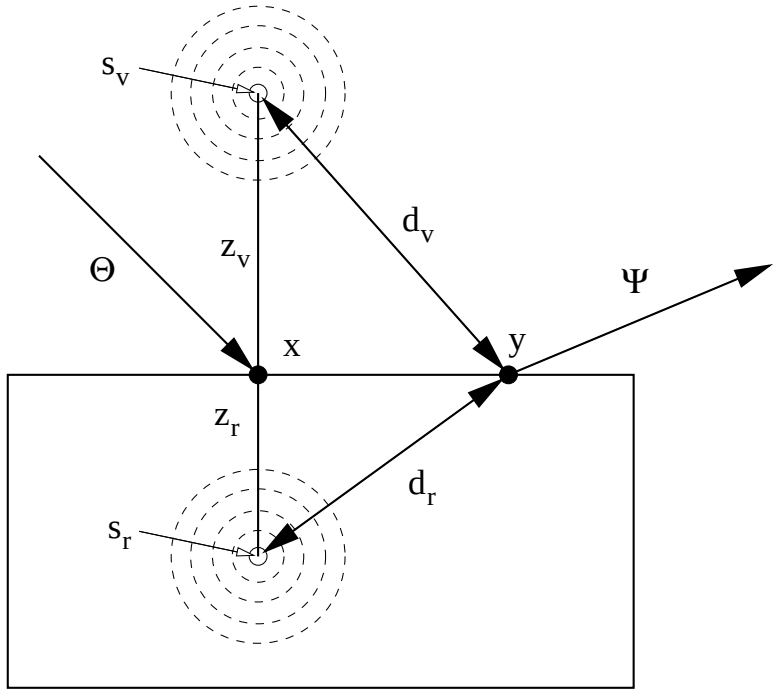


Figure 2.4: The BSSRDF model of Jensen et al. [2001] based on a dipole source approximation. A pair of imaginary point light sources is placed above and below the surface point x. The distances z_r and z_v are calculated from the reduced scattering coefficient σ_s' and the absorption coefficient σ_a of the material.

2.3 Dipole Approximation

Jensen et al. [2001] introduced a practical model for subsurface scattering based on a dipole source approximation of the diffusion equation. They compute the outgoing radiance as the sum of two exponential functions:

$$R_d(x,y) = \frac{\alpha'}{4\pi}\left[z_r(1+\sigma_{tr}d_r)\frac{e^{-\sigma_{tr}d_r}}{d_r^3} + z_v(1+\sigma_{tr}d_v)\frac{e^{-\sigma_{tr}d_v}}{d_v^3}\right]$$

$a' = \sigma_s'/\sigma_t'$ and $\sigma_{tr} = \sqrt{3\sigma_a\sigma_t'}$ are material properties, $\sigma_a, \sigma_s', \sigma_t'$ and σ_{tr} have been introduced in section 2.2. z_r and z_v are the distances of a pair of imaginary point light sources to the point of incidence x. d_r and d_v are the distances between the exiting light point y and these imaginary source points (see Figure 2.4 for an illustration).

2.4 Hierarchical Model for R_d

Storing and rendering the full 4D function R_d with high resolution is impractical. The nature of optically dense translucent objects makes R_d however well suitable for a hierarchical modeling approach: R_d will typically vary strongly in the proximity of x_i due to the exponential fall-off of radiosity inside the scattering volume. In distant areas, R_d varies quite smoothly. Sharp features are usually caused by inhomogeneous material properties within a small volume on and right below the object's surface at x_i and x_o. All photons pass through these volumes and their influence is significant (see Feng et al. [1993] for a study of photon path distributions within scattering material).

Lensch et al. [2003b] introduced therefore a three-part model in their rendering system: The irradiance for all surface areas is collected in a texture atlas in the form of a diffuse light map. Spatially varying filter kernels $K_{(u,v)}$ that are convolved with the light map model an object's impulse response in the immediate vicinity of the incoming surface location x_i. The global term is modeled by a vertex-to-vertex throughput factor matrix. Energy transport is performed by multiplication with the per-vertex irradiance. The global response is linearly interpolated between the vertices of the underlying triangle mesh (equivalent to Gouraud interpolation). An optional modulation texture T_ρ adds surface appearance detail by modulating the global response.

The general strategy is to use a model with high sampling density in the vicinity of the incoming impulse and a much coarser sampling in distant, smooth areas (still modulated by a detail texture T_ρ). The sample location could for example also be determined by adaptively subdividing the input mesh. Subdivision allows one to choose the size of the densely sampled area adaptively and to guarantee that global and corresponding local samples coincide whenever possible.

We implemented our renderer based on the technique of Lensch et al. [2003b] and will describe all further processing steps with respect to this rendering technique. The problems and the proposed solutions are however general and transferring them to an alternative hierarchical data structure should be straightforward.

Chapter 3
Previous Work

Ishimaru [1978] did comprehensive studies on the theory of how radiation propagates in randomly scattering media. This is of great interest in many applications (e.g. wave propagation in communications, medicine and remote sensing).

In medical and remote sensing applications the goal is to draw conclusions about the interior properties of some scattering media by observing its response to incoming radiation in the time and/or space domain. Obviously this requires solving a difficult inverse problem. The predominant goal in computer graphics is to acquire enough data to get a realistic visual model of an object. This also includes the object's response to incident illumination. Our approach is entirely based on external measurements and does not depend on the inversion.

3.1 Models for Translucent Objects

Translucent objects can be modeled by defining their basic physical properties such as the absorption coefficient σ_a and the scattering coefficient σ_s [Ishimaru 1978] for each point inside their volume. Rendering the object can then be done by physical simulation or appropriate approximations.

Alternatively, it is sufficient to record the visible effects of subsurface scattering. This information can then be used to render the object without knowledge of its internal structure. The bidirectional scattering-surface reflectance distribution function $S(x_i, \omega_i; x_o, \omega_o)$ (BSSRDF) [Nicodemus et al. 1977] relates irradiance at any surface position x_i to the reflected radiance at another surface position x_o for incoming and outgoing directions ω_i and ω_o. The complexity of this 8D function makes handling it and especially sampling it really hard. However, for optically dense material the directional dependency is negligible since the response is mainly dominated by multiple scattering. The BSSRDF model of Jensen et al. [2001] is therefore split into a directionally dependent single scattering term

$S^{(1)}$ and a directionally independent multiple scattering term S_d. Omitting the single scattering term, the BSSRDF of a translucent object can be collapsed to the 4D diffuse subsurface reflectance function $R_d(x_i, x_o)$ that depends only on the incoming and outgoing surface positions.

3.2 Data Acquisition

BRDF [Nicodemus et al. 1977] acquisition techniques such as proposed by [Marschner et al. 1999, Lensch et al. 2003a] are only able to capture the local reflectance properties of an object. They can not acquire the global light transport characteristics of a translucent object. But many other image-based acquisition and rendering techniques are able to record and reproduce at least some aspects of subsurface scattering objects.

Levoy and Hanrahan [1996] capture light fields, [Gortler et al. 1996] capture Lumigraphs which describe the flow of light through unobstructed space with fixed illumination. Both methods have in common that they are able to reproduce the appearance of translucent objects correctly but only for a fixed set of viewing positions and lighting conditions. Surface Light Fields as proposed by [Miller et al. 1998] and [Wood et al. 2000] assign RGB values to every ray leaving every point on the surface. This provides sufficient information to construct images of the object from arbitrary viewpoints but for a single fixed illumination. In contrast to that, reflectance fields [Debevec et al. 2000] and polynomial texture maps [Malzbender et al. 2001] capture objects from a single viewpoint but illuminated with a set of different point-lights.

Environment matting [Zongker et al. 1999, Chuang et al. 2000] allows recording of strongly translucent or transparent objects (such as glass) with all reflection and refraction properties for a single viewpoint. Objects acquired this way can then be placed into another environment using environment compositing, where it will reflect and refract light from the new scene. This technique was also combined with reflectance fields by [Matusik et al. 2002] to acquire an image base representation of translucent and refractive objects including their 3D shape. Even though some of the above techniques use varying illumination while recording the object, they all illuminate the whole object with distant light sources. Therefore they are unable to reproduce local changes in the illumination such as textured light sources and shadow boundaries projected onto an object.

Masselus et al. [2003] capture the reflectance field of a scene for a fixed viewpoint and arbitrary illumination parameterized as a 4D incident light field. This allows to light the scene with arbitrary light sources including local illumination within the relatively coarse resolution limit of the incident light field.

[Jensen et al. 2001] measure the scattering parameters for common materials

such as marble or skin. They illuminate a single point on the object's surface with a tightly focused beam of white light and measure the radiant exitance across the entire surface using a digital video camera. From this they compute the absorption coefficient σ_a and the reduced scattering coefficient σ'_s. Their results allow modeling of homogeneous objects made from these materials. Objects made of arbitrary, inhomogeneous material are beyond the scope of their measurement approach.

3.3 Rendering

A variety of rendering techniques such as finite element methods [Rushmeier and Torrance 1990, Sillion 1995, Blasi et al. 1993], bidirectional path tracing [Hanrahan and Krueger 1993, Lafortune and Willems 1996], photon mapping [Jensen and Christensen 1998, Dorsey et al. 1999], Monte Carlo simulations [Pharr and Hanrahan 2000, Jensen et al. 1999], or diffusion [Stam 1995, Stam 2001] are able to simulate the effects of subsurface scattering based on physical principles. Although these techniques are able to provide high quality results, the long time required to compute a solution often prohibits interactive rendering. Subsurface scattering can also be integrated in the framework of precomputed radiance transfer [Sloan et al. 2003].

Jensen et al. [2001] introduced a practical BSSRDF model that combines a dipole point source diffusion approximation for multiple scattering and an extension of the Hanrahan and Krueger BRDF model [Hanrahan and Krueger 1993] for the local, directional dependent effects caused by single scattering. The simplicity of the model implies drastically reduced rendering times compared to a full Monte Carlo simulation. Although the dipole approximation is only correct for the surface of a homogeneous, infinite half-space, it yields visually convincing results. The degree of realism can be further improved by adding a modulation texture to simulate spatially varying materials.

Rendering with BSSRDFs is computationally expensive since contributions from the whole surface of the object must be taken into account. Jensen and Buhler [2002] sample therefore the irradiance first and hierarchically evaluate the diffusion approximation in a second pass. Hao et al. [2003] use a local illumination model and precompute the transfer factors for all vertex neighborhoods (i.e. they compute it for all vertices within a certain distance to an "incident" vertex). Lensch et al. [2003b] precompute the local response as filter kernels for an illumination map and the global response as a vertex-to-vertex throughput factor matrix. Both terms are then evaluated in parallel at rendering time. Carr et al. [2003] use a multi-resolution atlas to organize a precomputed hierarchy of precomputed subsurface scattering links and evaluate it at rendering time using 3 passes on the

graphics board.

None of the above rendering methods except the one from Jensen et al. [2001] are based on extensive sets of measured input data. The main focus of our work is therefore to capture the exact behavior of real translucent objects including effects that cannot be captured by simplified models such as the dipole approximation (e.g., cracks in an object, volumetrically varying properties or hollow objects). The acquired model is comprehensive and versatile. We also discuss how it can be integrated into a variety of the above rendering algorithms and present the acquired models rendered by the method of Lensch et al. [2003b].

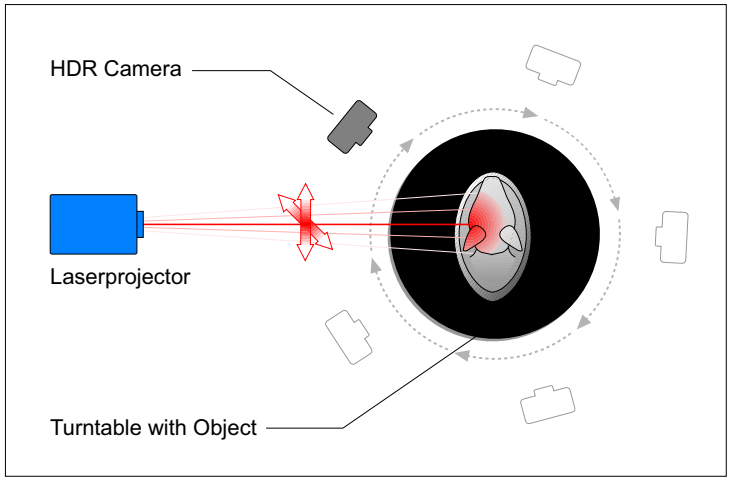
Chapter 4
Acquisition

Acquisition of the diffuse reflectance function R_d of an object seems to be a straightforward task: illuminate every point on the surface of the object and record its impulse response from all possible viewing directions. In practice however limitations of the available hardware, from the acquisition of the 3D shape of the object to the laser and camera systems that were used, make the acquisition a challenging task. This chapter first gives a short overview over the whole acquisition process. After that each individual part of the acquisition system with its features and drawbacks is described.

4.1 Measurement Overview

Figure 4.1 shows an overview over the DISCO setup. The target object is illuminated by a laser projection system that sweeps a laser spot over the object's surface (see Section 4.2). A high-dynamic range video camera (see Section 4.3) records an image for every location of the laser spot to capture the object's impulse response to the illumination. Since the camera only produces grey-scale images, each surface location has to be sequentially illuminated with a red, green and a blue laser in order to obtain color images (see Figure 4.2). With this setup of we can obviously capture color information only at three distinct wavelengths instead of the whole spectrum.

The object itself is placed on a turntable in order to illuminate as many surface regions as possible. After a complete sweep over the object's surface, the camera can be placed manually at different locations relative to the projection system to minimize occlusion and to record the object from all sides. For each view (a combination of camera and turntable position) an additional image is taken where the object is fully illuminated by spotlights. These images serve as silhouette



```
for all camera positions
  for all turntable rotations
    record silhouette image
    for all laser colors
      for all projector positions
          record impulse response image
```

Figure 4.1: Top: a schematic of the acquisition setup, Middle: a photograph of the acquisition setup in the MPI photo lab, Bottom: pseudo code of the acquisition process.

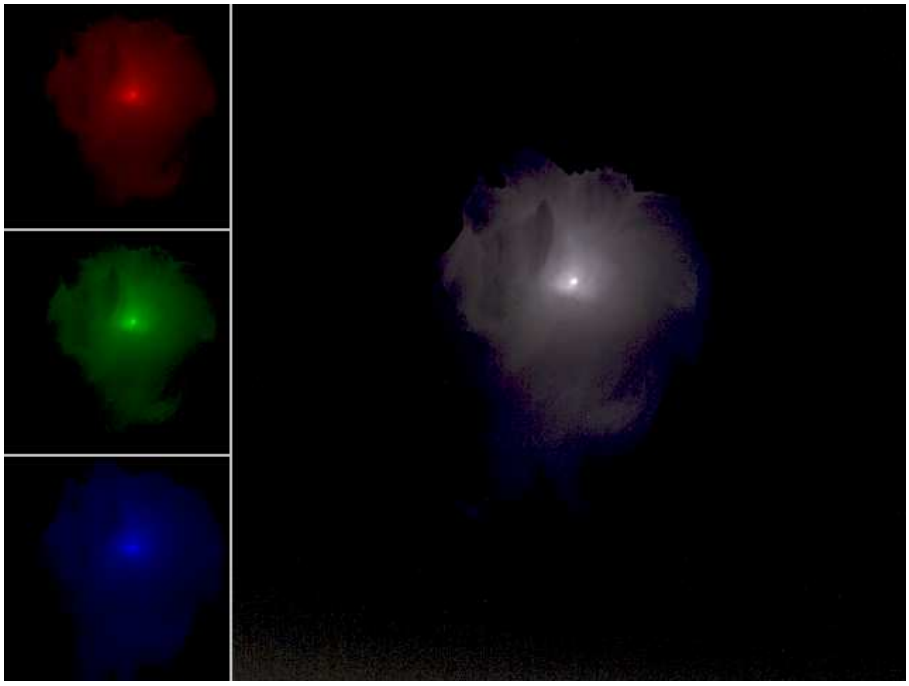


Figure 4.2: ***Left:*** **acquired high-dynamic range images with the red, green and blue lasers.** ***Right:*** **combined color image.**

images to register the objects position with respect to the camera (see Figure 4.3 for an example of a silhouette image).

The whole acquisition process runs at 15 fps which is half the speed of the camera. This is because we discard every other image and use the time in which the discarded images are taken to move the laser. The acquisition order in pseudo code is shown in Figure 4.1 on the bottom.

4.2 Laser System

The laser system we used was custom-built to our specifications by [Omicron-Laserage]. It consists of three individual lasers: a red diode laser and green and blue solid-state lasers with wavelengths of 635 nm, 532 nm, and 476 nm respectively. Each laser has an optical power of about 10 mW. All three lasers are coupled into a single optical fiber to ensure exact alignment of the resulting beams (see Figure 4.4). A collimator at the other end of the fiber focuses the beam to a size of about 2 mm within the working range. Care is taken in the whole laser system to reduce laser speckle to a minimum in order to avoid measurement artifacts. The position of the laser beam is controlled with a programmable laser

Figure 4.3: Example of a silhouette image. The acquired high-dynamic range image has been turned into a black-and-white image with a simple threshold operation.

projection system depicted in Figure 4.5.

In practice, the target object is illuminated with a narrow beam of light with finite width. Regarding the influence of the incident angle ω_i on the outgoing radiance $L^{\rightarrow}(x_o, \omega_o)$ we can therefore distinguish between two cases. If x_o is located outside the beam incident at x_i we can assume that the full energy (modulated by the Fresnel factor $F_t(\eta, \omega_i)$) enters the object at x_i and is scattered diffusely inside. The angle of incidence ω_i influences then only $F_t(\eta, \omega_i)$, the area foreshortening has to be omitted.

Within the area of incidence of the beam, the area foreshortening is definitely important since it scales the irradiance at each point. The shape of the peak around the incident beam will also change to some degree depending on the beam width and the angle of incidence ω_i.

Overall, the ideal measurement setup would therefore illuminate an object with an infinitesimal beam of light along the surface normal in order to avoid all these problems. But as this is not practical we have to compromise. One good method to limit the amount of error introduced is to consider only samples where the angle of incidence is below a certain threshold. The introduced noise must be dealt with in later stages of the processing pipeline.

An unknown factor in our measurements are the Fresnel factors $F_t(\eta, \omega)$. They

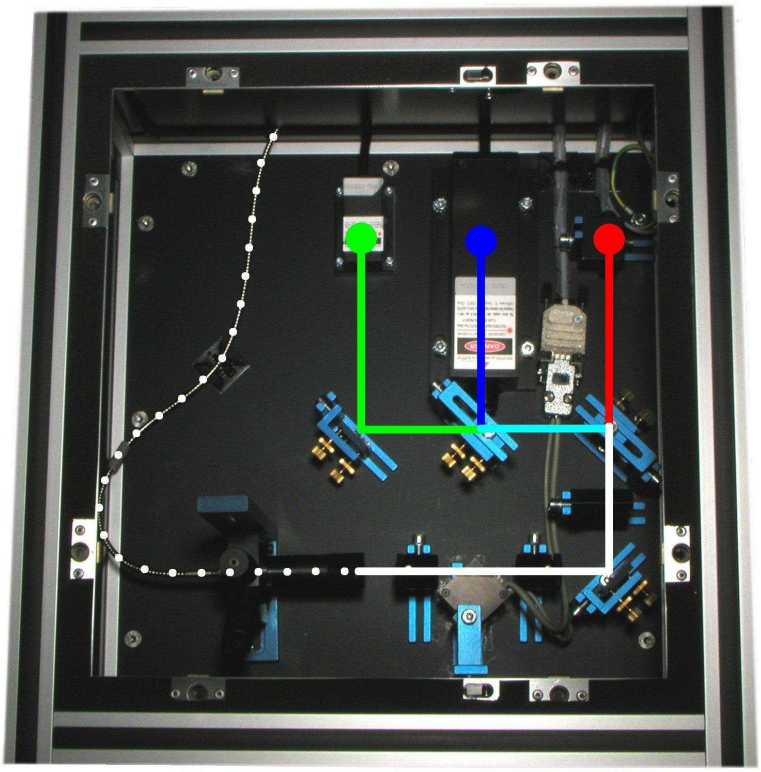


Figure 4.4: The internals of the laser system. It consists of a red diode laser and green and blue solid state lasers. Several semi-transparent mirrors align the three beams together before they are fed into the fiber.

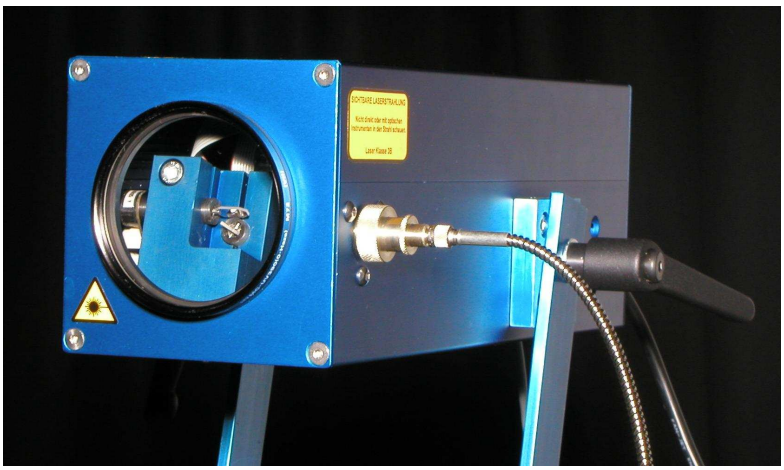


Figure 4.5: Photograph of the laser projection system. The two mirrors that deflect the laser beam and the fiber cable on the right side can be clearly seen.

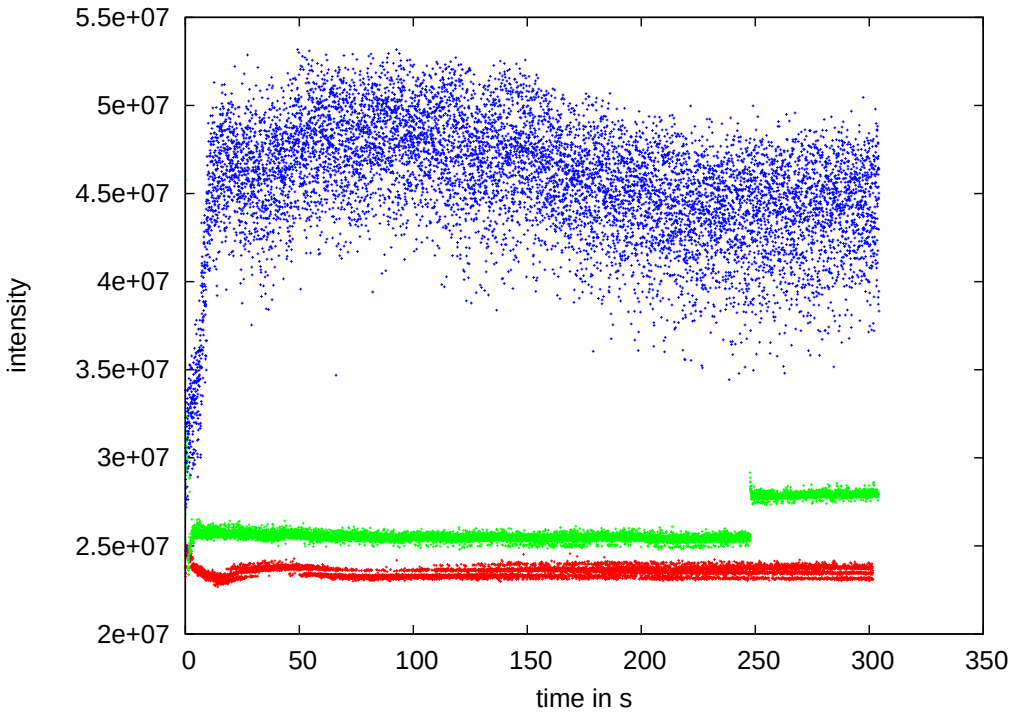


Figure 4.6: Intensity of the red, green and blue laser. The x axis shows the time in seconds after powering on the laser. The intensity as reported by the HDR camera is on the y axis.

represent only an approximation to the real behavior of a surface [Schlick 1994] and require at least knowledge of the relative index of refraction η. We follow here Jensen et al. [2001] and set $\eta = 1.3$ but acknowledge that this is not exact.

Godin et al. [2001] have shown that 3D laser range scanning of translucent objects suffers from a systematic bias since the peak location is moved for varying viewing and lighting directions. Since we also rely on peak detection, we expect that our measurements are also influenced by this behavior to some degree.

Figure 4.6 shows the development of the laser intensity over time. The red diode laser needs about a minute to stabilize its intensity and then maintains it with relatively little noise. The blue laser starts off relatively dark but reaches its final intensity quickly. However it is clearly visible that there is a lot of noise in its intensity. The green laser shows the same starting phase as the blue laser but its intensity remains quite stable and relatively noise free. But there is a clear jump in the intensity about 250 seconds after powering on the laser. This is due to mode hopping, a phenomenon which is hard to eliminate in solid state lasers. Mode hopping and also the variations in intensity could be either avoided by using more expensive lasers or could be taken into account by an additional calibration setup.

A photo diode connected to the measurement computer could for example monitor the laser intensity during the acquisition and thereby make it possible to normalize it afterwards. Another problem is the warmup phase of the lasers. Especially the blue laser shows a strong variation in intensity in the first few seconds after switching it on. The measurement software accounts for this by simply waiting one minute after powering on a laser.

4.3 High-Dynamic Range Video Capture

High-dynamic range (HDR) video capture can be achieved with specialized cameras and/or software (see Kang et al. [2003] for an overview). A specific requirement of our measurement setup is the quite extreme dynamic range for which we need linear radiance values without interfering quantization artifacts or blooming. In addition, care must be taken that the interesting features in the vicinity of the laser spot are not masked by lens flare or other artifacts caused by the high scene contrast. The use of a high quality lens is therefore mandatory.

Cameras with standard logarithmic CMOS chips are in principle well suited for the task at hand due to the exponential fall-off of subsurface scattering. Linear response can be achieved by a simple HDR calibration step [Debevec and Malik 1997, Robertson et al. 1999]. These cameras suffer however often from strong quantization artifacts as the images are quantized to 8–12 bits. Newer technologies such as the Photonfocus LINLOG technology should improve this situation.

In DISCO, we use a Silicon Vision Lars III high-dynamic range video camera equipped with a Jenoptik Lametar 2.8/25 mm lens (see Figure 4.7). The camera records gray-scale images with 768×496 pixel resolution at up to 30 fps and returns linear radiance values over a range of approximately 120 dB. The basic principle of this camera is that each pixel decides at fixed time steps (powers of two) whether it already received sufficient irradiance. This corresponds to an estimate whether the pixel will be overexposed at the next time step given constant irradiance. Each pixel records the exposure time and the amount of charge collected within this time. From this information linear radiance values are computed with high precision throughout the dynamic range.

4.4 Geometry Acquisition

Optical 3D scanning of translucent objects is challenging due to the non-local light reflection [Levoy et al. 2000, Godin et al. 2001, Matusik et al. 2002]. Even a human observer can have difficulties to visually detect fine shape details. To

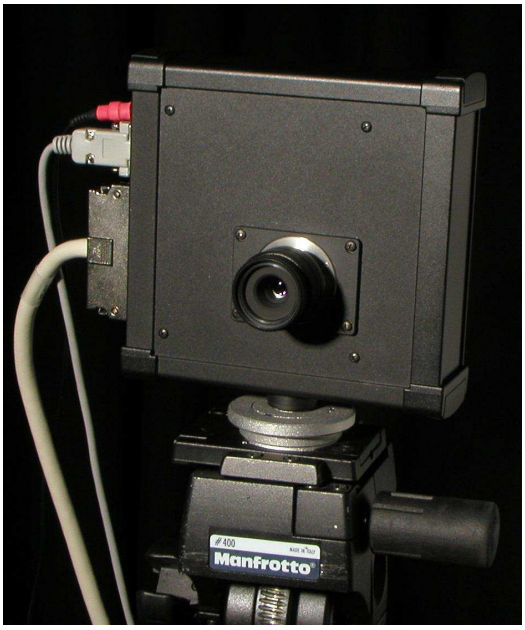


Figure 4.7: Photograph of the Silicon Vision Lars III high-dynamic range video camera

achieve an almost Lambertian reflection which is ideal for optical 3D scanning techniques the objects can be covered with a fine layer of matte spray paint or dust. Ideally, the surface detail should not be modified and the object should not be destroyed. We use a Minolta VI-910 laser scanner to acquire the geometry of the objects. We tried several sprays and performed measurements with the 3D scanner to see how the amount of noise in the scanning result changes and how much the geometry of the object changes. The results in Figure 4.8 confirm completely the work of [Godin et al. 2001]: Measurement noise and deviation increases with the translucency of the object.

MET-L-Check developer spray consists of very fine particles as it is designed to get into fine cracks in weldings in order to make them visible. Therefore it produces very thin coatings. It is also water dissolvable so it can be washed off afterwards.

Cyclododekan spray has the nice property that it dissolves completely and without residue within several days after its application. But it produces little flocks on the surface which add a significant amount of noise to the result compared to the MET-L-Check spray coating.

We also made sure that the sprays do not react with the object's material so that the object's appearance does not change.

In the end we used the MET-L-Check developer spray to cover the objects.

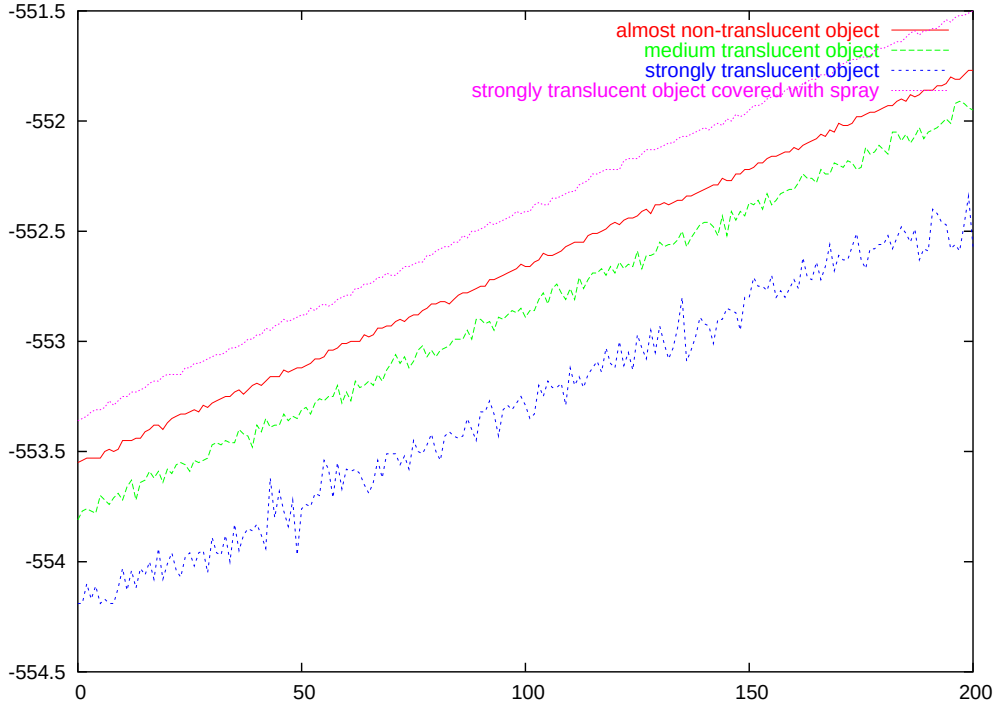


Figure 4.8: Difference in depth measurement between objects with different translucency. The plot shows on the y-axis the distance of the object to the scanner for one scan line. The deviation and the noise in the depth measurement increase with the translucency of the object. The difference between the object covered with spray and the almost not translucent object can be due to the thickness of the applied spray or because the almost not translucent object causes a little deviation in the measurement.

The measurements in Figure 4.8 show that noise is almost completely eliminated by the application of the spray and that the depth error is at most 0.3 mm.

The resulting point clouds were then transformed into the final triangle meshes using commercial geometry-processing tools.

4.5 Photometric Calibration

For the photometric calibration, we rely on the overall linearity of the camera output and assume that the laser power is constant over time. We then need to perform a white balancing step for the individual lasers, taking the spectral response of the camera into account. To this end, we sequentially illuminate a white, scattering calibration target with the three lasers and sum up the contribution of all image pixels.

Chapter 5
Data Processing

The data acquired by our acquisition system described in Chapter 4 can not be used directly to render subsurface scattering models. First of all the acquired data is only a series of impulse response images from which we need to extract the diffuse reflectance function R_d of the object. Second, the locations hit by the lasers do not correspond to the vertex positions of the 3D mesh, requiring resampling of the input data. Therefore a considerable amount of data processing is necessary after the acquisition process is completed. In this chapter we describe the steps that we take to turn the stream of input images into a final model of a subsurface scattering object.

First we describe how the series of images is stored efficiently so that it can be accessed fast enough to keep the rest of the pipeline busy. Then we show which information needs to be precomputed before the whole stream can be processed into our hierarchical model (see Section 2.4). The remaining sections of this chapter present the details of each part of the hierarchical model.

5.1 Data Compression and Efficient Processing

A typical model acquisition yields an uncompressed data volume of several hundreds of gigabytes. Since the biggest part of the impulse response images is dark, lossless data compression (e.g. using bzip2 compression) can reduce the size of the raw data by a factor of about 30 for a typical data set. The acquisition software compresses the data during acquisition to a few tens of gigabytes thus eliminating the need for large disk drives or RAID systems. The software keeps track of where to find a particular image in the compressed stream so that random access to arbitrary images remains possible. Nevertheless it is necessary to use efficient

algorithms to process the input data.

Each complete acquisition consists of a small number of views (combinations of camera position and turntable rotation, typically 20–30 views). Most of the information required in further processing steps is constant per view since the camera observes the object always from the same perspective. This includes the position of the object and the laser projector relative to the camera as well as various per-pixel parameters. We therefore precompute this per-view information to speed up further computation (see Section 5.2.1).

In estimating our hierarchical model, we then evaluate all images for a given turntable and laser projector position at the same time. These images show the object under identical illumination conditions and contain all information that is available for this illumination condition. We first decide whether the image tuple is valid (i.e. whether the laser spot is visible), and can then efficiently resample the data using the precomputed information. Our current implementation processes the input data streams with up to 50 fps on a PC with a 3 GHz Intel Xeon CPU which is more than three times the speed of the acquisition.

5.2 Precomputed Per-View Information

5.2.1 Per-Pixel Parameters

We precompute the Fresnel term and assign all pixels in the input image a confidence value based on the viewing and lighting directions. This confidence value is then used as weighting factor for the input data. We also reject at this stage all pixels that are close to a specular highlight or seen under grazing angles and generate the mapping from our texture atlas into the input images. The mapping from vertices to image coordinates and the visibility are also precomputed.

5.2.2 Position of the Target Object

We use a hardware-accelerated silhouette-based registration algorithm [Lensch et al. 2001] to recover the pose of the target object relative to the camera for each view. In order to use this algorithm on our silhouette images, we first have to determine the intrinsic parameters (especially the field of view) of the high-dynamic range camera [Bouguet 2003]. The algorithm rotates and translates the 3D model until its 2D projection (taking the intrinsic parameters of the camera into account) matches the silhouette of the object in the photograph as close as possible. Figure 5.1 illustrates the concept of the algorithm.

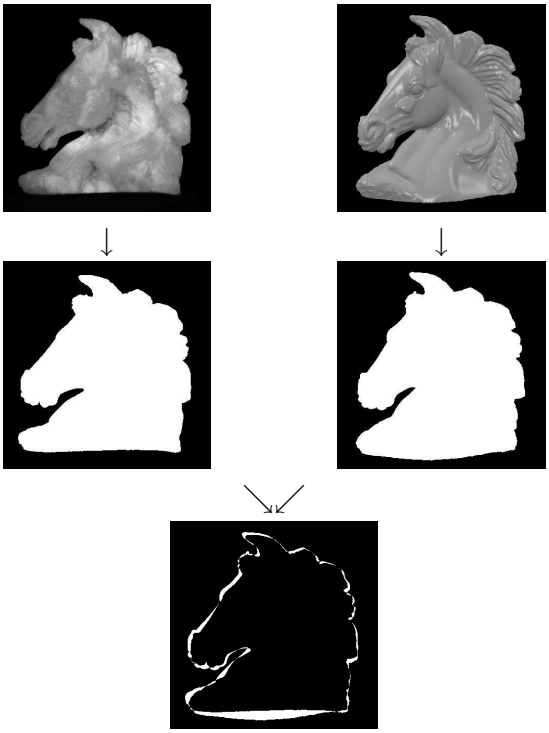


Figure 5.1: The acquired high-dynamic range image *Top Left* is automatically segmented into a silhouette image *Middle Left* via a normal thresholding operation. The *Top Right* image shows a view of the corresponding 3D model in a slightly different pose. The *Middle Right* image is a black and white projection of the 3D model in 2D. The *Bottom* image is the result of a pixel-wise XOR of the two Middle images. It shows all areas where the pose of the 3D model does not match the pose of the object in the high-dynamic range image correctly. The amount of white pixels in the Bottom image is a measurement of the mismatch.

5.2.3 Estimation of the Laser Position

In order to determine the Fresnel term and the confidence values described in Section 5.2.1, we need to know the position of the laser projection system in camera space. Estimating the position of the center of projection of the laser projector is possible with a non-degenerate set of laser spots seen for each view. Non-degenerate in this context means that a sufficient number of highlights on the surface have to provide enough depth information to estimate the distance of the laser projector to the object reliably. This corresponds to an extrinsic camera calibration [Faugeras 1993], the center of projection of the laser projector is calculated from the deflection settings of the laser and the corresponding 3-D spot positions.

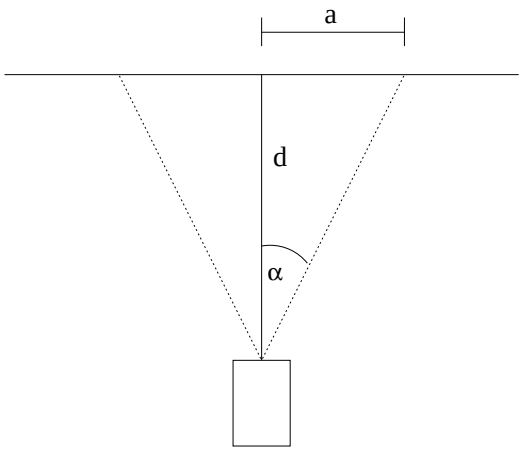


Figure 5.2: Measurement of the laser projector's step size. The full angular range (2α) of the laser projector can be calculated as $2 \cdot \arctan(\frac{a}{d})$.

After measuring the intrinsic and extrinsic parameters of the camera and calibrating the position of the object via the silhouette matching algorithm we are able to obtain the 3-D coordinates in the camera space for highlights seen on the object's surface (i.e., its center of projection) in the 2-D camera images.

The laser projection system can move the laser across a fixed angle in 4096 steps in each direction. In order to relate the values at the digital input of the projector to the spot locations on the object's surface, we need to determine these fixed angles first. To do this, we set the projector to point a laser beam perpendicular onto a planar surface at distance d with the deflection mirrors set to center positions (see Figure 5.2). Then we measure the distance a that the laser moves when we set it to full deflection. The full angular range of the laser projector in this direction can then be calculated as $2 \cdot \arctan(\frac{a}{d})$.

After that we need to determine the projection Matrix P of the laser projector that relates the values at its digital inputs (u, v) to 3D coordinates (x, y, z).

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \underbrace{\begin{bmatrix} r_1^T & t_1 \\ r_2^T & t_2 \\ r_3^T & t_3 \end{bmatrix}}_{P} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

where

$$r_j = \begin{bmatrix} r_{j,1} \\ r_{j,2} \\ r_{j,3} \end{bmatrix}, \quad j = \{1, 2, 3\}.$$

Simple algebraic operations show that

$$\frac{u}{w} = \frac{\left[r_1^T t_1\right] \left[xyz1\right]^T}{\left[r_3^T t_3\right] \left[xyz1\right]^T}$$

and

$$\frac{v}{w} = \frac{\left[r_2^T t_2\right] \left[xyz1\right]^T}{\left[r_3^T t_3\right] \left[xyz1\right]^T}.$$

Since the laser projection system only has 2 degrees of freedom we can set

$$w = t_3 = 1$$

and then calculate

$$\begin{aligned} \left[r_3^T t_3\right] \left[xyz1\right]^T u &= \left[r_1^T t_1\right] \left[xyz1\right]^T \\ u + \left[r_3^T\right] \left[xyz\right]^T u &= \left[r_1^T t_1\right] \left[xyz1\right]^T \\ u &= \left[r_1^T t_1\right] \left[xyz1\right]^T - \left[r_3^T\right] \left[xyz\right]^T u \end{aligned}$$

and

$$v = \left[r_2^T t_2\right] \left[xyz1\right]^T - \left[r_3^T\right] \left[xyz\right]^T v$$

Therefore we can express the elements of the projection matrix as the unknowns of a system of linear equations filling in the coordinates of the laser spots on the surface (x, y, z) and the respective mirror deflection settings (u, v):

$$\begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} x_1 & y_1 & z_1 & 1 & 0 & 0 & 0 & 0 & -x_1u & -y_1u & -z_1u \\ 0 & 0 & 0 & 0 & x_1 & y_1 & z_1 & 1 & -x_1v & -y_1v & -z_1v \\ x_2 & y_2 & z_2 & 2 & 0 & 0 & 0 & 0 & -x_2u & -y_2u & -z_2u \\ 0 & 0 & 0 & 0 & x_2 & y_2 & z_2 & 2 & -x_2v & -y_2v & -z_2v \\ x_3 & y_3 & z_3 & 3 & 0 & 0 & 0 & 0 & -x_3u & -y_3u & -z_3u \\ 0 & 0 & 0 & 0 & x_3 & y_3 & z_3 & 3 & -x_3v & -y_3v & -z_3v \\ & & & & & \vdots & & & & & \end{bmatrix} \begin{bmatrix} r_{1,1} \\ r_{1,2} \\ r_{1,3} \\ t_1 \\ r_{2,1} \\ r_{2,2} \\ r_{2,3} \\ t_2 \\ r_{3,1} \\ r_{3,2} \\ r_{3,3} \end{bmatrix}$$

We then use singular value decomposition (SVD) [Press et al. 1994] to solve this system of linear equations. This way we get a projection matrix which is scaled by some linear factor n:

$$n \cdot \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & t_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & t_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & t_3 \end{bmatrix}$$

In a standard projection matrix all eigenvalues are 1. This enables us to determine n and eliminate it. We perform another singular value decomposition on the following matrix.

$$\begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} \\ r_{2,1} & r_{2,2} & r_{2,3} \\ r_{3,1} & r_{3,2} & r_{3,3} \end{bmatrix}$$

Then we fit a factor of $\frac{1}{n}$ to the three singular values to bring them as close to 1 as possible. Finally, we rescale the whole matrix with $\frac{1}{n}$ and can extract the position of the laser from $t_1 \ldots t_3$.

5.3 Global Term Postprocessing

An object's diffuse subsurface reflectance function $R_d(x_i, x_o)$ away from the point of incidence in our model is represented by a global term. The discrete representation of this global term is the throughput factor matrix $\mathbf{F}$. The shape of the object is represented by a triangular surface mesh. $\mathbf{F}$ is filled with discretized impulse response functions - each impulse response function per measured (lit) vertex is stored in a column of $\mathbf{F}$. The outgoing radiosity ($B(\vec{x}_o)$) can then be calculated as the matrix-vector product of the irradiance vector ($E(\vec{x}_i)$) and the matrix $\mathbf{F}$:

$$B_v = \sum_u E_u \mathbf{F}_{u,v}$$

Light transport is of course symmetric, therefore we can also fill the corresponding row of $\mathbf{F}$. Normally $\mathbf{F}$ is only partially filled, since only a fixed set of surface locations x_i was illuminated and because only a subset of exiting surface locations x_o has been observed due to incomplete coverage and occlusion.

We estimate the complete matrix $\mathbf{F}$ based on the existing measurements in the postprocessing step. To interpolate missing values correctly, we need to take the distribution of energy within the material into account. In homogeneous material the energy falls off exponentially away from the point of incidence according to the diffusion approximation [Jensen et al. 2001]. But our method is aimed at inhomogeneous objects with behavior that deviates from the smooth diffusion approximation.

5.3.1 Resampling

The data that results from the acquisition is already a representation of subsurface reflectance functions because of our choice of measurement method. We record the impulse response of the object to light incident at one point on the surface. A high-dynamic range image of the object is a scaled subsurface reflectance function for light entering at point x_i. But the data only consists of discrete samples on the image plane and also only covers the part of the surface which is in view. Since we already know the 3D geometry of the whole measurement setup (see Section 5.2.1 for details) we can project the image data onto the object's surface represented as a triangular surface mesh where each triangle covers a similar surface area.

To resample the image data into the global part of the hierarchical model, we look up the bi-linearly interpolated image intensity values at each vertex location. We also determine the position of the incident light on the mesh. The light is assumed to contribute to the three vertices of the triangle enclosing the point of incidence. The energy is distributed to the three vertices according to the barycentric coordinates of the incident light point inside the triangle. We combine measurements of the same subsurface reflectance function observed from different viewpoints and with varying laser color. The result of the resampling are the columns of the throughput factor matrix $\mathbf{F}$ in *RGB* color space. The matrix is missing some entries due to unobserved surface area and is completely missing some columns for vertices that were never illuminated by the laser. The interpolation of the matrix values for these cases is described next.

5.3.2 Interpolation

Filling missing values in the columns of $\mathbf{F}$ is the task of function interpolation on an irregularly meshed surface. In the diffusion approximation for homogeneous material the function is a sum of exponential decays which non-trivially depend on distance. For inhomogeneous objects the function is more complicated and the interpolation has to potentially fill in large unseen areas.

Individual missing entries $\mathbf{F}_{r,k}$ in columns k are filled where we observe the incident light point $\mathbf{F}_{k,k}$. Estimated values are weighted averages of existing values, keeping the measured values fixed. In particular, the following iteration is solved:

$$\mathbf{F}_{r,k}^{t+1} = \mathbf{F}_{r,k}^{t} + \alpha(1 - c_r) \sum_{n \in \mathcal{N}} \psi \left(\frac{\mathbf{F}_{n,k} - \mathbf{F}_{r,k}}{e_{n,r}}, \sigma \right)$$

The neighbors n of a vertex r are its one-ring neighborhood $\mathcal{N}$ connected with edges of length $e_{n,r}$. ψ is a least-squares weighting function with $\psi(x, \sigma) = 2\frac{x}{\sigma^2}$ and median filtering with $\psi(x) = sign(x)$. c_r serves as a confidence value in the current transfer coefficient $\mathbf{F}_{r,k}$ and $c_r = 1$ keeps the measurement fixed.

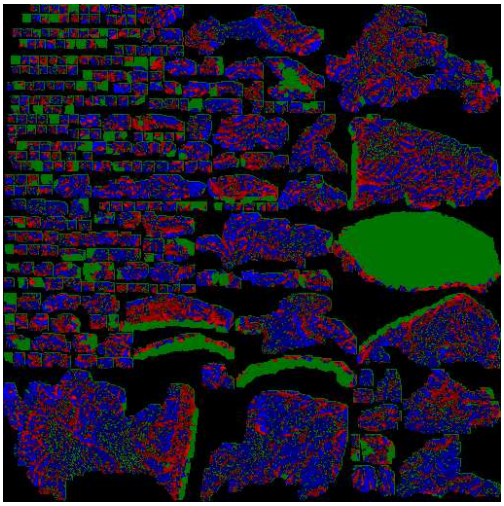


Figure 5.3: Recovered detail texture (color coded version of the green channel: blue areas become darker, red areas become brighter, green areas remain unchanged)

Subsurface reflectance functions for vertices (columns) which the laser did not reach are interpolated iteratively from neighboring vertices. The interpolation for missing complete columns of $\mathbf{F}$ uses the same filtering framework as the interpolation for missing entries within a column. The differences between the two are limited to different neighborhood definitions.

We make sure that $\mathbf{F}$ is symmetric by setting $\mathbf{F}^{t+1} = \frac{1}{2}\left(\mathbf{F}^t + (\mathbf{F}^t)^T\right)$ after each iteration as well as at the beginning and at the end of the filtering process.

5.4 Detail Texture

The detail represented in the throughput factor matrix $\mathbf{F}$ is limited by the resolution of the underlying mesh. We therefore increase the resolution of the subsurface reflectance functions with two kinds of textures. We estimate illumination dependent local high-resolution filter kernels (see Section 5.5) and a global illumination independent modulation texture T_ρ (see [Lensch et al. 2003b] for details). At rendering time, we multiply the global response with this modulation texture. Figure 5.3 shows an example of the recovered detail texture for the horse model.

Figure 5.4: The texture atlas with borders for the horse model.

5.5 Local Term Postprocessing

We follow Lensch et al. [2003b] and model local light propagation by fixed sized spatially varying filter kernels $K_{(u,v)}$ (7×7 pixels in all our examples) on a texture atlas. The filter kernels convolve the irradiance at texel locations for local energy transfer. We use a texture atlas with a border and with a low guaranteed upper limit for texture stretch (on the order of 2). Figure 5.4 shows an example texture atlas of size 512×512 pixel with borders. The filter kernels are estimated in the texture atlas domain based on images where the laser spot is visible.

5.5.1 Resampling

In general, the peak of the laser spot will fall between the discrete pixel locations in the texture atlas (texels). An example of this situation is shown in Figure 5.5. Therefore, if we map the peak of a filter kernel to the nearest neighbor texel of the laser spot, we introduce some error. Figure 5.6 (Top) shows an example of the peak of a filter kernel mapped to the nearest neighbor texel. All texels in an 8×8 pixel neighborhood around the peak location are plotted according to their 3D distance to the peak texel. A clearly structured fall-off is not visible.

If we however plot them according to their distance to the subpixel location of the laser spot (see Figure 5.6 Middle), the general shape of the exponential fall-off of subsurface scattering material is recognizable. Figure 5.6 (Right) shows an example of the dipole approximation of [Jensen et al. 2001] for comparison.

The measurement data also shows a strong variation due to surface detail at a given texel location. Surface detail is associated with a texel location and should

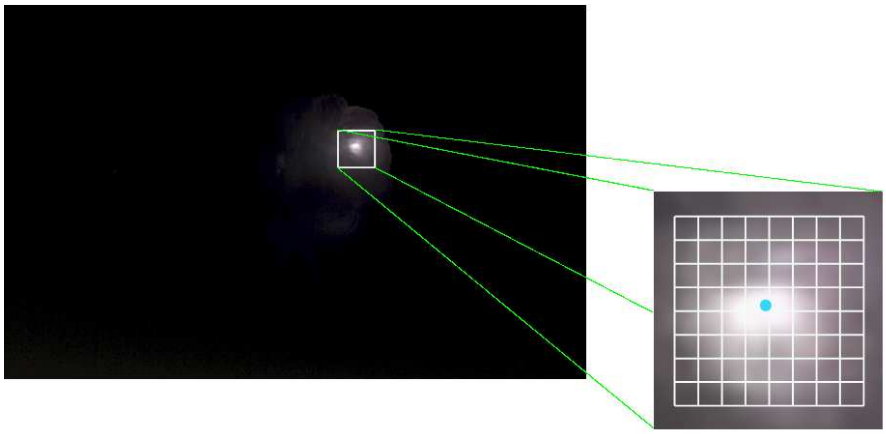


Figure 5.5: An acquired high-dynamic range image and the 8x8 pixel neighborhood around the point of incidence from which the filter kernels are computed.

not be resampled to preserve the sharpness of features.

In order to recover correct filter kernels, we therefore shift the peak location to all four texels neighboring the laser spot while keeping surface detail spatially fixed (see Figure 5.7 for an example). To seperate the illumination from the surface detail, we fit the approximation

$$f(d) = c_1 \cdot e^{\alpha_1 d} + c_2 \cdot e^{\alpha_2 d}$$

to the measured data. d is the 3-D distance from the location of the peak, c_1 and c_2 are fit per color channel. α_1 and α_2 are identical over all color channels. The difference between the measured data points and their approximation $f(d)$ is then encoded in a 8×8 pixel multiplicative correction texture. This texture can then be used to recover the surface detail after the location of the peak has been shifted. To compute the four filter kernels for the neighboring texel, we evaluate $f(d)$ at their respective center locations and divide the result by the correction texture. The four resulting filter kernels are then weighted according to their sub-pixel distance to the laser spot location.

5.5.2 Interpolation

In general, like for the global transfer matrix, we can not obtain a measurement for all the filter kernels, i.e. some kernels are not measured at all while others are measured partially. Therefore we need to do some interpolation to fill in the missing data.

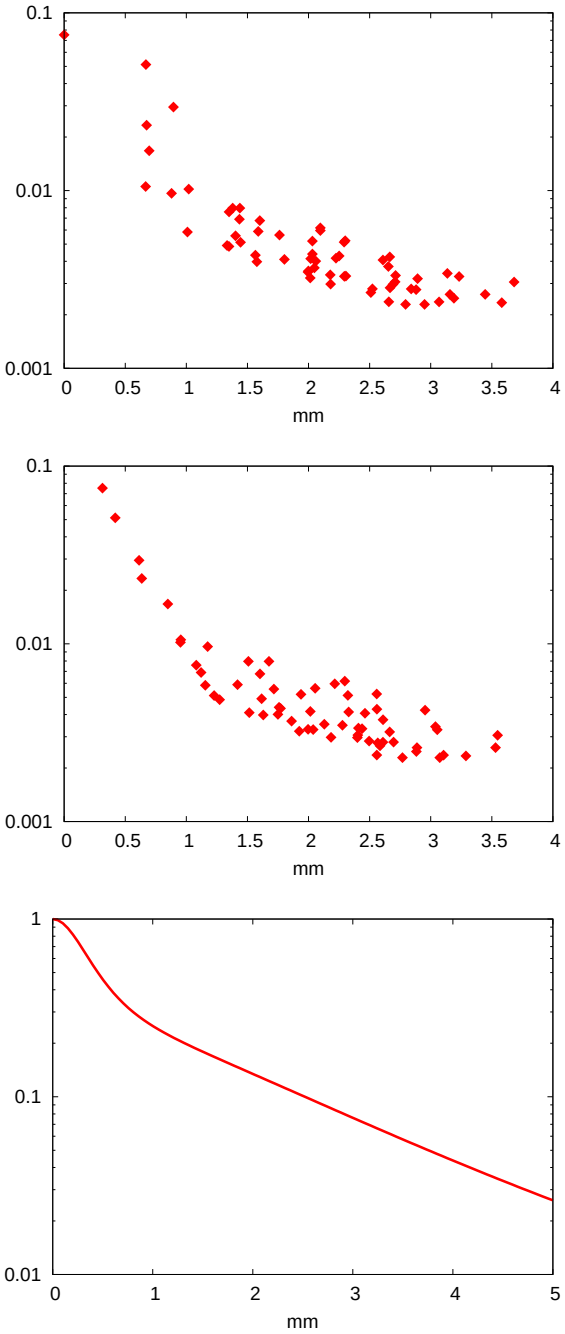


Figure 5.6: Building the filter kernels. ***Top:*** **Location of the laser peak mapped according to the distance to its' nearest neighbor texel. The plot shows monochrome luminance values from the 8×8 texel neighborhood around the highlight. The horizontal axis is the distance in mm to a neighboring texel in the texture atlas. The values show a fall-off with distance but no clear structure.** ***Middle:*** **If we plot the same data over the distance to the sub-pixel peak location, a fall-off resembling the sum of two exponential functions appears.** ***Bottom:*** **Plot of the dipole approximation for the material properties of marble from Jensen et al. [2001] for shape comparison.**

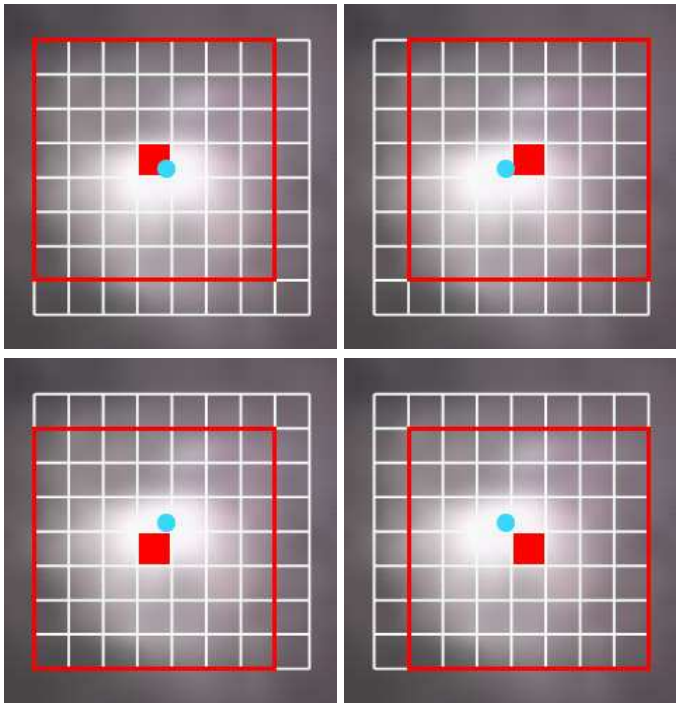


Figure 5.7: To recover the filter kernels, the peak location is shifted to all four neighboring texels around the point of incidence. The four resulting filter kernels are weighted according to their sub-pixel distance to the laser spot location.

Interpolation of filter kernels is performed by vector filtering over the surface of the object. Each 7×7 filter kernel is represented as a 49-vector. Since we have already fitted an exponential function (see Section 5.5.1) to the measured data of each filter kernel, completion of missing data within a filter kernel is simply done by evaluating the exponential function at the missing points.

For filter kernels that were not measured at all we use the same filtering framework as for the transfer coefficients in the matrix $\mathbf{F}$ but on the texture atlas domain instead of the triangular mesh. The neighborhood $\mathcal{N}$ in the texture atlas contains a texel's four-neighbors except for texture atlas boundaries. The boundaries need to be treated specially in order to ensure the correct neighborhood information. Each boundary texel links therefore to a corresponding non-boundary texel in a different map of the atlas that is used instead of the boundary texel during filtering. Filter kernels also differ in size since different areas in the atlas are not isometric.

Chapter 6

Results

We acquired the following objects to validate our approach: An alabaster horse sculpture, a rubber duck, and a starfruit (carambola) The alabaster horse shows a strong variation in its subsurface scattering properties. It consists of regions with varying translucency and contains cracks inside the material and has a very complex geometry. The base material of the rubber duck is uniform, strongly scattering plastics. The beak and the eyes are painted on the outside with glossy, opaque paint. Although the duck is hollow inside, incoming light is so strongly diffused at the surface that the assumption of multiple scattering remains valid. The starfruit is an example of a subsurface scattering biological object with a relatively specular reflecting surface. Figure 6.1 shows all objects under indoor illumination and illuminated by all three lasers at the same time (i.e. the object's impulse response). Table 6.1 summarizes important properties of the models.

We also did some comparisons to see how the final models and renderings cor-

	Horse	**Duck**	**Starfruit**
# input views	24	25	20
# input images	1.065.744	541.125	401.220
input size (compressed)	31G	14G	12G
acquisition time	20.5h	11.25h	8h
# vertices	8924	5002	5001
# filter kernels	82.390	115.151	112.538
processing time (resampling)	7.8h	3.6h	3.4h

Table 6.1: Some statistics about the acquired models.

Figure 6.1: The test objects under indoor illumination (top row) and illuminated by all three laser at the same time (bottom row)

Figure 6.2: ***Left:*** **Rendering of the horse model acquired by our DISCO method. Differences in material are clearly visible when the object is lit from behind.** ***Right:*** **Photograph of the horse sculpture taken under similar lighting conditions.**

respond to the real world. Figure 6.2 shows a rendering picture and a photograph taken under nearly the same lighting conditions. The alabaster horse has areas of differing translucency. There are more opaque regions at the head around the eyes and at the bottom of the model, while others like the muzzle and the mane are quite translucent. There is a rather sharp boundary between the opaque and translucent regions in the nose. This is due to a crack in the material. The model reproduces all these volume effects. Figure 6.3 demonstrates how the local light transport adds surface detail and gives an impression of the various optical densities at the object's surface. The structure is also visible in the global throughput factor matrix albeit smoother. Figure 6.3 also shows how the sharp boundary in lighting is preserved by the local filter kernels and how it is smoothed away by the global light transport. The side-by-side comparison of a rendering of the model and a photograph as in Figures 6.2 and 6.4 shows that our method recovers fine structure detail. The highly translucent veins in the stone as well as the strongly diffuse patches which are visible in the photograph are also present in the model. In figure 6.4, the slightly brownish region in the center of the horse is also captured well by our model, both, in color and shape.

The duck model in Figures 6.5 and 6.6 shows how our method deals with completely heterogeneous objects without any knowledge about the distribution of material inside the object. The duck is made of rubber and it is hollow on the inside. The beak and the eyes are painted on with opaque color paint, and the wings

Figure 6.3: ***Left:*** **irradiance,** ***Middle:*** **local light transport by filter kernels,** ***Right:*** **global light transport by the throughput factor matrix.**

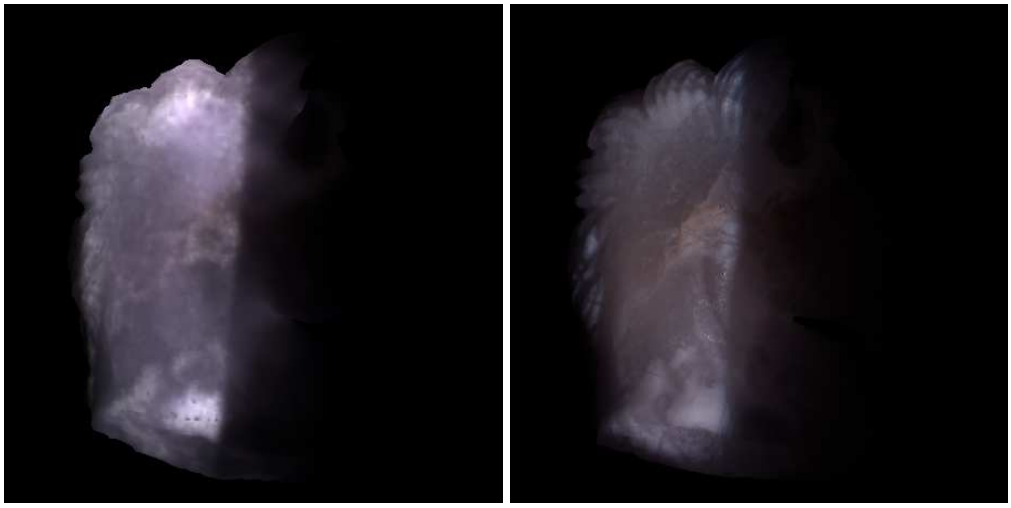


Figure 6.4: ***Left:*** **Combined rendering of local and global term. The local light transport adds surface detail and gives an impression of the material properties at the object's surface whose basic structure is also visible in the global term.** ***Right:*** **Photograph of the real object under similar conditions. A slide projector was used to produce the sharp shadow boundary.**

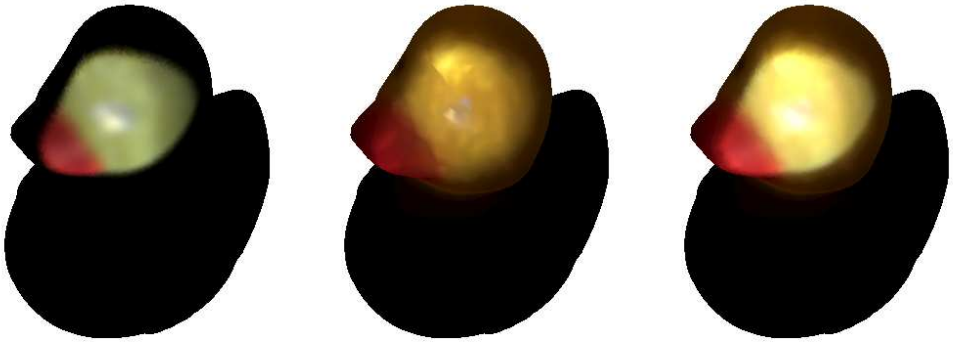


Figure 6.5: The duck model illuminated with a spotlight. *Left:* Local light transport by filter kernels. *Middle:* Global light transport by throughput factor matrix. *Right:* Combined rendering.

block more light than the rest of the body because of their different geometry. The model captures these major deviations from homogeneous behavior while being smooth overall. Figure 6.6 shows that the neck provides a border for global light transport: light illuminating the head is scattered mainly inside the head and light illuminating the body is scattered mainly inside the body.

The starfruit is a good example of a biological object with geometrically complex shape. It also features a relatively specular surface. The model in Figure 6.7 is of good fidelity despite missing a lot of input data due to the complex shape. Additional input images would fill the gaps better than our data interpolation technique.

6.1 Discussion

Like with every image-based rendering method the quality of the input images determines the amount of post-processing needed and ultimately the quality of the final model. Surface coverage is limited by the classical stereo vision occlusion problem. There is also a trade-off between making additional images and increased acquisition time. Maybe a view planning algorithm could help here. As discussed in Section 4.2 our low-cost laser system varies a lot in intensity requiring additional smoothing during the interpolation. This variation and the observed intensity drift due to thermal effects lead to color artifacts in the final renderings.

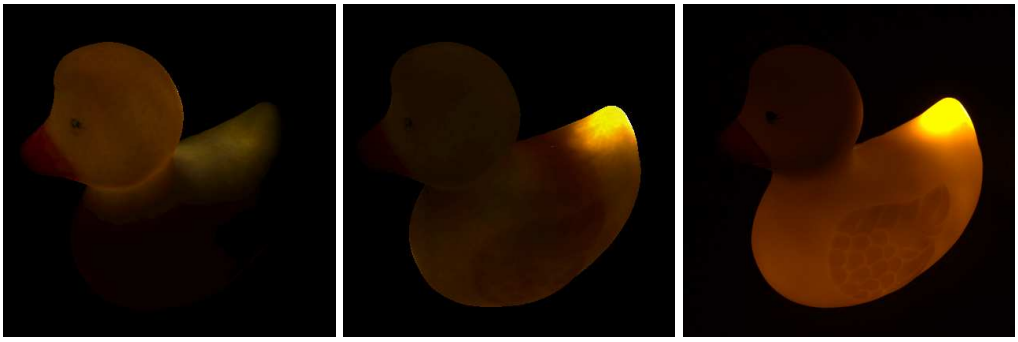


Figure 6.6: Details of the duck model. *Left:* Head of the duck illuminated with a small spotlight from behind. The beak and the eyes are clearly visible. *Middle:* Body of the duck illuminated with a small spotlight from the back. *Right:* Photograph of the duck taken under similar conditions. A slide projector was used to generate the small spotlight which illuminates the tail of the duck.

Figure 6.7: Starfruit model with both the global transfer function and the local term (7×7 filter kernels). *Left:* Illumination from the front right. *Right:* Illumination from behind.

Chapter 7

Conclusion and Future Work

We presented a complete framework for the acquisition and rendering of real-life, inhomogeneous, subsurface scattering objects.

Furthermore, we showed methods to consistently interpolate missing parts and to reduce noise in the measured data. We rendered the resulting models using the rendering framework of [Lensch et al. 2003b], although the models can also be rendered using other methods (see Section 3.3 for details). Since our acquisition method is capable of acquiring spatially varying objects, the final models can be rendered using arbitrary light source patterns. We validated our approach by acquiring three translucent objects which show strongly different behavior and also by showing side-by-side comparisons of photographs of the real objects and the rendered models under similar lighting conditions.

We introduced a hierarchical representation for our final models which is relatively small in size compared to the large amounts of data that the acquisition generates in the first place. Furthermore this hierarchical representation allows for an efficient (i.e. at interactive speeds) rendering of the final models.

However, there still remains some work for the future to be done. First of all, we would like to test our approach with some well-defined test targets and compare them to the resulting models. An acquisition planning algorithm could be employed in the acquisition process to maximize the surface coverage with a given amount of camera positions. We would also like to include the surface BRDF of the object in the final models in order to make them look even more realistic.

At the moment, our complete framework depends on the assumption that the target object scatters light diffusely. But in the future, we would like to be able to acquire objects where single scattering is the predominant behavior (e.g. honey, see also Section 2.2). This would greatly enlarge the class of measurable materials.

Since our final models are represented in a hierarchical fashion, the acquisition could be done in a hierarchical fashion too. This could greatly reduce the acquisition time and the amount of input data which has to be processed in order to obtain the final model.

There is also room for improvements on the engineering side. Technological advances could both speed up the acquisition and improve the quality at the same time. For instance a color high-dynamic range camera could reduce the acquisition time by two thirds. Color reproduction can be improved using lasers that keep their intensity more stable over time and of course by using an image sensor that provides more resolution and produces less noise.

Bibliography

[Blasi et al. 1993] BLASI, P., SAËC, B. L., AND SCHLICK, C. 1993. A Rendering Algorithm for Discrete Volume Density Objects. *Computer Graphics Forum 13*, 3, 201–210.

[Bouguet 2003] BOUGUET, J.-Y., 2003. Camera Calibration Toolbox for Matlab. See http://www.vision.caltech.edu/bouguetj.

[Carr et al. 2003] CARR, N. A., HALL, J. D., AND HART, J. C. 2003. GPU Algorithms for Radiosity and Subsurface Scattering. In *Graphics Hardware 2003*, 51–59.

[Chuang et al. 2000] CHUANG, Y.-Y., ZONGKER, D. E., HINDORFF, J., CURLESS, B., SALESIN, D. H., AND SZELISKI, R. 2000. Environment Matting Extensions: Towards Higher Accuracy and Real-time Capture. In *SIGGRAPH 2000*, 121–130.

[Debevec and Malik 1997] DEBEVEC, P., AND MALIK, J. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. In *SIGGRAPH 1997*, 369–378.

[Debevec et al. 2000] DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the Reflectance Field of a Human Face. In *SIGGRAPH 2000*, 145–156.

[Dorsey et al. 1999] DORSEY, J., EDELMAN, A., JENSEN, H. W., LEGAKIS, J., AND PEDERSEN, H. K. 1999. Modeling and Rendering of Weathered Stone. In *SIGGRAPH 1999*, 225–234.

[Faugeras 1993] FAUGERAS, O. 1993. *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, Cambridge, MA.

[Feng et al. 1993] FENG, S., ZENG, F., AND CHANCE, B. 1993. Monte Carlo Simulations of Photon Migration Path Distributions in Multiple Scattering Media. In *Photon Migration and Imaging in Random Media and Tissues, Proc. of SPIE Vol. 1888*, 78–89.

[Godin et al. 2001] GODIN, G., RIOUX, M., BERALDIN, J.-A., LEVOY, M., COURNOYER, L., AND BLAIS, F. 2001. An Assessment of Laser Range Measurement on Marble Surfaces. In *5th Conf. on Optical 3D Measurement Techniques*.

[Goesele et al. 2004] GOESELE, M., LENSCH, H. P. A., LANG, J., FUCHS, C., AND SEIDEL, H.-P. 2004. DISCO - Acquisition of Translucent Objects. In *SIGGRAPH 2004*.

[Goesele 2004] GOESELE, M. 2004. *New Acquisition Techniques for Real Objects and Light Sources in Computer Graphics*.

[Gortler et al. 1996] GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. 1996. The Lumigraph. In *Computer Graphics Proceedings (ACM SIGGRAPH 96)*, 43–54.

[Hanrahan and Krueger 1993] HANRAHAN, P., AND KRUEGER, W. 1993. Reflection from Layered Surfaces due to Subsurface Scattering. In *SIGGRAPH 1993*, 165–174.

[Hao et al. 2003] HAO, X., BABY, T., AND VARSHNEY, A. 2003. Interactive Subsurface Scattering for Translucent Meshes. In *Proc. of the 2003 Symposium on Interactive 3D Graphics*, 75–82.

[Heidrich 1999] HEIDRICH, W. 1999. *High-quality Shading and Lighting for Hardware-accelerated Rendering*.

[Ishimaru 1978] ISHIMARU, A. 1978. *Wave Propagation and Scattering in Random Media*. Academic Press.

[Jensen and Buhler 2002] JENSEN, H. W., AND BUHLER, J. 2002. A Rapid Hierarchical Rendering Technique for Translucent Materials. In *SIGGRAPH 2002*, 576–581.

[Jensen and Christensen 1998] JENSEN, H. W., AND CHRISTENSEN, P. H. 1998. Efficient Simulation of Light Transport in Scences with Participating Media using Photon Maps. In *SIGGRAPH 1998*, 311–320.

[Jensen et al. 1999] JENSEN, H. W., LEGAKIS, J., AND DORSEY, J. 1999. Rendering of Wet Materials. In *Rendering Workshop 1999*, 281–290.

[Jensen et al. 2001] JENSEN, H. W., MARSCHNER, S. R., LEVOY, M., AND HANRAHAN, P. 2001. A Practical Model for Subsurface Light Transport. In *SIGGRAPH 2001*, 511–518.

[Kang et al. 2003] KANG, S. B., UYTTENDAELE, M., WINDER, S., AND SZELISKI, R. 2003. High Dynamic Range Video. *ACM Trans. Graph. 22*, 3, 319–325.

[Lafortune and Willems 1996] LAFORTUNE, E. P. F., AND WILLEMS, Y. D. 1996. Rendering Participating Media with Bidirectional Path Tracing. In *Rendering Workshop 1996*, 91–100.

[Lensch et al. 2001] LENSCH, H., HEIDRICH, W., AND SEIDEL, H.-P. 2001. Silhouette-based Algorithm for Texture Registration and Stitching. *Graphical Models 63*, 4, 245–262.

[Lensch et al. 2003a] LENSCH, H. P. A., KAUTZ, J., GOESELE, M., HEIDRICH, W., AND SEIDEL, H.-P. 2003. Image-based Reconstruction of Spatial Appearance and Geometric Detail. *ACM Trans. Graph. 22*, 2, 234–257.

[Lensch et al. 2003b] LENSCH, H. P., GOESELE, M., BEKAERT, P., KAUTZ, J., MAGNOR, M. A., LANG, J., AND SEIDEL, H.-P. 2003. Interactive Rendering of Translucent Objects. *Computer Graphics Forum 22*, 2, 195–206.

[Lensch 2004] LENSCH, H. P. 2004. *Efficient, Image-Based Appearance Acquisition of Real-World Objects*.

[Levoy and Hanrahan 1996] LEVOY, M., AND HANRAHAN, P. 1996. Light Field Rendering. In *Computer Graphics Proceedings (ACM SIGGRAPH 96)*, 31–42.

[Levoy et al. 2000] LEVOY, M., PULLI, K., CURLESS, B., RUSINKIEWICZ, S., KOLLER, D., PEREIRA, L., GINZTON, M., ANDERSON, S., DAVIS, J., GINSBERG, J., SHADE, J., AND FULK, D. 2000. The Digital Michelangelo Project. In *SIGGRAPH 2000*, 131–144.

[Malzbender et al. 2001] MALZBENDER, T., GELB, D., AND WOLTERS, H. 2001. Polynomial Texture Maps. In *SIGGRAPH 2001*, 519–528.

[Marschner et al. 1999] MARSCHNER, S. R., WESTIN, S. H., LAFORTUNE, E. P. F., TORRANCE, K. E., AND GREENBERG, D. P. 1999. Image-based BRDF Measurement Including Human Skin. In *Rendering Workshop 1999*, 139–152.

[Masselus et al. 2003] MASSELUS, V., PEERS, P., DUTRÉ, P., AND WILLEMS, Y. D. 2003. Relighting with 4D Incident Light Fields. *ACM Trans. Graph. 22*, 3, 613–620.

[Matusik et al. 2002] MATUSIK, W., PFISTER, H., ZIEGLER, R., NGAN, A., AND MCMILLAN, L. 2002. Acquisition and Rendering of Transparent and Refractive Objects. In *Rendering Workshop 2002*, 267–278.

[Miller et al. 1998] MILLER, G. S. P., RUBIN, S., AND PONCELEON, D. 1998. Lazy Decompression of Surface Light Fields for Precomputed Global Illumination. In *Rendering Workshop 1998*, 281–292.

[Nicodemus et al. 1977] NICODEMUS, F. E., RICHMOND, J. C., HSIA, J. J., GINSBERG, I. W., AND LIMPERIS, T. 1977. Geometrical Considerations and Nomenclature for Reflectance. National Bureau of Standards.

[Omicron-Laserage] OMICRON-LASERAGE. Omicron-Laserage homepage. http://www.omicron-laserage.de.

[Pharr and Hanrahan 2000] PHARR, M., AND HANRAHAN, P. 2000. Monte Carlo Evaluation of Non-linear Scattering Equations for Subsurface Reflection. In *SIGGRAPH 2000*, 75–84.

[Press et al. 1994] PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., AND FLANNERY, B. P. 1994. *Numerical*

Recipes in C : The Art of Scientific Computing, 2nd ed. Cambridge Univ. Press.

[Robertson et al. 1999] ROBERTSON, M. A., BORMAN, S., AND STEVENSON, R. L. 1999. Dynamic Range Improvement Through Multiple Exposures. In *Proc. of the Int. Conf. on Image Processing (ICIP'99)*, 159–163.

[Rushmeier and Torrance 1990] RUSHMEIER, H. E., AND TORRANCE, K. E. 1990. Extending the Radiosity Method to Include Specularly Reflecting and Translucent Materials. *ACM Trans. Graph. 9*, 1, 1–27.

[Schlick 1994] SCHLICK, C. 1994. An Inexpensive BRDF Model for Physically-based Rendering. *Computer Graphics Forum 13*, 3, 233–246.

[Sillion 1995] SILLION, F. X. 1995. A Unified Hierarchical Algorithm for Global Illumination with Scattering Volumes and Object Clusters. *IEEE Trans. Visualization and Computer Graphics 1*, 3, 240–254.

[Sloan et al. 2003] SLOAN, P.-P., HALL, J., HART, J., AND SNYDER, J. 2003. Clustered Principal Components for Precomputed Radiance Transfer. *ACM Trans. Graph. 22*, 3, 382–391.

[Stam 1995] STAM, J. 1995. Multiple Scattering as a Diffusion Process. In *Rendering Workshop 1995*, 51–58.

[Stam 2001] STAM, J. 2001. An Illumination Model for a Skin Layer Bounded by Rough Surfaces. In *Rendering Workshop 2001*, 39–52.

[Wood et al. 2000] WOOD, D. N., AZUMA, D. I., ALDINGER, K., CURLESS, B., DUCHAMP, T., SALESIN, D. H., AND STUETZLE, W. 2000. Surface Light Fields for 3D Photography. In *SIGGRAPH 2000*, 287–296.

[Zongker et al. 1999] ZONGKER, D. E., WERNER, D. M., CURLESS, B., AND SALESIN, D. H. 1999. Environment Matting and Compositing. In *SIGGRAPH 1999*, 205–214.